Université Bordeaux Montaigne

École Doctorale Montaigne Humanités (ED 480)

THÈSE DE DOCTORAT EN « SCIENCES DE L'INFORMATION ET DE LA COMMUNICATION »

E-réputation :

Des modèles théoriques aux méthodes de mesure

Présentée et soutenue publiquement le 15 décembre 2021 par

Liliane BARA

Sous la direction de Valérie Carayol et Amar Lakel

Membres du jury

Yanita Andonova, Professeur en SIC, Université Sorbonne Paris Nord (Paris 13) - Rapporteur
Francine Charest, Professeur titulaire en communication, Université Laval, Québec, Canada - Rapporteur
Camille Alloing, Professeur en communication, Université du Québec à Montréal, Canada
Jean-Claude Domenget, Maître de conférences HDR en Sciences de l'Information et de la Communication, Université de Franche-Comté - Laboratoire ELLIADD - EA 466
Amar Lakel, Maître de conférences en Sciences de l'Information et de la Communication, Université Bordeaux Montaigne laboratoire MICA - Co directeur de thèse
Valérie Carayol, Professeur des Universités en Sciences de l'information et de la communication, Université Bordeaux Montaigne laboratoire MICA - Directrice de thèse

L'Université Bordeaux Montaigne n'entend donner aucune approbation ni improbation aux opinions émises dans les thèses. Celles-ci doivent être considérées comme propres à leurs auteurs.

« Les machines un jour pourront résoudre tous les problèmes, mais jamais aucune d'entre elles ne pourra en poser un »

Albert Einstein

À mon défunt père,

À ma mère,

À Yves, Myriam et Sandrine!

REMERCIEMENTS

Avant tout propos, je tiens à adresser mes sincères remerciements à ceux qui de près ou de loin ont contribué à la réussite de ce projet de vie, et sans qui ces travaux n'auraient sans doute pas pu voir le jour.

Mes remerciements vont tout d'abord aux membres du jury pour l'honneur qu'ils me font de porter un regard critique sur mon travail, et pour leurs recherches qui m'ont incitée à la réflexion et ont contribué à éclairer ma thèse. Merci à Camille Alloing pour sa générosité, ses conseils avisés lors de nos échanges et ses nombreuses réflexions sur l'objet de ma recherche. Je remercie Yanita Andonova pour son implication et ses recherches sur la communication des organisations et la question numérique, Francine Charest et Jean-Claude Domenget pour leur disponibilité et leurs écrits sur la construction et la gestion de l'e-réputation sur les plateformes socionumériques.

À mes directeurs de thèse : merci à Valérie Carayol pour la confiance placée en moi dès l'annonce de mon projet de recherche. Votre patience, votre générosité et votre disponibilité auront su m'aider à rebondir surtout dans les moments d'incertitude. À Amar Lakel qui m'a fait découvrir les nouveaux enjeux réputationnels et a contribué à faire germer l'idée de cette recherche, merci pour sa bienveillance et son soutien sans faille.

À tous ceux et celles rencontrés lors de ce parcours solitaire : aux membres de l'axe COS du MICA, aux doctorants et doctorantes avec qui j'ai partagé cette expérience, aux chercheurs, à l'équipe de la "Saint Glinglin", à mes étudiants, merci pour la bonne humeur et ces fructueux échanges. Une pensée particulièrement aux professionnels qui, malgré le contexte lors de nos entrevues, m'ont fait confiance, ont accepté de partager avec moi leurs expériences, et ont contribué à enrichir mon corpus de recherche.

Enfin, à mes proches, merci pour votre soutien indéfectible. Je remercie mes fidèles amis pour leur présence dans les moments parfois difficiles, et qui, malgré la distance pour certains, ont toujours été présents.

À ma mère, mon frère Yves et mes sœurs Sandrine et Myriam : merci pour votre amour inconditionnel, de toujours avoir su trouver les mots, et pour ce que vous êtes. À mon père, puisses-tu de là-haut, être fier… À vous tous, merci.

RÉSUMÉ

E-réputation : des modèles théoriques aux méthodes de mesure

Depuis l'apparition du terme "e-réputation" dans un article scientifique en 2000 (McDonald et Slawson), force est de constater que, l'étude de ce phénomène est encore à l'état embryonnaire, et de nombreuses pistes restent à être explorées, notamment sur les approches mobilisées afin de le mesurer, de l'évaluer. L'objectif principal de nos travaux de recherche est de déterminer si les professionnels de la réputation numérique s'appuient, peu ou prou, dans leurs offres de mesure, sur des modèles académiques ou conceptuels. Déterminer un éventuel rapport entre les deux approches suppose un questionnement à la fois empirique et théorique.

Sur le plan théorique, à partir d'une taxonomie des modèles conceptuels de la réputation, nous nous intéresserons aux dimensions structurelles et méthodologiques mises en œuvre dans le processus d'évaluation. Ces dimensions nous permettront d'identifier des catégories de modèles.

L'approche empirique vise, sur la base d'entrevues réalisées auprès de professionnels, à circonscrire le marché de l'e-réputation en termes d'acteurs, mais aussi et surtout de prestations de service. Ces dernières, et plus particulièrement les méthodes de mesure déployées, feront par la suite l'objet d'une mise en regard avec les modèles théoriques issus de la littérature scientifique essentiellement anglo-saxonne.

Mots clés : E-réputation, mesure, audit, réputation, modèles théoriques, confiance numérique, identité numérique, fouille d'opinion, analyse de sentiments, cartographie web, web social, pratiques professionnelles

ABSTRACT

Online reputation : from conceptual models to professional practices

Since the first appearance of the term “online reputation” in a scientific article by McDonald and Slawson in 2000, it is clear that, even if its use is very widespread today, the phenomenon is still in an embryonic state, and several avenues remain to be explored. In particular, the approaches used to assess, evaluate reputation. Our research aims to determine whether reputation professionals rely more or less on conceptual models in their measurement strategies. Determining a possible relationship between the two approaches supposes both empirical and theoretical questioning.

At a theoretical level, through a taxonomy of conceptual reputation models, we will focus on the structural and methodological dimensions implemented in the assessment process. This will allow us to determine categories to classify reputation models.
The empirical approach aims, on the basis of interviews with professionals, to define the reputation market map in terms of agencies but also and above all, in terms of services provided. Those ones, and more particularly the measurement methods used, will subsequently be the subject of a comparative analysis with theoretical models.

Key-words : Online reputation, reputation, assessment, audit, conceptual models, online trust, digital identity, opinion mining, sentiment analysis, data visualisation, social web, professional practices

SOMMAIRE

REMERCIEMENTS 3
RÉSUMÉ 5
ABSTRACT 6
SOMMAIRE 7
LISTE DES PRINCIPALES ABRÉVIATIONS 11
INTRODUCTION GÉNÉRALE 13
1. Rappel historique 17
1.1. L'émergence du world wide web ! 18
1.2. L'avènement du web social dit 2.0 19
1.3. Le web des données 21
2. Contexte de la recherche 24
3. Objectif de la recherche 27
4. Problématique 28
5. Hypothèses de travail 30
6. Présentation du plan de la thèse 32
PARTIE I : DE LA RÉPUTATION À L'E-RÉPUTATION 34
CHAPITRE 1 : LA RÉPUTATION, UN OBJET DES SCIENCES HUMAINES ET SOCIALES 36
Introduction du chapitre 1 37
1. La réputation : historique, étymologie et définition 38
2. La réputation comme ressource et capital : approches quantifiantes 41
3. La réputation comme construit social de l'évaluateur ou du récepteur 45
4. Les espaces publics de la réputation : le paradigme info communicationnel 48
5. La réputation : une définition stabilisée à travers les différentes disciplines 50
Conclusion du chapitre 1 57
CHAPITRE 2 : DE L' E-RÉPUTATION À L'INFLUENCE 58
Introduction du chapitre 2 59
1. La trace numérique comme source de l'e-réputation 62
2. La construction des identités numériques 67
3. La construction de la confiance numérique 72
3.1. La confiance : un concept abstrait, complexe et multidimensionnel… 72
3.2. La confiance comme concept composite : entre risque, certitude et délégation 74
3.3. Les interactions au coeur de la construction de la confiance 81
4. La mesure de l'influence numérique 85
4.1. La détection de profils d'influence 89
4.2. La classification des influenceurs selon leur typologie 92
4.3. L'attribution d'un score d'influence 93

Conclusion du chapitre 2 95
CHAPITRE 3 : DES MEDIAS SOCIAUX À LEUR ANALYSE 97
Introduction du chapitre 3 98
1. Le web 2.0 et les plateformes collaboratives : l'avènement du "consom'acteur" 100
2. L'émergence de la fouille d'opinion et du sentiment analysis dans l'analyse des réseaux sociaux (ARS) 104
2.1. Le paramétrage 105
2.2. La collecte et le traitement des données : préparation du corpus et archivage 107
2.3. La phase d'analyse : le dilemme de l'analyse de sentiment 110
3. La visualisation des données : le boom des cartographies dans l'analyse des réseaux sociaux 113
Conclusion du chapitre 3 116
Conclusion de la partie I 117
PARTIE II : DES MODÈLES THÉORIQUES AUX MÉTHODES DE MESURE 120
Introduction de la partie II 121
CHAPITRE 4 : LES MODÈLES CONCEPTUELS DE MESURE DE LA RÉPUTATION EN LIGNE, PROPOSITION D'UNE TAXONOMIE 122
Introduction du chapitre 4 123
1. Revue de la littérature : les approches scientifiques de classification 126
1.1. Les dimensions de classification de Sabater et Sierra, 2005 126
1.2. Les dimensions de classification de Koutrouli et Tsalgatidou 2006 / 2012 129
1.3. Les dimensions de classification de Marti and Garcia-Molina, 2006 130
1.4. Les dimensions de classification de Wang and Vassileva, 2007 132
1.5. Les dimensions de classification de Hoffman et al., 2009 133
1.6. Les dimensions de classification de Yuan et al., 2012 138
1.7. Les dimensions de classification de Vavilis et al., 2014 141
1.8. Les dimensions de classification de Hendrikx et al., 2015 145
1.9. Les dimensions de classification de Granatyr et al. 2015 147
2. La mesure de la réputation en théorie : approche contextuelle et méthodologique 152
2.1. Les contextes d'utilisation des modèles 153
2.2. La phase de collecte : des sources de données à leur catégorisation 156
2.2.a. Le sourcing : des sources et typologies de données collectées 156
2.2.b. La catégorisation : les métriques utilisées dans le classement 160
2.3. La phase d'analyse des données : des métriques de réputation à la méthode de calcul utilisée 162
2.3.a. Des métriques de réputation comme objet de mesure 162
2.3.b. Les méthodes de calcul utilisées 169
3. Proposition d'une nouvelle taxonomie : nos dimensions de classification 179
3.1. Les modèles primaires ou sommaires 179
3.2. Les modèles ad hoc 180

3.3. Les modèles sophistiqués 181
3.4. Les approches modulaires / génériques (flexibles) 182
Conclusion du chapitre 4 184
CHAPITRE 5 : LE MARCHÉ DE LA MESURE DE LA RÉPUTATION EN LIGNE 185
Introduction du chapitre 5 186
1. L'e-réputation : un marché en pleine mutation 187
1.1. Les techniques et méthodes de collecte des données (les données utilisées) : échantillon versus quantité exponentielle de données 187
1.2. Les techniques et méthodes de mesure : approche traditionnelle versus algorithmique.193
1.3. Des problématiques liées à la réputation : l'objet sur lequel porte la mesure (image versus totalité des attributs) 195
2. Les agences et le marché : un marché français en pleine expansion....196
3. Les services proposés : une offre foisonnante et sur mesure 207
3.1. L'analyse de la présence en ligne : entre intelligence économique, veille, audit et social listening 208
3.2. La gestion de la réputation : de la gestion de crise au recours légal 224
3.2.a. La modération de contenus sur les media sociaux 224
3.2.b. La suppression ou le flooding (noyage) des contenus négatifs dans les résultats de recherche 227
3.2.c. Le recours légal en cas de diffamation 229
3.3. La construction de la réputation : l'approche par popularité 232
Conclusion du chapitre 5 236
Conclusion de la partie II 237
PARTIE III : UNE ENQUÊTE AUPRÈS DES PROFESSIONNELS DE LA RÉPUTATION EN LIGNE....239
Introduction de la partie III 240
CHAPITRE 6 : LA MÉTHODOLOGIE DE TRAVAIL 241
Introduction du chapitre 6 242
1. De la nécessité d'une approche quantifiante : utilisation d'une grille de lecture....243
1.1. Le contexte d'utilisation ou l'objectif visé : la phase de cadrage 244
1.2. La structure (de la solution ou de la méthode) 245
1.3. La méthode de mesure utilisée 247
1.4. Les données de mesure....249
2. Une approche inductive par entretiens semi directifs 250
2.1. Déroulé des entretiens 253
2.2. Proposition de catégories d'analyse : du processus de réduction des données 255
Conclusion du chapitre 6 263
CHAPITRE 7 : LE TERRAIN ET LES RÉSULTATS....264
Introduction du chapitre 7 265

1. Échantillonnage : d'un processus d'analyse exploratoire du web à la constitution du corpus 266
1.1. Le cadrage de la recherche 266
1.2. La constitution du corpus 268
1.3. La phase de traitement : de la prise de contact à la collecte....269
2. Présentation des résultats et analyse des données 270
2.1. L'e-réputation, un objet intangible nécessitant une approche pluridisciplinaire....270
2.2. L'e-réputation comme transposition de la réputation via l'interface numérique?....277
2.3. L'e-réputation, une offre sur mesure 280
2.4. La mesure de l'e-réputation, une méthode en construction 284
2.5. Différents types de livrables....289
2.6. La mesure de l'e-réputation, un avenir mitigé?....291
3. Interprétation des résultats et discussion....300
3.1. Les modèles théoriques sur lesquels sont construites les méthodes de mesure : des questions de recherche à la problématique….....300
Conclusion du chapitre 7....307
Conclusion de la partie III....308
CONCLUSION GÉNÉRALE ET PERSPECTIVES....309
1. Quelles réponses avons-nous pu apporter à nos hypothèses ? 311
2. Quelles premières conclusions? 317
3. Quelles sont les limites de nos travaux et les perspectives futures? 321
BIBLIOGRAPHIE 324
ANNEXES 352
Annexe 1 : Tableau de classification des modèles théoriques de mesure de l'e-réputation...353
Annexe 2 : Le marché de l'e-réputation....364
Annexe 3 : La grille de lecture....375
Annexe 4 : Le guide d'entretien....380
Annexe 5 : Les 20 catégories d'unités de sens....384
LISTE DES FIGURES....385
LISTE DES TABLEAUX 387
LEXIQUE....388
TABLE DES MATIÈRES 395

LISTE DES PRINCIPALES ABRÉVIATIONS

ADA : Authoritative Node Detection

API : Application Programming Interface

ARM : Adaptative and Robust Model

ARS : Analyse des Réseaux Sociaux

ART : Agent Reputation and Trust

BDI : Belief Desire Intention

CNIL : Commission Nationale de l'Informatique et des Libertés

CORE : COllaborative REputation

CSS : Cascading Style Sheets

C2C : Customer to Customer

DA : Domain Authority

DHTML : Dynamic HyperText Markup Language

DR : Domain Rating

FR TRUST : Fuzzy Reputation based trust

GMB : Google My Business

HTML : HyperText Markup Language

HTTP : HyperText Transfer Protocol

IA : Intelligence Artificielle

IoT : Internet of Things

KOL : Key Opinion Leader

LAN : Local Area Network

LIAR : Liar Identification for Agent Reputation

MANET : Mobile Ad hoc Network

ML : Machine Learning

MAS : Multi Agent Systems

MORE : Merged Opinions Reputation Model

NLP : Natural Language Processing

PA : Page Authority

PATROL : comPrehensive reputAtion -based TRust mOdel

PBN : Private Blog Networks

P2P : Peer to peer

PREP : Probabilistic Reputation model

Q&A : Questions and Answers

RD : Referring Domains

REPAGE : Reputation and ImAGE

SCT : Social Comparison Theory

SEA : Search Engine Advertising

SEM : Search Engine Marketing

SEO : Search Engine Optimization

SERP : Search Engine Result Page

SIC : Sciences de l'Information et de la Communication

SL : Subjective Logique

SMM : Social Media Management

TAL : Traitement Automatique du Langage

TB : Total Backlink

TCP / IP : Transmission Control Protocol / Internet Protocol

TIC : Technologies de l'Information et de la Communication

TF / CF : Trust Flow / Citation Flow

TRAVOS : Trust and Reputation model for Agent-based Virtual OrganisationS

TREET : Trust and Reputation Experimentation and Evaluation Testbed

TR : TrustRank

UGC : User Generated Content

URI / L : Uniform Resource Identifier / Language

VANET : Vehicular Ad hoc Network

INTRODUCTION GÉNÉRALE

« ***2,8 millions d'euros*** », c'est la somme allouée par le gouvernement en 2021[1] afin de mettre en place une stratégie de gestion de sa réputation sur le web. En effet, comme chaque année, et selon son accord cadre dont l'objectif vise à « *fournir des outils d'écoute du web social (social listening) et de mesure d'impact social sur les réseaux sociaux afin de permettre au Gouvernement de détecter, analyser et mesurer les préoccupations et attentes des internautes vis-à-vis de son action en particulier et de l'actualité en général* »[2], le Service d'Information du Gouvernement (SIG) a ouvert un appel d'offres de recherche de prestataires de services en *social listening*, offre pourvue par trois des acteurs du marché de la réputation.

Lorsqu'on exécute la requête « *e-réputation* ou *e réputation* » dans les générateurs ou outils de planification de mots clés (qui utilisent les suggestions des outils d'auto complétion ou complétion automatique[3] des moteurs de recherche), on obtient plus d'une centaine de suggestions de mots clés qui y sont associés (204 pour la plateforme *keyword tool*[4], 190 pour *Keyword shitter*[5], 163 pour *Answer the public*[6]) :

---

[1] Appel d'offre de recherche de prestataires en e-réputation ouvert en décembre 2020 et clos le 26 mars 2021 remporté par trois acteurs du marché (Visibrain acteur principal avec 1,7 millions d'euros https://www.visibrain.com/fr/ , Newship à Dublin 900 000 euros https://www.newswhip.com et Talkwalker 200 000 https://www.talkwalker.com/fr). Accessible sur https://www.boamp.fr/avis/detail/21-47280/officiel

[2] Ibid

[3] Construits à partir d'algorithmes d'intelligence artificielle, ils utilisent la prédiction de textes afin de proposer une suite à une requête initiée en partie via par exemple l'analyse de celles précédentes

[4] Liste des mots clés associés à la requête e-réputation sur la plateforme keyword tool https://keywordtool.io/fr/search/keywords/google/2786605?category=web&keyword=e%20réputation&country=FR&language=fr#suggestions

[5] Liste des mots clés associés à la requête e-réputation sur la plateforme https://keywordshitter.com

[6] Liste des mots clés associés à la requête e-réputation sur la plateforme answer the public https://answerthepublic.com/reports/136ef7a3-1caf-4f70-8b3b-eb5ae10b1882



Figure n°1 : Nuage de mots clés créé à partir des suggestions sur la plateforme « nuage de mots »[7]

Si on s'intéresse de plus près aux suggestions, elles sont diverses et variées et relatives notamment entre autres à ces thématiques[8] : « *la définition du concept, son utilité, l'audit, son impact, les avis clients, l'e-réputation appliquée aux entreprises ou à une personne sous forme de personal branding pour ce dernier, les outils*[9] *et/ ou logiciels (gratuits), les agences, les articles, livres et cours traitants du concept, l'e-réputation et les avis Google, son analyse, les avantages qui y sont liés, l'atteinte à l'e-réputation, le community management, les influenceurs, l'identité numérique, les enjeux, les étapes de sa mise en oeuvre, les formations sur la question, l'historique, l'image, l'aspect légal (la jurisprudence, la CNIL, les avocats spécialisés sur la question, etc.), les indicateurs (KPI), la gestion, le marketing, les nouvelles solutions, le référencement, les médias sociaux (Google, Twitter, Facebook,*

[7] Plateforme qui permet la création de nuages de mots clés. Accessible sur https://nuagedemots.co . Il est important de souligner que les propositions sont les mêmes sur les trois plateformes et certains résultats sont répétés

[8] Nous ne présenterons pas l'ensemble des associations mais celles les plus récurrentes sur les trois plateformes. Sur la proposition en elle même elle se présente sous forme de requête avec le terme e-réputation suivi du mot ou expression utilisé

[9] Typologie d'outils non spécifiés dans les requêtes

Instagram, Youtube, blog), le nettoyage de la réputation, la notoriété, les risques, la relation client / fidélisation, les stratégies d'influence, Tripadvisor, la veille, les traces, etc. ».

L'ensemble de ces suggestions répond essentiellement aux questions : « *Quelle est sa définition ? Comment vérifier / surveiller son e-réputation ? Comment la protéger / soigner ? Comment la gérer ? Comment la construire / l'améliorer ? Pourquoi en avoir une?* »[10].
Dans l'absolu et *a priori,* on constate que le terme soulève de nombreuses interrogations auprès des internautes avec un volume de recherches mensuelles au mois de mars 2021 dépassant les 5 000 (5 400[11]) requêtes.

Au vu de ces premiers éléments, plusieurs constats peuvent d'emblée être faits : Si la question de la définition du concept est encore au coeur des interrogations comme nous pouvons le constater, l'ensemble des propositions visent à définir un cadre afin de circonscrire le concept, avec notamment des éléments de réponse sur : le cadrage contextuel (contexte théorique qui s'articule autour d'un ensemble de concepts : l'identité numérique, l'image, l'influence, la notion de traces) , l'approche méthodologique (qui *a priori* utilise des approches de communication, marketing avec un aspect légal, allant d'une stratégie de veille en e-réputation en passant par le nettoyage de sa réputation ou encore les stratégies de gestion de la relation client) et dans une vision plus large son origine, ce qui nous renvoie à internet et à l'essor du web social.

[10] Qu'on retrouve dans l'onglet "question" de Keyword tool

[11] Tendances issues de la plateforme Ubersuggest. https://app.neilpatel.com/fr/ubersuggest/overview?lang=fr&locId=2250&keyword=e+réputation

1. Rappel historique

Les origines de l'informatique électronique « *remontent au moins aux années quarante, lorsque des appareils furent inventés pour aider le calcul mathématique : d'où le nom de « computer?» («?calculateur »=« ordinateur?» en français), appellation donnée aux personnes qui réalisaient ces calculs à la main (de « compute» = «?calculer?») »*[12]. À l'origine, internet était un réseau militaire conçu en 1960 et destiné au département américain de la défense. En 1967, le projet de mise en réseau des ordinateurs ARPA (Advanced Research Projects Agency) donnera le réseau ARPANET (Advanced Research Projects Agency Network) connu aujourd'hui sous le nom d'internet et dont l'objectif de départ était, « dans un contexte de guerre froide, de créer un réseau de télécommunication militaire, de structure décentralisée et redondante, capable de continuer à fonctionner malgré des coupures de lignes ou la destruction de certains systèmes » (Ghernaouti et Dufour, 2017[13]). Une première connexion est réalisée en 1969 et on pouvait compter en 1971, une quinzaine de nœuds connectés aux Etats Unis. À partir de 1973, plusieurs chercheurs de l'ARPA se sont attelés à développer de nombreux protocoles (dont le TCP, l'IP) afin d'améliorer les connexions entre les ordinateurs, mais aussi et surtout entre les divers réseaux et notamment sur les questions de fiabilité et de contrôle des erreurs. En 1983, à partir de la combinaison de protocoles TCP/IP, ARPANET connaîtra une évolution et une expansion avec l'arrivée d'Ethernet, permettant de « *relier en local des ordinateurs entre eux à haute vitesse* » (Cerruzi, 2012).

En parallèle de ces évolutions, plusieurs réseaux académiques ont été développés dont BITNET en 1981 (avec pour objectif de relier les ordinateurs centraux d'IBM sur les campus universitaires, Ceruzzi) et Usenet en 1980 qui, à l'époque était déployé sous Linux, n'était pas connecté à internet, mais qui aura un énorme impact sur l'internet d'aujourd'hui. Le nom « internet » sous son appellation actuelle l'est devenu en 1983, avec l'ouverture d'internet aux entreprises commerciales avec lesquelles coopère la National Science Foundation (NSF),

---

[12] Paul E. Ceruzzi (2012) « A history of modern computing, 2nd edition ». Aux origines américaines de l'Internet : projets militaires, intérêts commerciaux, désirs de communauté. https://www-cairn-info.ezproxy.u-bordeaux-montaigne.fr/revue-le-temps-des-medias-2012-1-page-15.htm

[13] Ghernaouti S., Dufour A., Chapitre premier - *Des origines aux réalités de l'internet* (2017) https://www-cairn-info.ezproxy.u-bordeaux-montaigne.fr/internet--9782130795001-page-7.htm

à la suite de la décision du ministère de la Défense de séparer ARPANET en deux, dont un réseau pour utilisation interne. Internet qui, sous l'impulsion de la NSF était utilisé pour la recherche et dans le domaine universitaire au départ et reliait en 1990 environ 200 universités, sera la même année déployée à d'autres services gouvernementaux (la NASA, le ministère de l'Énergie, etc.). La même année, Internet s'est imposé au grand public grâce au *world wide web*.

1.1. L'émergence du world wide web !

Le *world wide web*, communément appelé web, est un système hypermédia fonctionnant sur Internet et permettant de consulter, via un navigateur spécifique, des pages web accessibles en ligne, c'est à dire connectées au réseau informatique internet (Thomas Chaimbault, 2007[14]). En 1989, Tim Berners Lee dans l'optique que des scientifiques travaillant dans les universités et les instituts du monde entier puissent échanger des informations instantanément, évoqua l'idée d'une « toile », où chacun pourrait « naviguer » de contenus en contenus. En 1990, est officiellement né le *World Wide Web* "toile mondiale", tant dans sa dimension que dans ses contenus (Ertzscheid, 2013) reposant sur trois technologies fondamentales : le protocole *HTTP* (*Hyper Text Transfer Protocol*) qui visait à faciliter l'échange de dossiers provenant d'une variété de sources, le langage *HTML* (*HyperText Markup Language*) qui est un sous-ensemble d'un langage de formatage et l'URI/L (*Uniform Resource Identifier / Language*) qui pouvait diriger un ordinateur vers toutes les localisations sur internet avec la même facilité.

Le web connaîtra une croissance exponentielle à partir de 1993 avec la création du premier logiciel de navigation web Mosaic, qui sera suivi en 1994 par Netscape Navigator et le système de standardisation des technologies *World Wide Web Consortium* (W3C). Pour Chaimbault, le web à l'origine comprenait des pages statiques reliées entre elles par des liens hypertextes rarement mises à jour, ce qu'il a nommé le web 1.0. Les sociétés fonctionnaient alors pour lui selon un modèle économique estimant que le web était fait de publications, non de participations, et que les annonceurs étaient les véritables acteurs moteurs. Entre 1996 et

[14] Proposition de définition du concept "world wide web" par Chaimbault accessible sur https://www.enssib.fr/bibliotheque-numerique/documents/2-web-2-0-l-avenir-du-web.pdf

2000, on assiste à une pléthore de services accessibles via internet : logiciels, portails, moteurs de recherches, etc., favorisant la naissance de ce qu'on a appelé les « dotcom » : une structure de pages fondée cette fois sur un web dynamique où des systèmes de gestion de contenus servaient des pages web créées à la volée à partir d'une base de données en constante évolution. Dès lors, on parle de langages de script et du *DHTML*[15], d'un « web 1.5 ». Le web qui se voit plus « dynamique » et moins statique, garde néanmoins sa visée centrée sur les utilisateurs.

1.2. L'avènement du web social dit 2.0

Le web 2.0 représente une tendance dans l'utilisation du *World Wide Web* qui favorise la créativité, le partage et l'interaction entre les utilisateurs. Le concept est apparu en octobre 2004 lors d'une conférence *brainstorming* entre la société O'Reilly Media (du nom de son fondateur Tim O'Reilly) et la société MediaLive International au cours de laquelle Dale Dougherty, membre d'O'Reilly Media, suggéra que, loin de s'être effondré suite à l'explosion de la bulle internet en 2001, le web n'avait jamais semblé aussi important et novateur (Thomas Chaimbault, 2007). Chaimbault note que la pléthore de sites et applications issues du web 2.0 utilise des améliorations technologiques, ergonomiques, sémantiques, un *business model* innovant et reposant surtout sur un renversement de la logique *top-down* du web initial : alors que ce dernier « descendait » vers l'usager pour lui proposer contenus et services, le web 2.0 mettait l'accent sur une nouvelle forme d'interactivité qui place l'usager au centre de l'internet et se veut plus social et collaboratif.

Dans le web social, une grande part de l'activité consiste à échanger des informations sur des sujets de conversation, qui peuvent être, mais ne sont pas toujours des documents au sens traditionnel du terme : partage de textes, des images, de la musique, des vidéos, mais aussi information autour de sujets : des personnes, des lieux, des organisations, des événements ou simplement des centres d'intérêt (Sajus et al., 2009[16]). Le web 2.0 est donc devenu un web de

---

[15] Dynamic HyperText Markup Language (méthode qui permet de rendre les pages web dynamiques)

[16] Sajus Bertrand, Cardon Dominique, Levrel Julien *et al.*, « *Web 2.0, et après ? Critique et prospective* », *Documentaliste-Sciences de l'Information*, 2009/1 (Vol. 46), p. 54-66. DOI : 10.3917/docsi.461.0054. URL : https://www.cairn.info/revue-documentaliste-sciences-de-l-information-2009-1-page-54.htm

mémoire pour Sajus, car avec lui, s'est répandue l'habitude d'exploiter des informations sans ressentir le besoin de les enregistrer pour les réutiliser. Ce qui est une révolution, car avec les plateformes de blogs ou les sites de partage de vidéos, c'est désormais la production et le stockage des documents par les internautes qui s'effectuent à distance. "*D'espace de communication et de création*" comme il le souligne, « *le web est ainsi devenu un espace de mémoire, abritant une somme de connaissances et d'expériences inaccessibles hors ligne* »[17].

O'Reilly définit sept principes qui caractérisent le web 2.0 et ses applications : le web en tant que plateforme, exploiter l'intelligence collective (implication des utilisateurs devenus "*prosommateurs : producteurs et consommateurs*" avec l'utilisation des outils collectifs Wikipedia, Digg par exemple), la puissance est dans les données (capacité de stockage et mise à disposition des données : exemple de Google), la fin des cycles des *releases* (évaluation en tant réel de l'efficacité de nouvelles fonctionnalités avec des tests bêta), des modèles de programmation légers (rapidité et simplicité d'utilisation des modèles de programmation, logiciels *open source*, etc.), le logiciel se libère du pc (sites *responsive design* adaptables sur tous supports : téléphone, consoles de jeux, etc.), l'expérience utilisateur enrichie (ergonomie des interfaces avec les langages (Ajax: par Jesse James Garrett, ou Dhtml).

Dans la même perspective, C. Alloing (2013) reprend l'idée de Rebillard (2007, p. 132) selon laquelle le web 2.0 loin d'être un simple dispositif technique, repose sur une « *configuration sociotechnique* » soit une « *modalité évolutive d'agencement social d'une technologie résultant des relations entre groupes sociaux engagés dans sa conception, son utilisation, et sa représentation, et (historiquement) structuré par ses modalités antérieures comme par les logiques macrosociales environnant son développement* ». Dispositifs socio techniques qui sont regroupés en cinq points par Proulx et al. (2010, pp. 15-16) : la capacité des utilisateurs à créer, « remixer » et partager des contenus[18] ; l'accessibilité et la grande facilité d'utilisation des plateformes (d'un point de vue de l'appropriation technique) ; un mode de collaboration

[17] Ibid

[18] Aspect participatif, contributif du web 2.0 avec l'internaute comme producteur de contenu et non simple consommateur

se fondant sur « la force des coopérations faibles » (Aguiton et Cardon, 2007)[19]; un modèle économique basé sur la transformation des contenus produits par les utilisateurs (et leurs données personnelles) en valeur économique, ce que les auteurs voient comme le mécanisme clé d'une forme de « capitalisme informationnel » et une pluralité de pratiques et d'usages des outils techniques (blogs, forums entre autres).

Ertzscheid (2013) introduit quant à lui la notion de "*world live web*" qu'il qualifie de web instantané donnant en temps réel l'état des dernières informations publiées, le service Google News et microcontenus (billets de blogs par exemple) étant perçus comme les pionniers du *world life web*. Au critère de pertinence utilisé par les moteurs auparavant, vient s'ajouter un autre tout aussi important, celui de la capacité à rendre compte en temps réel de l'évolution des documents publiés sur le web, et ce, quelle que soit leur « granularité ». Là où les moteurs de recherche classiques n'indexaient que des « pages » comme il le souligne, les moteurs de recherche de la blogosphère se mettent à indexer des « billets », comme autant de fragments documentaires « *sitôt publiés, sitôt indexés, sitôt accessibles* ».

1.3. Le web des données

Cinq ans après la publication de son texte "*What is Web 2.0*" (2005), Tim O'Reilly, aidé de John Battelle, s'est intéressé à son évolution[20]. Cette nouvelle dimension du web qui s'appuie en priorité sur les plateformes de réseaux socionumériques, s'intéresse à l'exploitation des données personnelles des internautes et étend la portée du web 2.0 au-delà de la frontière des ordinateurs et des utilisateurs pour lui trouver des domaines d'application dans le monde réel. C'est la maturation du web 2.0, avec l'intégration des métadonnées, du temps réel, de la réalité augmentée, des objets connectés, des informations universelles, structurées et open sources, du web sémantique.

Le fonctionnement des applications rattachées au web2 (géolocalisation, personnalisation, réalité augmentée) s'appuie en grande partie sur les données fournies par les utilisateurs de

---

[19] C'est-à-dire la collaboration entre individus sans intentionnalités fortes de former un collectif

[20] Dans sa publication « *Web Squared. Web 2.0 Five Years On* » (2009). https://conferences.oreilly.com/web2summit/web2009/public/schedule/detail/10194

manière volontaire ou non (Rebillard Franck, 2011), et ces pratiques communicationnelles débouchent sur des « systèmes coopératifs de données », des bases de données reliées entre elles et alimentées en permanence à l'aide des traces laissées par les internautes (O'Reilly et Battelle, 2009, p. 8). Ces traces, selon O'Reilly et Battelle, accumulées depuis l'époque du Web 2.0, ont rendu donc possibles les développements liés au Web2.

Dans la même perspective, O. Ertzscheid (2013), souligne que du "*world wide web*", on est passé par le "*world live web*" pour aujourd'hui en arriver au "*world life web*". Le "*world life web*" est apparu pour lui avec l'essor des "réseaux sociaux" tels que Facebook et Myspace et celui des mondes virtuels (*Second Life*). Ce « 3e âge documentaire » comme il le nomme, après l'adressage des documents et leur niveau de granularité (de plus en plus fin) pose la principale question « de la sociabilité et du caractère indexable, remixable de notre identité numérique et des traces qu'elle laisse sur le réseau ». Il distingue de ce fait, trois sphères d'où sont extraites les informations : la sphère publique (notre métier par exemple), la sphère privée (nos relations, nos amis), enfin et surtout la sphère intime (nos préférences politiques, sexuelles ou religieuses). Cette nouvelle ère du web, qualifiée de « web sémantique » ou d' « ère post-PC » par Ghernaouti et Dufour (2017), est apparue selon eux en 2007, et parler de web 3.0 sous-tend pour eux, prendre en compte l'internet mobile (avec les smartphones, les tablettes et leurs caractéristiques).

L'évolution du réseau internet est de fait marquée par cinq (5) grandes ères : la première, la genèse ou l'ère précédant l'arrivée du web de 1960 à 1990 ponctuée par trois principales phases. La première phase est relative à l'éclosion de la nébuleuse internet avec notamment le projet ARPANET et le développement des protocoles tels le TCP/IP, la deuxième à "l'impulsion de la micro informatique" avec l'émergence des PC, des réseaux locaux (LAN) et des technologies de télécommunications (Compuserve, Minitel, etc.) et la troisième, l'apparition du web avec notamment la création en 1991 des serveurs Gopher[21] permettant la

[21] Ghernaouti S., Dufour A., *Chapitre premier - Des origines aux réalités de l'internet* (2017) https://www-cairn-info.ezproxy.u-bordeaux-montaigne.fr/internet--9782130795001-page-7.htm . Même s'ils n'existent plus aujourd'hui, les serveurs Gopher ont été les prémisses de la création des liens hypertextes

mise à disposition de ressources sur divers serveurs[22].

La seconde ère est relative à l'arrivée du web des années 1995 à 2000 avec notamment l'entrée en bourse de la société Netscape en 1995. On assistera dès lors à l'émergence du commerce électronique, des investissements dans les sociétés liées à internet, des télécommunications, au financement d'infrastructures, etc. Cette ère est également marquée par la création d'entreprises de services[23] avec l'apparition des "dotcom"[24] et les crash boursiers dans les années 2000 (notamment le NASDAQ) d'où l'éclatement de la bulle internet, impliquant une restructuration du secteur, et affectant les fournisseurs de matériel informatique et de télécommunication.

La troisième ère est relative à l'éclosion des réseaux sociaux de 2000 à 2007 qui a marqué l'ancrage d'internet dans les usages avec l'arrivée des sites d'information, des moteurs de recherche, des blogs, des sites de e-commerce, etc. La même année, apparaissent les risques liés à ces développements tels que les spams, virus, etc.

Le développement des Technologies de l'Information et de la Communication (TIC), dans tous les domaines d'activité, marque un tournant majeur dans les échanges, les achats et la perception de la société en général en brouillant les frontières géographiques et temporelles et en mettant en exergue la proximité et l'instantanéité des échanges. L'apparition des plateformes de réseaux sociaux, d'échange, blogs, wikis, etc., à partir de 2003, et l'avènement du web 2.0 d'O'Reilly en 2005 propulsa le web vers une dimension coopérative, participative via notamment les contenus générés par les utilisateurs (*UGC*). Le passage de l'internet mobile des années 2007 à nos jours, qualifié de web sémantique et marqué par

---

[22] Cette phase est également marquée par l'apparition du world wide web de Tim Berners-Lee et Robert Cailliau en 1991 avec le développement des premiers logiciels clients et serveurs web, du premier logiciel de navigation Mosaic sous Unix, windows et Mac en 1993, l'expansion des usages du web avec les premières versions de navigateurs fonctionnels sur des ordinateurs personnels et développés par Netscape en 1994 et la création la même année du World Wide Web Consortium.

[23] Des logiciels, des moteurs de recherche, des portails, des sites d'information, des magazines électroniques, du commerce en ligne, etc., de 1996 à 2000

[24] Expression utilisée par Ghernaouti & Dufour (2017) dans leur chapitre 1er « *Des origines aux réalités de l'internet* » pour désigner les entreprises utilisant le suffixe ".com" dans leurs adresses web

« l'internet mobile »[25], utilise les smartphones, tablettes et phablettes, et a été rendu possible avec notamment l'arrivée des smartphones et leurs diverses fonctionnalités dont l'iPhone en 2007, les tablettes numériques vers 2010 et des progrès dans les débits (2G, 3G, 4G, etc.).

Enfin, la période actuelle est marquée par la transformation numérique, l'internet des objets (IoT) et l'Intelligence Artificielle (IA)[26].

2. Contexte de la recherche

Ces cinq principales ères ont vu, avec notamment le développement des médias sociaux et les contenus générés par les internautes (*UGC*), apparaître une nouvelle typologie d'internautes qui ont pris la parole pour parler d'eux, mais aussi et surtout des autres. Personne n'est désormais à l'abri d'un avis (positif comme négatif) sur ses actions ou sa personne. L'internaute, qui dès lors semble adopter la posture « d'influenceur » actif et de relais d'information, est susceptible de concourir au développement et au maintien ou non de la réputation des individus comme des collectifs et des organisations.

On attribue l'apparition du terme e-réputation en 2000 dans un article scientifique de McDonald et Slawson « *Reputation in an internet auction market » (*et dont la dernière version a été mise à jour en 2007)[27]. À partir de là, plusieurs événements, dont le salon Erepday[28] (dont le premier date de 2011 et qui traite des stratégies de gestion de sa réputation en ligne, de branding et de la relation client 2.0), ou encore le premier colloque français organisé en 2013[29] sur l'e-réputation et les traces numériques ont vu le jour. D'un point de

[25] Accessible sur la plateforme web www.w3.org

[26] La première caractéristique de cette ère est relative à l'avènement de la réalité augmentée et virtuelle avec notamment les *Google glasses*, Google VR, Microsoft applicables à une pléthore de domaines : la domotique, l'industrie avec les robots, les technologies embarquées (*wearables*), les *smart cities*, etc. La troisième englobe les intelligences artificielles avec notamment les assistants vocaux, les systèmes de recommandation, la prédiction, etc. Enfin, la dernière caractéristique liée à la bio informatique qui est plus spécifiquement rattachée à "*l'élaboration de systèmes informatiques capables de s'intégrer physiquement au corps humain*" tels les prothèses par exemple (Ghernaouti et Dufour, 2017).

[27] Premier article de McDonald et Slawson sur l'e-réputation accessible via https://onlinelibrary.wiley.com/doi/abs/10.1093/ei/40.4.633

[28] Page web du salon Erepday https://www.erepday.fr

[29] Page web du premier colloque français sur l'e-réputation http://e-reputation-lecolloque.org

vue académique, on peut citer les différentes thèses dont la première en France débutée par N. Chazaud en 2005 et enfin, plusieurs auteurs et chercheurs de la littérature scientifique (Chun, 2005 ; Davis, 2015 ; Ertzscheid, 2013 ; Merzeau, 2012 ; Georges, 2009 ; Boullier et Lohard, 2015 ; Severo, 2017) qui se sont intéressés à l'objet e-réputation dans sa globalité : l'identité numérique, les traces numériques, l'*opinion mining* et le *sentiment analysis*, la visualisation de données, la *corporate reputation*, le management de la réputation, etc.

On retrouve également des *e-books,* guides, études de cas et livres blancs sur les risques et enjeux liés à la réputation, les typologies de menaces auxquelles peuvent être exposées les organisations, les bonnes pratiques en termes de gestion de sa réputation en ligne, les techniques de veille, etc., édités pour la plupart par les organisations spécialisées dans la mesure de la réputation : Digimind[30], Synthesio[31], Meltwater[32], Brandwatch,[33] Visibrain[34], AT internet[35], etc. La littérature scientifique anglo-saxonne quant à elle, en sus d'écrits sur le management de la réputation, la *corporate reputation*, les systèmes de réputation, etc., a produit une myriade d'articles sur les modèles théoriques de mesure de la réputation et de la confiance en ligne (nous en proposons une taxonomie et une analyse dans la partie II). En effet, plusieurs chercheurs anglo-saxons pour la plupart, Yu et Singh (2000) ; Dellarocas (2003) ; Zacharia et Maes (1999) ; Sabater et Sierra (2001, 2005) ; Hoffmann et al. (2009) ; Hendrikx et al. (2015) ; Granatyr et al. (2015), etc., se sont intéressés à la suite de McDonald

---

[30] Exemples de ressources sur les questions liées à l'e-réputation proposées par Digimind sur leur site web
https://www.digimind.com/fr/ressources
https://www.leslivresblancs.fr/livre/internet-ntic/e-reputation/20-bonnes-pratiques-1-en-e-reputation
https://www.leslivresblancs.fr/livre/internet-ntic/e-reputation/e-reputation-ecoutez-et-analysez-le-buzz-digital

[31] Exemples de ressources sur les questions liées à l'e-réputation proposées par Synthesio sur leur site web
https://www.synthesio.com/fr/resources/

[32] Exemples de ressources sur les questions liées à l'e-réputation proposées par Meltwater sur leur site web
https://www.meltwater.com/fr/insights/

[33] Exemples de ressources sur les questions liées à l'e-réputation proposées par Brandwatch sur leur site web
https://www.brandwatch.com/fr/guides/

[34] Exemples de ressources sur les questions liées à l'e-réputation proposées par Visibrain sur leur site web
https://www.visibrain.com/fr/ressources/livres-blancs/

[35] Exemples de ressources sur les questions liées à l'e-réputation proposées par AT Internet sur leur site web
https://www.atinternet.com/ressources/

et Slawson, à la problématique de la fabrique et de la mesure de la réputation en ligne en proposant des modèles théoriques.

Au sens de l'individu, la mesure de la réputation peut prendre la forme du "*personal branding*" ou marketing de soi (Montoya et Vandehey, 2002 ; Chritton, 2012 ; Reynolds, 2013 ; Kapidzic, 2018) et va de la simple "*googlelisation*"[36] aux techniques de mise en visibilité et stratégies d'influence. La gestion de la réputation des organisations (*online reputation management*) elle, revêt une forme plurielle, avec des pratiques et méthodes qui utilisent de fait une approche pluridimensionnelle.

La notion « d'e-réputation », même si elle est très largement utilisée aujourd'hui, semble être une notion plurielle et polymorphe (comme nous le verrons tout au long de notre recherche) et ne fait pas encore totalement l'unanimité auprès de ses praticiens et utilisateurs. Force est de constater qu'assez peu de travaux en français ont été réalisés sur la question. Qu'elle soit perçue comme image, identité, relative aux traces numériques, ou comme transposition de la réputation au numérique, plusieurs auteurs (Georges, 2009 ; Babkine et al, 2011 ; Guigou, 2012 ; Rosoor, 2012 ; Camille Alloing, 2016 ; Severo et al, 2017 ; Galinon-Melenec, 2019, etc.) s'accordent sur le fait qu'elle concerne autant les individus que les entités morales, et qu'il faudrait de fait y prêter une attention particulière. Elle connaît un essor avec la multiplicité des plateformes numériques et des interactions entre les internautes, plateformes qui, du reste, sont la résultante de l'essor du réseau informatique internet et du web. Surveiller sa réputation en ligne (*online)* devient alors un enjeu majeur tant pour les individus que les organisations, et les méthodes et outils de mesure sur les réseaux numériques et le web se multiplient (Alloing, 2015 ; Boulier et Lohard, 2015) autant que les prestataires de service.

Même si, nous le verrons, la littérature anglo-saxonne est foisonnante sur les modèles théoriques qui servent à mesurer la réputation, peu a été dit et fait sur le rapport entre ces

---

[36] Phénomène qui vient du terme googliser défini par le larousse en ligne comme le fait de «*rechercher des informations (en particulier sur quelqu'un) sur internet en utilisant le moteur de recherche Google*», accessible sur https://www.larousse.fr/dictionnaires/francais/googliser/10910928
Certains auteurs (Lazuly, 2003 ; Lockergnome, 2005 ; Pinte, 2005) soulignent le côté "néfaste" de la googlelisation, les sources n'étant pas toujours qualifiées et adaptées à chaque domaine

derniers (les modèles) et les méthodes actuelles de mesure et d'audit d'e-réputation proposées par les professionnels. Cette carence nous a incitée à mener des recherches sur la question dans le domaine des Sciences de l'Information et de la Communication (SIC) et plus particulièrement dans le champ de la communication organisationnelle, domaine au sein duquel les questions de réputation et d'image sont l'objet d'études. Nous n'avons de fait pas pour objectif d'effectuer un énième recensement des bonnes pratiques en termes de gestion de la réputation ni de faire l'apanage de son importance, mais plutôt de questionner un éventuel rapport entre savoirs académiques et pratiques des professionnels.

3. Objectif de la recherche

Notre travail, axé sur l'**e-réputation,** traite plus particulièrement des modèles théoriques qui ont inspiré les méthodes d'audit proposées par les professionnels de l'e-réputation, à travers leur offre.

La notion d'objectif, qui ici désigne un but, une finalité, est le moteur de notre recherche qui jalonnera la stratégie à déployer tout au long de notre travail: formulation de notre problématique, choix des hypothèses de travail et cadrage méthodologique.

Notre travail de recherche a pour objectif principal d'**analyser les modèles théoriques d'e-réputation existants dans la littérature scientifique et les méthodes d'audit comprises dans l'offre proposée par les professionnels de l'e-réputation afin de questionner un éventuel rapport entre ces derniers.**

Pour y parvenir, nous nous proposons de :

- Démontrer comment la notion de réputation se transforme dans le passage à l'e-réputation et comment les sciences de l'information et de la communication peuvent la conceptualiser en prenant en compte ces métamorphoses;
- Définir l'audit d'e-réputation afin d'identifier les différentes méthodes proposées par les professionnels dans leur (s) offre (s);

- Expliquer en quoi l'usage de nouveaux outils numériques a permis de nouveaux types de mesures et pratiques, transformant de fait les pratiques d'audit;
- Identifier, dans la littérature scientifique, les modèles théoriques de mesure de la réputation et en analyser les caractéristiques;
- Identifier dans le champ professionnel les principales caractéristiques des audits proposés et leur relation avec les modèles académiques recensés.

Ces différentes analyses nous permettront de répondre à notre objectif central.

4. Problématique

Notre recherche est basée sur une approche inductive. Dans cette partie, nous exposerons et détaillerons d'une part la problématique centrale et les sous-questions permettant de l'étayer, ce qui nous permettra d'avancer nos hypothèses de travail.

Problématique principale : les professionnels de l'e-réputation ont-ils recours dans leurs stratégies d'audit à des savoirs académiques? Existe-t-il un rapport entre les modèles théoriques et les pratiques professionnelles, dans la mesure de la réputation en ligne?

Questionner le rapport entre savoirs académiques et pratiques professionnelles dans la mise en place d'une stratégie d'audit en e-réputation sous-entend de prendre en compte plusieurs éléments:

- Les caractéristiques et composantes des modèles existants dans la littérature scientifique;
- Les méthodes et stratégies d'audit en e-réputation proposées sur le marché par les professionnels, ce qui implique *de facto* une analyse du statut de ces professionnels et de l'offre globale qu'ils proposent;
- Le rapport entre les différents modèles théoriques et les pratiques professionnelles.

Afin d'étayer notre problématique principale, il s'avère intéressant de s'attarder sur quelques questions subsidiaires :

Question 1 : Dans quelle mesure internet et le web social ont-ils favorisé la mise en place et le développement de nouveaux types de mesures et pratiques?
Cette problématique met en exergue la nécessité d'observer les mutations de l'audit à l'ère du numérique. Il s'agit ici d'analyser les nouveaux outils numériques existants sur le marché et les métamorphoses entraînées par ces derniers sur les méthodes d'audit préexistantes.

Question 2 : Quelle perception ont les praticiens du concept "d'e-réputation"?
Il s'agit ici de questionner le concept clé de notre travail de recherche, "l'e-réputation", par le prisme de ses praticiens. Cette question implique une prise en compte du point de vue de différents auteurs (issus à la fois de la littérature scientifique et professionnelle) tant dans leur définition, leur emploi de la notion que ses différents champs d'application.

Question 3 : Sous quelle (s) forme (s) se déclinent les stratégies d'audit en e-réputation proposées par les professionnels, et quelle (s) méthode (s) utilisent-ils?
Afin de répondre à cette problématique, nous effectuerons une analyse comparative des pratiques des professionnels dans la mesure de la réputation. Nous utiliserons pour ce faire, une première approche notamment l'analyse de l'offre que proposent ces derniers à travers leurs plateformes numériques (site web entre autres) et pour conforter les informations recueillies, des entretiens semi-directifs seront réalisés auprès d'agences spécialisées en e-réputation.
À partir des différentes questions formulées ci-dessus, nous proposons quelques hypothèses de travail.

5. Hypothèses de travail

Hypothèse 1 : Cette hypothèse se base sur les différentes significations et déclinaisons accolées à la notion qui à la fois est perçue dans certains contextes comme relevant de l'image, de la notoriété, ayant attrait à l'honneur ou comme résultante d'une opinion commune, etc. On peut postuler que l'e-réputation est la résultante d'une évaluation d'un « objet », qu'il s'agisse d'une personne au sens d'individualité ou d'organisation ou encore d'un produit ou service. L'intérêt dont elle fait l'objet, en tant que bien immatériel désormais monnayable économiquement, expliquerait l'engouement des agences à en proposer des mesures.

Hypothèse 2 : Avec l'essor du web social, la production et la diffusion exponentielle de "données" semblent avoir favorisé le développement de nouvelles méthodes et pratiques de la part des professionnels de la réputation, tant dans la collecte que dans l'analyse des informations collectées.

Du web statique, on est passé au web social ou 2.0, pour aujourd'hui, arriver au web des données. Ce passage d'un web statique au web sémantique a entraîné des mutations dans les interactions qu'entretiennent les individus entre eux et avec les organisations sur la toile, impactant *de facto* les pratiques des professionnels de la réputation qui redoublent d'ingéniosité dans leurs pratiques. Cette hypothèse s'appuie donc sur les nouveaux outils, les nouvelles pratiques des praticiens dans la collecte des données qui seront par la suite utilisées dans leurs stratégies d'audit. On peut supposer une forme d'opportunisme dans l'utilisation des métriques rendues possibles par les technologies et la recherche d'un effet de « scientificité » liées à l'utilisation de mesures quantitatives nombreuses.

Hypothèse 3 : Dans la littérature scientifique (surtout anglo-saxonne), il existe une myriade de modèles théoriques de mesure de la réputation et on peut supposer que ces derniers ont évolué à travers le temps tant dans leurs caractéristiques que leurs applications.
Il semble exister un lien étroit entre réputation et confiance, les deux étant dans certains cas assimilés avec pour visée, la réduction du facteur risque dans l'établissement d'une relation. On peut supposer que les définitions et dynamiques de la « confiance » dans des environnements économiques et sociopolitiques différents peuvent expliquer des variations dans les modèles théoriques existants.

Hypothèse 4 : L'offre proposée par les professionnels de la réputation et qui s'articule autour du maître mot "le sur mesure", semble différer d'un praticien à l'autre et cela, selon plusieurs dimensions : leur statut, les besoins et attentes du marché au sein duquel ils évoluent et notamment ceux des commanditaires, l'offre proposée par les concurrents, ce qui nous amène à questionner l'existence d'une éventuelle méthode "universelle" de mesure et d'évaluation de l'e-réputation.
Notre dernière hypothèse sur la stratégie d'audit d'e-réputation que proposent les professionnels nécessite de s'intéresser tant au statut de ces professionnels, qu'à leur offre globale, facteurs susceptibles d'influer sur les résultats issus de la mesure de la réputation. On peut supposer une mobilisation différenciée et stratégique des ressources théoriques, technologiques et d'accès aux données des acteurs de terrains dans la recherche d'un modèle économique commercialisable, plus qu'une recherche de validité des résultats, au regard des mesures convoquées dans la mesure de l'e-réputation.

6. Présentation du plan de la thèse

Notre recherche, qui a pour objet l'e-réputation et plus particulièrement ***les modèles théoriques qui ont inspiré les méthodes d'audit proposées par les professionnels de la réputation***, se présente comme suit:
La première partie de notre travail s'articule autour de l'état de l'art sur les notions de réputation et e-réputation et celles connexes.

Parler d'e-réputation revient dans une première approche à s'intéresser au concept de réputation (au sens traditionnel). À travers la littérature, plusieurs auteurs ont déjà proposé un état de l'art de la notion de réputation telle que perçue à travers les différents champs disciplinaires. Notre travail (la partie I) visera plutôt à appréhender l'objet réputation comme "mesure" ou "évaluation" et plus précisément à identifier dans chaque cas de figure, l'objet de mesure. Nous reviendrons sur un bref historique de la notion avant de proposer une analyse des différents paradigmes avec toujours pour objectif la mise en exergue de l'objet de mesure dans chaque cas, pour, à la fin, proposer une analyse critique des différents travaux.

Nous proposerons par la suite (chapitre 2), une revue de l'objet e-réputation. Après un bref retour sur l'historique du concept, nous nous intéresserons aux notions connexes dans le passage de la réputation à l'e-réputation et plus particulièrement au concept d'influence. L'influence est souvent associée à une collecte de traces numériques et à la construction d'une forme d'identité numérique avec pour finalité clé la mise en place d'une relation de confiance dont nous analyserons la construction.

Enfin, nous analyserons dans le chapitre 3, à partir des médias et réseaux socionumériques (plateformes collaboratives au sein desquelles peut se construire la réputation *a priori*), les différents procédés utilisés dans la mesure de la réputation en ligne, modalités d'évaluation et métriques nous permettant d'approfondir nos hypothèses 1 et 2.

Dans la seconde partie, nous aborderons d'une part la mesure de la réputation telle que conçue par les différents théoriciens de la littérature scientifique anglo-saxonne et notamment les différents modèles théoriques ayant inspiré les méthodes actuelles d'audit. Nous ferons par la suite une analyse critique des modèles et de leurs limites, afin de proposer, une nouvelle taxonomie de classification.

Le chapitre 5 présentera une analyse comparative du marché de la mesure de la réputation qui, nous le verrons, est un marché foisonnant qui connaît un essor depuis les années 2009. Nous proposerons une taxonomie des prestataires qui composent le marché français[37] de la réputation, pour ensuite analyser les différents services proposés par ces derniers (veille, SEO/SEM, webmarketing, RP 2.0, déréférencement, flooding par exemple). Un des points clés de cette partie sera la proposition d'une analyse des méthodes d'audit existantes sur le marché, nous permettant d'approfondir notre hypothèse 1.

Enfin, et dans la troisième partie, nous présenterons l'analyse menée auprès d'une sélection d'entreprises françaises de mesure de la réputation qui vise à identifier et à affiner notre réflexion sur la mobilisation de modèles théoriques de mesure de la réputation dans les pratiques professionnelles. Cette partie devrait nous permettre de répondre à l'hypothèse 4.

[37] Nous avons choisi de nous intéresser aux entreprises françaises qui évoluent dans le domaine de la réputation ou à celles internationales qui possèdent un siège en France pour des raisons de faisabilité et d' accessibilité

PARTIE I : DE LA RÉPUTATION À L'E-RÉPUTATION

Introduction de la partie I

Dans cette première partie, et afin d'appréhender le concept d' e-réputation dans sa globalité, nous nous intéresserons plus particulièrement aux principales notions connexes qui gravitent autour du concept. Cette partie se présente comme suit:

L'un des points clés sur lequel nous nous attarderons est le concept de réputation. Après une introduction sur la notion de réputation, nous reviendrons d'abord sur l'historique du concept qui nous servira de cadre de départ afin d'appréhender les définitions du concept et de proposer une revue de la littérature scientifique. Cette revue ne vise pas à proposer les différentes définitions accolées à la notion de réputation à travers les différents champs disciplinaires, mais à questionner l'objet de mesure pris en compte dans chacune des conceptions du concept.

Dans un second temps, nous proposerons une analyse du concept clé de la thèse, l'e-réputation. La première partie sera consacrée à un retour historique sur la notion, ce qui nous amènera à questionner les différents concepts clés accolés à l'objet e-réputation : traces, identité, confiance et influence et une éventuelle transposition de notions dans le passage de la réputation à l'e-réputation.

Enfin, le troisième chapitre sera consacré à la place des médias et réseaux socionumériques dans l'essor de l'e-réputation, avec un focus sur l'analyse des réseaux sociaux.

CHAPITRE 1 : LA RÉPUTATION, UN OBJET DES SCIENCES HUMAINES ET SOCIALES

Introduction du chapitre 1

Parler d'e-réputation revient dans une première approche, à s'intéresser à l'objet réputation qui a été étudié tant en économie, sociologie, psychologie, sciences de la gestion, philosophie qu'en SIC. À travers ces différents champs, la réputation semble revêtir une multitude de composantes et significations.

Dans ce chapitre, nous ne souhaitons pas proposer une énième revue de littérature de la réputation (tant dans sa définition que de ses champs d'application) perçue à travers ces différentes disciplines, mais plutôt effectuer une analyse comparative de la notion de réputation perçue comme "évaluation" ou mesure avec les différents éléments pris en compte dans le processus d'évaluation. Pour y parvenir, nous nous proposons d'analyser dans un premier temps les différentes composantes ou déclinaisons du concept de réputation, ce qui nous permettra par la suite d'effectuer une analyse comparative de celles-ci (les composantes) pour enfin, identifier, pour chacune, l'objet sur lequel se base la mesure.

Après un retour historique à travers les champs disciplinaires, nous proposerons une analyse des différentes perceptions qui sont accolées au concept. Ce premier cadre nous permettra par la suite de proposer une analyse comparative du concept de réputation perçu comme évaluation, et d'appréhender notamment "l'objet" sur lequel porte la mesure de la réputation.

1. La réputation : historique, étymologie et définition

Emprunté au latin *reputatio*, le terme réputation vient de "puto qui signifie compter, estimer, considérer et du préfixe re- qui marque la répétition", Origgi (2007). Elle s'apparente de fait à une estimation constante de la valeur d'une personne ou d'un objet ; à une forme de considération commune ou de jugement basé sur un "examen" des actions de cet agent. Étymologiquement, le terme réputation revêt plusieurs significations et usages à travers l'histoire et le temps, et le CNRTL nous en propose quelques-unes : on parlera dans les années 1370 de la réputation perçue comme « *célébrité d'une personne ou d'une chose en raison de sa valeur* »[38] ou encore « *renommée*[39]*, bonne opinion que le public a de quelqu'un, de quelque chose* »[40]. Dès lors, on assistera à la création d'une pléthore d'expressions:

- "*jouer sa réputation*" ou "jouer avec sa réputation" en 1544 avec la réputation ici perçue comme une valeur sacrée : l'honneur;
- "*être en réputation de...*" dans les années 1650 qui signifie "*être fortement connu pour*…" s'apparentant ici à la renommée;
- "*ne connaître que de réputation*" en 1671 qui évoluera en 1784 pour donner l'expression courante actuelle "*connaître quelqu'un de réputation*" relatif notamment à l'ouï-dire, à ce qui se dit par autrui;
- "*avoir une réputation de...*" en 1786 pour qualifier le fait *d'être connu pour une chose ou un attribut singulier*;
- "vivre sur sa réputation" dans les années 1804 relatif au fait de *tirer partie (bénéficier) de l'opinion favorable donnée de soi à autrui à un moment donné;*
- *perdre quelqu'un de réputation* qui renvoie à ternir la réputation ou porter atteinte à la réputation de quelqu'un ou quelque chose.

Ce retour historique pose d'emblée le cadre de l'objet réputation et fait ressortir plusieurs significations: la réputation d'abord perçue comme estimation constante et qui, de fait, fait référence à l'action de compter, la réputation comme relevant du ouï-dire et donc mettant en

---

[38] Définition du CNRTL, accessible sur http://www.cnrtl.fr/etymologie/réputation

[39] la renommée est ici perçue au sens de la moralité

[40] Définition du CNRTL, accessible sur http://www.cnrtl.fr/etymologie/réputation

exergue le fait que la réputation d'un objet émane d'un autre objet, et enfin la réputation comme jugement basé sur les attributs, les faits et gestes d'un objet. Afin d'affiner la perception du concept, nous proposons une analyse de plusieurs définitions.

Fombrun (1997) fait partie des premiers à proposer une définition de la notion de réputation dans le champ économique qu'il perçoit comme « *A perceptual representation of a company's past actions and future prospects that describes the firm's overall appeal to all of its key constituents when compared with other leading revals* »[41]. La réputation qui ici est perçue comme « un outil de classement » d'une organisation par rapport à ses concurrents et donc qualifiée de *corporate reputation*, n'est pas maîtrisée par celle-ci, et est issue de l'ensemble des perceptions des actions passées et projections futures de l'organisation de la part de tous ses « publics ».

À partir de cette première définition, Boistel (2014) proposera l'une des définitions les plus reprises de l'objet réputation comme « *l'opinion bonne ou mauvaise que se font les différents publics de l'entreprise à partir du résultat de toutes leurs expériences, croyances, sentiments, connaissances et impressions qui constituent alors un tout perçu, construit à partir d'un ensemble d'items véhiculés par l'entreprise donnant une réputation forte si ceux-ci sont perçus de façon identique par les différents publics de l'entreprise et s'ils sont admis par les salariés* »[42]. Comme il le souligne lui-même, la définition soulève certaines interrogations relatives notamment au nombre de publics à prendre en compte dans la mesure de la réputation, mais aussi et surtout à la valeur des items positifs et négatifs et leur équivalence.

Cette notion de réputation comme *opinion* se retrouve également chez D'Almeida (2007) pour qui la réputation qui est *une opinion* (au sens d'opinion publique) est issue de l'interaction sociale entre plusieurs entités, et vise un « jugement », une évaluation qu'elle soit positive ou négative. Cette perception de la réputation comme évaluation se retrouve également chez plusieurs auteurs.

---

[41] Fombrun, Cj, et Cbm van Riel. « *The Reputational Landscape* ». *Corporate Reputation Review* 1, nº 2 (juillet 1997): 5-13. https://doi.org/10.1057/palgrave.crr.1540024. Traduction par nos soins " la réputation est une représentation en termes de perception des actions passées d'une entreprise et la prédiction de celles futures, qui permettent de définir, et ce à partir de tous ces constituants clés, l'attrait d'une entreprise par rapport à ses principaux concurrents".

[42] Boistel, Philippe. « Réputation : un concept à définir ». *Communication et organisation*, nº 46 (1 décembre 2014): 211-24. https://doi.org/10.4000/communicationorganisation.4796.

C'est notamment le cas pour Camille Alloing (2015) pour qui cette forme d'évaluation, qui peut être *individuelle ou collective*, est issue des informations collectées[43] à partir de plusieurs sources[44] et qui, par la suite, passent à travers des filtres. Origgi (2013) va au-delà, et introduit la notion de *traces* et perçoit la réputation plutôt comme «l*a valeur informationnelle de nos actions, la trace que laisse inévitablement notre conduite dans les opinions des autres chaque fois que nous interagissons avec le monde social, trace qui oriente leurs actions vis-à-vis de nous.* »[45]. Proposition qui met en exergue plusieurs éléments: la réputation comme construit social, la réputation comme agrégation de "traces" et la réputation comme jugement, évaluation, et qui est un facteur de reconnaissance sociale.

La notion de réputation comme nous le voyons se décline sous une pléthore de significations dans la plupart des lexiques, mais dans une acception générale, ces différentes propositions se retrouvent en certains points dans celle proposée par le Centre National de Ressources Textuelles et Lexicales qui dans son dictionnaire définit la réputation comme "*l'opinion favorable ou défavorable attachée à quelqu'un ou quelque chose; le fait d'être connu ou fortement connu pour ; connaître par ouï-dire, l'honorabilité, la renommée de bonne moralité ; la notoriété, le renom qui affecte une personne ou une chose*"[46]. Ainsi décrite, la notion de réputation met en exergue un ensemble d'actions, d'actes, de "traces" (dont nous proposerons un cadre d'analyse dans le chapitre 2, notamment dans le passage à l'e-réputation) qui, pris individuellement et cumulés, concourent à façonner une opinion "commune" souvent accolée à la notoriété, la renommée, et qui, *in fine* est perçue comme mesure, évaluation.

Eu égard aux différentes approches et à son caractère complexe, la recontextualisation du concept à travers son évolution historique et ses différentes significations nous amène à nous questionner sur la perception qu'ont les scientifiques de la notion à travers les différents champs disciplinaires.

---

[43] Qualifiées de "signaux informationnels" par Alloing (2013)

[44] Qui, pour lui, englobent les médias, les expériences personnelles ou autres

[45] Origgi, Gloria. « *Un certain regard* ». *Communications* n° 93, no 2 (13 novembre 2013): 101-20. https://www.cairn.info/revue-communications-2013-2-page-101.htm.

[46] Définition du concept de réputation proposée par le CNRTL http://www.cnrtl.fr/definition/réputation

2. La réputation comme ressource et capital : approches quantifiantes

L'approche quantitative de la réputation nous est proposée à travers le prisme des économistes et des gestionnaires, et met en exergue la réputation comme relevant autant d'une ***représentation ou perception*** (Fombrun, 1996 ; 1997), d'un ***actif intangible*** (Hall, 1992 ; Grant,1996 ; Scott et Walsham, 2005), d'un ***signal*** (*Ackerlof, 1970),* d'une ***garantie de contrat (***Cornell et Shapiro, 1987*),* et d'un ***jugement (***Barnett et al., 2006 ; Gray, Balmer, 1998 ; Da Camara, 2007).

La notion d'*actif intangible* de la réputation fait écho à son caractère stratégique (Hall, 1992 ; Grant,1996) et holistique pour l'entreprise. De par les actions qu'elle met en place, l'entreprise s'assure une certaine « place » dans son environnement et dans l'esprit de ses parties prenantes. La réputation ici, qui est une représentation collective, reprend l'antécédent et les résultats d'une organisation démontrant l'aptitude de celle-ci à fournir des résultats pertinents à ses diverses parties prenantes. Utilisée comme indicateur de la position de l'entreprise tant en interne auprès des employés qu'auprès des parties prenantes en externe, elle opère dans les environnements concurrentiels institutionnels. Perçue comme ressource intangible de l'entreprise, la réputation, qui implique une certaine forme de « transparence » avec les différentes parties prenantes, est ici source d'avantage concurrentiel et influe sur la performance de celle-ci, lui conférant un caractère de mesure, d'évaluation[47].

Outre les actions mises en place par l'entreprise, la réputation porte également sur les différents attributs de celle-ci et convoque de fait plusieurs notions connexes auxquelles elle est souvent assimilée.

- L'image voulue, qui se dit de la manière dont une organisation souhaite être perçue par ses différentes parties prenantes (Whetten & Mackey, 2002, p.401). En tant que projection, elle est façonnée par l'organisation puis externalisée par la suite auprès des différentes parties prenantes et appelée à évoluer (Dowling 2002 ; Bromley, 2000) ;

[47] Evaluation qui dans certains cas peut prendre la forme d'un classement

- L'identité,[48] relative à l'essence même d'une organisation ou d'un individu, ce qu'il est au sens de Whetten & Mackey (2002), et donc différente de la réputation qui est la manière dont elle est perçue ;

- La crédibilité, qui est fortement liée à la réputation et à l'influence fait référence au concept de confiance (sur lequel nous reviendrons plus en détail dans le chapitre 2) ;

- La légitimité qui fait référence au sens de Suchman (1995) aux normes sociales et suppose une adéquation entre les actions d'une entité et les valeurs (croyances et normes sociales de son environnement) ;

- Et la notoriété qui est souvent accolée au concept de célébrité et a trait au fait d'être connu et donc à une certaine forme de reconnaissance pas toujours liée à une compétence type au sens de Rindova & al. (2005[49]).

Les différents attributs identitaires qui se présentent sous forme de « signaux informationnels » affectifs[50] et cognitifs au sens de Fombrun et Van Riel (1997), façonnent une sorte d'identité auprès des publics, ce qui nécessite une analyse minutieuse, les significations pouvant différer d'un individu à un autre. Se met donc en place une certaine boucle au sein de laquelle « l'image » de l'organisation est directement perceptible par les collaborateurs, ces derniers contribuant et étant à l'origine de l'image de l'organisation. Cette boucle se réitère chaque fois que de nouvelles informations apparaissent sur l'organisation ou que de nouveaux *stakeholders* sont exposés aux signaux provenant de l'organisation.

---

[48] Nous y reviendrons à travers notamment la construction des identités numériques

[49] Rindova, Violina & Williamson, Ian & Petkova, Antoaneta. (2005). *Being Good or Being Known: An Empirical Examination of the Dimensions, Antecedents, and Consequences of Organizational Reputation*. Academy of Management Journal. 48. 10.5465/AMJ.2005.19573108. Accessible en ligne https://www.mv.helsinki.fi/home/aula/Top20/Being%20good%20or%20being%20known.pdf

[50] Relatif au sentiment analysis qui peut être décliné qualitativement en plusieurs sentiments

Cette première conception se voit complétée par celle de la réputation perçue comme représentation ou perception (Fombrun, 1996) qui met en exergue trois principales caractéristiques de la réputation: la première est liée aux perceptions, ce qui sous-tend que la réputation qui peut se construire sur des « items » non réels (la perception différant de la réalité), n'est pas le fruit de l'entreprise et lui échappe[51] ; en découle donc la deuxième caractéristique de la réputation qui est conçue comme un construit social, issu de la somme des perceptions de tous les *stakeholders* et faisant d'elle un « concept de réception », un « concept collectif de type macroéconomique »; et enfin la dernière caractéristique relative à la valeur de la réputation qui peut être soit positive ou négative, et perçue ici comme outil de classement et de comparaison des entreprises[52] .

La notion de réputation fait ici appel à trois constituants clés dont le premier est lié à la notion de représentation (tant sur les attributs identitaires que sur les actions passées, données qui échappent souvent au contrôle de l'organisation et sont parfois construites sur le ouï-dire). Que ces "*items*" soient l'apanage de l'entreprise ou non, les *stakeholders* dans un souci d'information, de confirmation ou d'infirmation d'une assertion, établissent des relations avec les autres parties prenantes de l'entreprise. Ce qui conduit donc à la seconde caractéristique évoquée par Fombrun et reprise par Boistel soit la notion de construit social. De par les échanges qu'entretiennent les stakeholders entre eux, ils produisent une forme de "perception commune" relative à l'organisation en question et basée sur la perception que chacun se fait de celle-ci. Cette perception commune, qui peut être positive ou négative, permet une analyse comparative de l'entreprise par rapport à ses concurrents. La réputation comme élément, à la fois de mesure (de la place qu'occupe une organisation) et mesurable (la valeur qui lui est attribuée) apparaît chez Fombrun comme la somme des "images" des parties prenantes d'une organisation. La réputation revêt ici un caractère d'évaluation, utilisé comme "outil" de classement, de comparaison.

---

[51] Perspective constructiviste à rapprocher de la sociologie

[52] La dimension de prospection/prédictibilité, de classement, et de facteurs favorisant la confiance sont des outputs de la réputation

La réputation comme évaluation se retrouve également chez d'autres auteurs en gestion sous forme de jugement. C'est notamment le cas de Barnett *et al.*, (2006)[53] pour qui elle est un jugement sur les différents attributs d'une organisation, jugement de la part de ceux qu'ils ont appelés « les observateurs » de l'organisation, et qui s'effectue tant en interne (notamment sur les aspects financiers) que sur le plan externe (aspects sociaux, environnementaux, etc.), et ce, sur une certaine temporalité. La réputation de l'entreprise perçue ici comme une évaluation « globale » des *stakeholders* se base sur l'adéquation entre l'être (le comportement, les actions de l'organisation) et le paraître (notamment la communication qu'elle met en place) de l'organisation (Gray, Balmer, 1998). Da Camara (2007) parle de concept holistique de la réputation, car elle englobe le jugement des *stakeholders* tant sur les actions de l'organisation que sa performance et s'appuie à la fois sur les attributs donc les composantes intrinsèques de l'entreprise et sur tous les aspects de celle-ci.

Enfin, la dernière valeur quantitative de la réputation est relative à celle de signal informatif qui met en exergue le « contenu » de l'information qui est véhiculé, informations qui reprennent les actions passées de l'entreprise et par extrapolation sous-tendent celles à venir dans la mesure où s'installe, de manière implicite entre l'entreprise et les différentes parties prenantes, ce que Cornell et Schapiro (1987) qualifient de « garantie de contrat ». Cet « engagement » implicite obligera l'entreprise qui est ici légitimée, à maintenir une constance dans les produits et services proposés afin d'éviter de « rompre ce pacte » et de perdre cette « réputation », qui garantira aux deux parties des « économies dans les coûts de transaction » (Holmstrom et Tirole 1989). Dans ce cas de figure, la réputation relève de l'évaluation, d'une forme de perception générale des *stakeholders* sur une organisation à partir des actions et informations sur celle-ci et revêt deux formes : positive ou négative. Une bonne réputation pour l'organisation suppose de fait un faible risque dans l'établissement d'une relation et surtout une fiabilité tant dans ses actions que dans les informations fournies ; ce qui oblige l'entreprise en question, afin de conserver cette réputation, à adapter ses actions futures.

---

[53] Barnett, Michael L, John M Jermier, et Barbara A Lafferty. « *Corporate Reputation: The Definitional Landscape* ». *Corporate Reputation Review* 9, nº 1 (mars 2006): 26-38. https://doi.org/10.1057/palgrave.crr.1550012.

A contrario, une réputation "défavorable" dénoterait auprès de toutes les parties prenantes de l'organisation un manque de fiabilité et donc un risque élevé d'interagir avec celle-ci, surtout si celle-ci sert ses propres intérêts (Cornell et Shapiro, 1987).

La réputation telle que perçue ici, nous le voyons, permet de réduire les risques et d'établir de fait une relation de confiance (nous y reviendrons dans la deuxième partie de ce chapitre) entre les différentes parties prenantes autant qu'elle est synonyme d'anticipation, de prédictions, les actions passées permettant de prédire celles à venir. L'évaluation se fera donc tant sur le contenu de l'information véhiculée par l'entreprise à travers ses communications, mais aussi et surtout sur son offre générale.

Cette perception de la réputation comme signal informatif est également reprise par certains auteurs de la théorie du jeu (Ching *et al*., 1992 ; Lahno, 1995 ; Whitmeyer, 2000 ; Ferris *et al.,* 2003) pour qui ce sont les actions passées qui sont à l'origine de la réputation. Pour ces derniers, les actions précédentes d'un joueur permettent de prévoir celles à venir, et il serait donc possible de prédire les possibilités d'interaction entre un joueur et les autres *stakeholders*. Un minimum d'interaction et de connaissance des différents acteurs s'avère donc nécessaire afin d'appréhender les différentes possibilités de jeu.

3. La réputation comme construit social de l'évaluateur ou du récepteur

La réputation perçue comme un ***construit social*** présuppose d'une ***source d'information déficiente*** et est un ***construit social***.

Cette première assertion part du postulat d'un manque d'informations sur un acteur précis, amenant les agents à avoir recours à d'autres sources possibles et plausibles d'information. Ici perçue comme un ensemble de "croyances collectivement partagées" et servant d'aide à la décision, la réputation agrège l'appréciation des autres acteurs (Bourdieu, 2000). Ce qui signifie, et comme le note Schapiro (1987), que les individus dans une situation d'évaluation et étant confrontés à un manque d'informations, auront recours à d'autres sources telles les journalistes, leaders d'opinion, etc. La réputation d'un agent évaluera sur la base des

informations recueillies, son prestige, mais prendra aussi et surtout en compte le contexte social dans lequel il évolue. D'entrée de jeu, on perçoit le lien étroit entre réputation et évaluation, et la mesure de la réputation s'effectue encore ici (comme cela peut être le cas dans certaines autres disciplines comme nous le verrons) à partir des actions passées d'un acteur et de leur perception par les différents agents qui interagissent avec lui. La réputation ici est donc une mesure pour un groupe social donné et il n'y a, de fait, pas de réputation absolue.

Cette forme de vision commune fait de la réputation un construit social dans la mesure où elle agrège l'ensemble des perceptions. Lang & Lang, (1988) et Camic (1992) parleront de « convention collective dominante » construite à partir des informations qu'un agent dispose sur un autre agent. La mesure de la réputation prend ici en compte trois principales caractéristiques : la notion d'évaluation permanente qui relève de la dimension spatio-temporelle de la réputation[54], les normes sociales (qui pour les sociologues influent sur les actions prises) et les attentes (qui sont subjectives et relatives à un contexte spécifique).

La réputation utilise des indicateurs (nous y reviendrons dans la partie II) afin de légitimer la position d'un acteur par rapport à un autre, et peut donc prendre la forme de classements dont la finalité est la reconnaissance sociale. Interaction sociale et réputation vont de pair, et pour Hahn *et al.* (2007), la réputation relève d'une « dimension structurelle de pratiques sociales ». Ils distinguent trois niveaux de réputation dans un macro phénomène : la réputation comme phénomène micro social qui est un ensemble d'évaluations à un niveau plutôt restreint, la réputation de méso-niveau qui correspond à l'estime sociale donc à une plus grande portée et la réputation sociale qui correspond au prestige à une « reconnaissance publique ».

Il est important ici de prendre en compte les notions d'appartenance, de risque et de projection ou anticipation qu'on retrouve également chez les économistes et gestionnaires avec notamment la notion de garantie de contrat. En psychologie sociale, la réputation est pensée comme l'évaluation du facteur risque dans l'établissement d'une relation entre

[54] La base de la réputation, nous l'avons mentionné, étant l'interaction sociale qui se déroule dans un temps t et selon un objet spécifique, l'opération de mesure qui s'effectue donc dans cet intervalle temps t et "objet", se réitère autant de fois qu'il ya d'interactions autour d'un objet ou de personnes différentes impliquées

plusieurs entités. Pour ces penseurs (Kelley, 1973[55] ; Humphrey, 1976[56] ; Funder, 1999[57] ; Emler, 1994[58] ; Emler, 2005[59] ; Emler & al., 2007[60] ; Anderson and Shirako, 2008[61]), les comportements des individus dans la perception des tiers sont régis et influencés par une pluralité de sources (variations pouvant être liées à l'acteur, à la situation singulière au cours de laquelle a lieu le comportement et les circonstances qui l'entourent par exemple pour Kelley[62]) et qu'il convient d'appréhender et d'analyser ces diverses variances afin de comprendre le processus de construction de ce "jugement". Les actions antérieures d'un acteur sont donc utilisées comme unité de mesure afin de prévoir celles à venir. La réputation ici est intrinsèquement liée aux notions de "confiance, appartenance et risque" (Castelfranchi & Falcone, 1998 ; Marsh, 2000 ; Gambetta, 2000 ; Grandison & Sloman, 2000 ; Luhmann, 2000 ; McKnight & Chervany 2001 ; Aljazzaf et al., 2010), concepts sur lesquels nous reviendrons dans le chapitre suivant.

---

55 Harold Kelley, « The processes of Causal Attribution », *American Psychologist*, 28, 1973, p. 107-128

56 Nicholas Humphrey, « The Social Function of the Intellect », *in* P. P. G. Bateson et R. A. Hinde (Eds), *Growing Points in Ethology*, Cambridge, Cambridge University Press, 1976.

57 David Funder, *Personality Judgment : A Realistic Approach to Person Perception*, San Diego, CA, Academic Press, 1999

58 Nicholas Emler, « Gossip, Reputation and Social Adaptation », *in* R. Goodman et A. Ben Ze'ev (Eds), *Good gossip*, Kansas, Kansas University Press, 1994, p. 117-133

59 Nicholas Emler, « A Social Psychology of Reputation », *European Review of Social Psychology*, 1, 1990, p. 171-193 ; Robin Dunbar, « Gossip in Evolutionary Perspective », *Review of General Psychology*, 82, 2005, p. 100-110.

60 Nicholas Emler, Hammond Tarry et Angela Saint James, « Post-Conventional Moral Reasoning and Reputation », *Journal of Research in Personality*, 41, 2007, p. 76-89

61 Cameron Anderson et Aiwa Shirako, « Are Individuals' Reputations Related to Their History of Behaviour ? », *Journal of Personality and Social Psychology*, 94, 2008, p. 320-333

62 Notamment dans son modèle de covariation qui part du postulat que les conclusions sur un objet sont un ensemble de données brutes issues des observations de ses actions et réactions, conclusions qui fournies par un observateur aux utilisateurs finaux et qui peuvent présenter un biais dans la mesure où elles peuvent d'entrée de jeu être influencées par ses motivations personnelles (la politesse, la loyauté pour Burt, le besoin de lui porter atteinte pour un concurrent par exemple, etc.)

4. Les espaces publics de la réputation : le paradigme info communicationnel

Parler d'espaces publics de la réputation revient à s'intéresser au paradigme infocommunicationnel et plus particulièrement aux travaux en Sciences de l'information et de la communication, discipline au sein de laquelle s'insère notre travail de recherche et cependant parent pauvre dans le domaine de la réputation.

L'un des travaux en SIC sur la réputation a été proposé par Nicole D'Almeida (2007) qui apporte une nouvelle vision et introduit l'idée des espaces publics et plus précisément d'espace social de construction de la réputation. La réputation pour D'Almeida recouvre à la fois les « processus de formation de l'opinion » et leur diffusion dans l'espace public et médiatique, et est associée à la notion de visibilité de l'organisation soit relative aux attributs identitaires qu'elle met en avant.

L'opinion, qui renvoie selon elle au concept *d'opinion publique*, est dénuée de la notion de risque, elle permet d'établir une relation entre les entités et les diverses parties prenantes dans les divers espaces où se jouent la mesure de la réputation, l'espace public étant une plateforme d'échange où se fabrique une certaine réputation.

L'interaction sociale revêt ici une place importante dans l'établissement de relations, ce qui fait de la réputation pour D'Almeida un construit social qui s'appuie sur deux notions connexes : l'imputation et la députation. Par imputation, elle entend l'évaluation positive ou négative d'une action imputée à une entité qui se joue donc a postériori, car impliquant « une obligation de faire » et donc en retour une réponse, et par députation, la mission d'intérêt public endossée par un ou plusieurs individus et qui sous-tend une « volonté de faire ». La réputation revêt ici un caractère «politique», elle est construite par les interactions (dialogues, débats, etc.) entre les différents individus à travers le temps (action présente qui sera mesurée sur le long terme), nécessitant de fait « *une égale capacité de chacun à produire des jugements, à examiner ou proposer du possible* », D'Almeida (2007, p.117).

L'objet réputation a également été traité en SIC par Camille Alloing. Pour lui (2015)[63], la réputation relève d'une forme «d'évaluation commune ou individuelle» d'un organisme. Cette mesure, part de la captation de signaux dits informationnels qui, pour lui, peuvent être des jugements ou opinions et qui se (re) présentent sous forme d'image ou de "marqueurs".
Nicolas Chazaud lui, dans sa thèse (2008)[64], utilise plutôt le concept "d'asymétrie d'informations" et définit la réputation comme un construit d'informations échangées dans un groupe dans l'optique de pallier à l'asymétrie ou au manque d'informations entre l'entité évaluée et le groupe et ce, à partir de critères d'évaluation partagés. Chazaud met ici en exergue deux notions importantes : la notion d'asymétrie d'informations et celle de critères d'évaluation.

La réputation perçue sous l'angle informationnel a également été étudiée en philosophie. Pour Gloria Origgi (2013)[65] qui reprend Adam Smith dans la *Théorie des sentiments moraux* (1759), la réputation est un ciment des relations humaines et sociales, instaurant de fait un besoin de coopération, d'approbation, de consensus. Elle distingue « deux mécanismes réputationnels : le premier est basé sur ce qu'elle qualifie « d'intérêt égoïste dans les interactions répétées », et le second, le besoin d'approbation. Parler de réputation revient ici à s'intéresser aux concepts de construit social, de traces et de jugement. La réputation comme construit social met en exergue la notion d'interaction relative au paradigme communicationnel, qui est ici et comme nous avons pu le constater pour plusieurs autres auteurs l'élément moteur de production d'informations, informations qui, seront utilisées dans l'attribution de la réputation. La réputation relève également pour Origgi de l'agrégation de "traces", ici perçues comme l'ensemble des actions issues des interactions avec les autres. Ces traces, qui peuvent être volontaires ou non, servent à évaluer un objet, à attribuer un jugement sur un individu dans la société. D'où la dernière caractéristique de la réputation, l'évaluation, qui est systématique à chaque interaction et est un facteur de reconnaissance

---

[63] Alloing, Camille. « *Votre entreprise plus nette sur le Net* » . *Comment les professionnels fabriquent-ils les publics et les imaginaires de l'e-réputation ?* » *Communication. Information médias théories pratiques*, nº vol. 33/2 (11 décembre 2015). https://doi.org/10.4000/communication.6060.

[64] Chazaud N., *Réputation d'entreprise et veille informationnelle : vers un modèle anticipatif de gestion des risques de réputation sur internet*, Thèse de doctorat, Université Montpellier 1, 2008.

[65] Origgi, G., « *Un certain regard* ». *Communications* n° 93, nº 2 (13 novembre 2013): 101-20. https://www.cairn.info/revue-communications-2013-2-page-101.htm.

sociale. Comme le souligne Origgi « *toute interaction sociale comporte une dimension d'évaluation, de jugement réciproque, une perception de qui nous sommes donnée par notre conduite, la constitution d'une image dans les yeux des autres* »[66]. Cette évaluation qui revêt une forme d'étiquette, nous permet une adaptation perpétuelle à autrui dans la société et est un facteur de stabilité pour notre réputation.

La réputation comme construit d'information a été également étudiée en philosophie par Erin Coyne dans sa thèse (2010)[67]. Pour elle, la réputation d'une entité est la résultante de plusieurs facteurs informationnels perçus par ses « cibles » : les informations sur l'entité en question (subjectives et objectives), le contexte dans lequel elle évolue, les interactions qu'elle entretient avec les publics (*third parties*) ou entre les publics et l'entité évaluée, et les observations de l'évaluateur. Ces différentes informations permettront à l'évaluateur soit de juger directement de la réputation d'une entité à partir des informations dont il dispose et d'agir en conséquence, soit de poser les jalons qui une fois ces informations confirmées ou infirmées (par les interactions directes qu'il entretiendra avec l'entité, les observations des attitudes des autres, etc.) serviront de mesure de la réputation de l'entité.

5. La réputation : une définition stabilisée à travers les différentes disciplines

L'objet réputation, nous le voyons, englobe une forme plurielle de significations à travers les différents champs disciplinaires. Comme nous l'avons mentionné, notre intérêt dans cette partie se porte sur l'objet de mesure pris en compte pour chacune des disciplines :

- La sociologie, qui perçoit la réputation comme un construit social à l'origine d'une déficience de sources d'informations, utilise pour objet de mesure le contexte social dans lequel évolue un agent. Par contexte social, il s'agit autant des ***normes sociales***

[66] Ibid.

[67] Coyne, E. (2010). *Reputation as Information: A Multilevel Approach to Reputation in Organizations* [Doctoral dissertation, Ohio State University]. OhioLINK Electronic Theses and Dissertations Center. http://rave.ohiolink.edu/etdc/view?acc_num=osu1276696814

que des ***attentes*** d'un individu vis-à-vis d'un autre. Pour les sociologues, le contexte dans lequel évolue un agent influe de manière indéniable sur ses différentes prises de position (choix d'interactions, avis, etc.) et l'attribution de sa réputation dépend en partie des normes et attentes sociétales. La réputation ici, qui est une forme de « convention collective » (Lang et Lang 1988 ; Camic, 1992) prenant en compte l'ensemble des perceptions d'un agent à partir des interactions qu'il a entretenues avec un autre, peut donc s'effectuer soit au niveau micro social, méso social ou macro social et a pour finalité première la reconnaissance sociale ;

- Pour les économistes, parler de réputation renvoie à la fois aux notions de signal, d'actif intangible et de garantie de contrat. La réputation ici, qui est celle d'une entreprise ou organisation, se base à la fois sur les ***offres*** (produits et/ou services) et les ***actions passées*** d'une entité. La réputation qui est perçue comme un enjeu concurrentiel pour l'organisation visant à asseoir son positionnement par rapport à ses concurrents, est agrégée à la fois par l'organisation (de manière implicite à travers son offre et ses différentes actions), mais aussi et surtout par les parties prenantes. La notion de garantie de contrat (Cornell et Shapiro 19982) est donc ici comme nous l'avons vu, essentielle, car pour conserver ou avoir une "bonne réputation", l'organisation s'assure de proposer ou conserver la même ligne de conduite, la réputation évaluant les actions passées. Une des notions clés qui intervient est donc celle de "prédiction" à partir de l'existant à laquelle est corrélée celle de **confiance** ;

- Les gestionnaires dont la vision rejoint en certains points celle partagée par les économistes sont ceux ayant produit le plus d'écrits sur la réputation. Pour eux, l'objet réputation peut prendre la forme de perception ou représentation, d'actif intangible, mais aussi et surtout elle relève d'une évaluation, d'un jugement de valeur sur une organisation. Ils introduisent la notion de *corporate reputation* qui se fonde sur la mesure à la fois des ***attributs identitaires*** et des ***actions passées*** d'une organisation. En gestion, la réputation part du postulat d'une asymétrie ou d'un manque d'informations sur une organisation, ce qui amène ses parties prenantes à rechercher, auprès d'autres parties prenantes ayant déjà interagi avec l'organisation,

des données de référence. L'une des données importantes sur laquelle repose la mesure de la réputation en sciences de la gestion est l'ensemble des attributs identitaires (Fombrun, 1996 ; Bromley, 2000 ; Barnett et al, 2006 ; Boistel, 2014, etc.). Par attributs identitaires, les chercheurs entendent l'identité, l'image, la culture et la marque d'une organisation, données qui, pour eux, sont perçues sous forme de représentation, donc subjectives et échappant *de facto* au contrôle de l'organisation. À ces informations sont corrélées les différentes actions menées par l'organisation tant sur les aspects sociétaux, financiers, environnementaux, etc., d'où la notion de caractère intangible qu'emploient les gestionnaires et qui désigne le caractère holistique de la réputation. La mesure de la réputation permet à l'organisation de se positionner par rapport à ses concurrents et peut prendre la forme (comme cela peut être le cas en économie), de classement;

- En SIC, la réputation, qui a peu été étudiée, prend la forme d'une évaluation basée sur l'***opinion*** (D'Almeida, 2017) et les différents ***signaux informationnels*** (Alloing, 2015) émis par une entité. D'entrée de jeu, on perçoit en SIC que la réputation est une évaluation de l'information véhiculée par une entité sous forme de signal, jugement ou d'opinion;

- Ce point de vue est partagé également par les philosophes, qui étudient les signaux informationnels et notamment le concept de ***traces*** (Origgi, 2013) laissées de manière intentionnelle ou non. La philosophie part du postulat que toute interaction sociale comporte une dimension d'évaluation, et met en exergue le caractère holistique de la réputation qui est perçue comme construit social, agrégation de "traces" et comme jugement. La mesure de la réputation d'une entité qui utilise autant les informations personnelles, le contexte dans lequel elle évolue, les interactions qu'elle entretient avec les autres entités a pour principal objectif la reconnaissance;

- Les psychologues, quant à eux, et notamment la psychologie sociale, la perçoit par le prisme du risque (risque d'établir une relation avec un agent entre autres) et notamment à travers la mesure des ***états mentaux,*** ce qui implique une certaine forme

de confiance accordée à ceux qui rendent compte, aux motivations subjectives à l'origine de leurs conclusions.

Le tableau ci-dessous résume les différentes prises de position sur le concept :

Disciplines	**Perception**	**Objet de mesure**
Economie	- Actif intangible - Signal - Garantie de contrat	- Offre (biens / services) - Actions passées
Sociologie	- Construit social	- Interaction sociale - Contexte (normes, attentes)
Gestion	- Perception / Représentation - Actif intangible - Évaluation / Jugement	- Actions passées - Attributs identitaires
SIC	- Evaluation	- Opinions - Informations
Psychologie	- Risque	- Etats mentaux - Confiance
Philosophie	- Signal - Evaluation	- Traces (opinions et actions)

Tableau n°1 : Approche comparative par discipline de la perception et de l'objet pris en compte dans la mesure de la réputation

À partir de ces différentes prises de position, plusieurs observations peuvent être faites d'emblée :

- Au sein même de certaines disciplines, on note une pluralité de perceptions du concept de réputation : en économie (actif intangible, signal, garantie de contrat), en gestion (perception, actif intangible, évaluation ou jugement), en philosophie (signal, évaluation). Cette pluralité de significations intra disciplinaire permet autant de recouper certaines prises de position que de les différencier. C'est notamment le cas pour les économistes et les gestionnaires, dont la perception du concept est

convergente en certains points et divergente en d'autres. L'un des points de consensus est leur perception de la réputation au sens de *corporate reputation* comme actif intangible de l'entreprise, dont la finalité est le classement, le positionnement dans un environnement concurrentiel. Pour se faire donc, la mesure de la réputation s'effectue autant sur les différentes actions posées par l'organisation que sur ces différents attributs identitaires (en gestion) et son offre (économie). S'ils s'accordent en certains points sur l'objet de mesure, leurs avis divergent sur les acteurs à l'origine de la production des données utilisées dans la mesure de la réputation. Pour les économistes, la réputation avec sa notion associée de garantie de contrat, met en exergue autant une implication de l'entreprise (dans la mise à niveau perpétuelle de son offre) que de celle des parties prenantes.

Ce qui pour eux, fait de la réputation, un construit de l'entreprise et de ses différentes parties prenantes. Vision que ne partagent pas les gestionnaires pour qui la réputation se fonde en partie sur les différents attributs identitaires de l'organisation. Ces attributs identitaires (identité, image, culture, marque, etc.) étant une perception/ représentation au sens de Fombrun (1996, 1997), échappent donc de fait au contrôle de l'organisation, ce qui pour eux, fait de la réputation, le fruit des parties prenantes uniquement. Il existe également chez les sociologues, les psychologues et les philosophes, même si leur perception de la réputation est différente, une certaine forme de consensus notamment sur la finalité de la réputation. Pour les sociologues pour qui elle est un construit social dont l'origine vient d'une déficience de sources d'informations, la visée première est la reconnaissance sociale, notamment via les différentes normes et attentes des acteurs.

Notion de reconnaissance qu'on retrouve également chez les psychologues, pour qui la réputation qui étudie les états mentaux d'un agent vise à réduire les risques liés à la décision d'interagir avec lui, et à établir une relation de confiance, gage de reconnaissance. La reconnaissance pour les philosophes est matérialisée par la notion d'étiquette qu'ils attribuent aux différents agents notamment via les traces laissées par ces derniers et qui servent d'éléments de mesure. La notion de reconnaissance prend également tout son sens chez les économistes et gestionnaires sous la forme de "reconnaissance publique" à travers les différents classements qu'ils proposent;

- Une seconde observation qui ressort de notre analyse est que l'une des notions clés, commune à toutes les disciplines, est celle d'*interaction*. Que ce soit en économie, en sociologie, en sciences de la gestion, en SIC, en psychologie ou en philosophie, la mesure de la réputation d'un acteur suppose qu'il y ait au préalable une interaction. Interaction, qui peut s'effectuer dans plusieurs contextes : social et numérique. L'interaction, qui se fonde en partie sur une asymétrie ou un manque d'informations sur un objet spécifique, vise à collecter des données sur ce dernier qui selon l'agent sur lequel porte la mesure peuvent prendre diverses formes. Et qui dit interaction, dit à minima la présence de deux agents;

- Sur les agents, on distingue à partir de notre analyse et selon les différentes disciplines, deux types d'acteurs pris en compte dans un processus de mesure de la réputation : l'individu et l'organisation. Sur l'individu, la mesure de la réputation permet d'évaluer autant ses interactions sociales, ses états mentaux que les différentes traces qu'il laisse. La mesure de la réputation pour les organisations ou *corporate reputation* étudiée par plusieurs disciplines (économie, gestion et SIC entre autres), se fonde elle autant sur l'organisation elle-même et ses différentes actions (attributs identitaires, aspects sociaux, environnementaux par exemple) que sur son offre (biens et services);

Pour qu'il y ait donc réputation selon nous, et à partir des différents travaux sus-cités, plusieurs composantes s'avèrent essentielles : un contexte, la présence de deux ou plusieurs agents, un objet de mesure et un historique servant de base à la mesure.

- le **contexte**

Par contexte, nous entendons l'environnement dans lequel s'effectue la mesure de la réputation. Celui-ci peut être soit virtuel (via notamment les médias sociaux, les moteurs de recherches, les forums, etc.) ou non (interaction sociale notamment);

- les **agents en présence**

Mesurer la réputation suppose de prime abord et comme nous l'avons vu, qu'il y ait interaction. Et parler d'interaction revient à identifier à minima 2 acteurs (un destinataire en quête d'informations et un émetteur source de l'information). Que la mesure de la réputation s'effectue dans un contexte numérique et/ou social, plusieurs agents peuvent intervenir dans la production des données qui seront par la suite utilisées dans le processus d'évaluation;

- l'**objet de mesure**

La mesure de la réputation, nous l'avons vu, se base sur un "objet" qui peut, selon les cas, être une ***personne*** (physique ou morale), un ***produit ou un service*** *ou* ***une organisation***. Selon le type d'acteur sur lequel porte la mesure de la réputation, le contexte, le type d'informations à collecter et les différentes sources de collecte peuvent être différents;

- l'**historique**

La notion d'historique renvoie à l'ensemble des informations qui sont collectées dans la mesure de la réputation. Les données peuvent se rapporter à un commentaire, un avis, une note, et faire référence à des (inter) actions passées.



Figure n°2 : Ce que suppose la réputation

Conclusion du chapitre 1

Ce que nous avons voulu montrer dans ce chapitre c'est que l'objet réputation implique une évaluation, avec pour chacun des champs disciplinaires, l'attribut sur lequel porte la mesure. En effet, qu'elle soit perçue comme résultante d'interactions sociales, comme construit d'opinion ou d'information, comme signal, comme actif intangible de l'entreprise et corrélée aux notions d'identité, d'image et de notoriété, ou encore à travers les notions de risque, de traces, tous les auteurs s'accordent à dire qu'elle vise une évaluation.

Cette forme d'évaluation porte, selon la discipline, sur divers attributs : l'offre (produits ou services), les attributs identitaires, les actions passées, les états mentaux, les signaux informationnels, etc. Nous avons donc tenté, dans cette partie et à partir des différentes prises de position, d'identifier et de définir les éléments mesurés dans chaque conception.

À partir des différents concepts qui contribuent à façonner la réputation "traditionnelle", nous proposerons dans le point suivant, un cadre d'analyse des différents éléments qui semblent, dans une acception commune, avoir été transposés au passage de la réputation à la e-réputation et tenterons d'expliquer la manière dont ces différentes caractéristiques se transforment dans le processus de numérisation. Comment change la conception de la réputation avec la recrudescence des outils numériques ? Certains éléments notionnels ont-ils mué ou été abandonnés ? De nouveaux concepts sont-ils apparus?

CHAPITRE 2 : DE L' E-RÉPUTATION À L'INFLUENCE

Introduction du chapitre 2

Même si Howard Rheingold s'intéressait déjà dans les années 1990 aux cybercommunautés[68] et à l'e-réputation, le terme *online reputation* est apparu dans un écrit de McDonald et Slawson (2000). Dans leur article « *Reputation in an internet Auction Market* »[69], les auteurs partent du postulat que les notions de réputation et de confiance sont intrinsèquement liés, et s'intéressent aux scores de réputation des vendeurs d'eBay. Ces scores de réputation qui sont importants dans la stratégie de prix (*pricing*) des vendeurs, reposeraient à la fois sur les avis laissés par les acheteurs, l'historique des vendeurs, les étoiles accordées, etc. Ils ont donc proposé un modèle mathématique dans l'optique d'établir une relation entre la réputation des vendeurs, le prix des produits et les offres proposées. La notion d'e-réputation associée à celle de confiance (initialement mobilisée par les psychologues) a très largement été étudiée dans la littérature. Les chercheurs qui s'y sont intéressés ont tenté de modéliser la réputation en ligne à travers les interactions que les « agents » entretiennent sur ces plateformes numériques. On assistera dès lors à l'apparition de modèles informatiques de mesure de la réputation avec la notion de confiance au centre des interactions (et dont nous proposons une lecture dans la deuxième partie).

Le terme sera, par la suite, repris dans un article scientifique de Chun et Davis (2001)[70] sous sa forme actuelle « e-réputation ». Les auteurs, dans leur écrit, s'intéressent, à partir de l'analyse de sites web d'organisations présents dans le classement de Forbes, à la construction de la réputation en ligne d'une organisation à travers son site web. La notion d'e-réputation est ici intrinsèquement liée aux attributs « moraux » que possède une organisation (vision, valeurs, etc.) et qui, pour Chun et Davis, et comme nous le constatons sur la quasi-totalité des sites institutionnels encore aujourd'hui, méritent d'être minutieusement pensés et mis en exergue, créant du lien avec les visiteurs d'un site web.

---

[68] Il a notamment écrit le livre "*The virtual community : Homesteading on the Electronic Frontier, 1993* https://dlc.dlib.indiana.edu/dlc/bitstream/handle/10535/18/The_Virtual_Community.pdf?sequence=1
D'autres ressources sont accessibles sur son site web http://rheingold.com

[69] McDonald, Cynthia G. & Slawson, V. Carlos, (2002), *Reputation in An Internet Auction Market*, *Economic Inquiry*, 40, issue 4, p. 633-650, https://EconPapers.repec.org/RePEc:oup:ecinqu:v:40:y:2002:i:4:p:633-650.

[70] Chun, R., et G. Davies. « *E-reputation: The role of mission and vision statements in positioning strategy* ». *ResearchGate*8, no 4 (1 mai 2001): 315-33. https://doi.org/10.1057/palgrave.bm.2540031.

On assistera par la suite à l'apparition de travaux de recherche dans les différents champs disciplinaires. Nicolas Chazaud, docteur en SIC, dans sa thèse de doctorat (2008)[71], s'intéresse à la gestion de la réputation en ligne et plus particulièrement aux risques associés aux traces numériques des organisations dans un environnement toujours plus concurrentiel. Pour lui, qui dit gestion de sa réputation sur internet dit veille stratégique et intelligence économique. Il propose donc, un modèle de gestion de la réputation en ligne se basant sur la veille et l'analyse de l'environnement informationnel d'une organisation.

Cette démarche qu'il qualifie de « surveillance », et qui pour Chazaud se veut « anticipative » et non « réactive » s'effectue en deux principales étapes. La première phase qui consiste en une surveillance environnementale de l'organisation englobe à la fois une analyse de l'environnement interne et externe de celle-ci. La notion d'environnement externe est ici relative à toutes les « parties prenantes » de l'organisation, qu'il s'agisse des clients, des partenaires ou des concurrents. Il s'agit donc d'identifier et analyser les attentes, les besoins et le comportement des clients et partenaires, de même que l'offre de produits ou service proposée par les concurrents.

La seconde phase de l'analyse environnementale, interne, se focalise sur l'identification des valeurs et des « failles » de l'organisation (produits ou services, performances, etc.). Cette surveillance une fois effective servira pour Chazaud, dans l'hypothèse où les informations collectées présenteraient un risque pour l'organisation, à évaluer la teneur ou gravité et à proposer des stratégies de gestion adéquates. Omar Hasan dans sa thèse en informatique (2010)[72] se propose d'étudier les différents systèmes de réputation dont la préservation de la vie privée est l'enjeu majeur et d'identifier les problèmes liés aux méthodes de recommandation sur ces différentes plateformes. En 2013, Camille Alloing[73] propose un

[71] Chazaud, N., *Réputation d'entreprise et veille informationnelle : vers un modèle anticipatif de gestion des risques de réputation sur internet*,Thèse de doctorat, Université de Montpellier 1, 2008.

[72] Hasan O., *Privacy Preserving Reputation Systems for Decentralized Environments*, Institut National des Sciences Appliquées de Lyon, Informatique et Mathématiques, 2010, 168p. Accessible sur http://theses.insa-lyon.fr/publication/2010ISAL0066/these.pdf

[73] Alloing C., *Processus de veille par infomédiation sociale pour construire l'e-réputation d'une organisation. Approche par agents-facilitateurs appliquée à la DSIC de La Poste*. Sciences de l'information et de la communication. Université de Poitiers, 2013. Français. NNT : 526 . tel- 00915004. Accessible sur https://tel.archives-ouvertes.fr/tel-00915004/document

modèle et un processus de veille visant autant à construire qu'à gérer l'e-réputation d'une organisation. Céline Fueyo en marketing (2015)[74] quant à elle s'intéresse plutôt à la gestion de l'e-réputation des organisations dans le secteur de l'automobile et plus précisément à l'introduction des réseaux sociaux professionnels de leurs salariés dans leur stratégie d'e-réputation et d'influence. À partir d'études quantitatives, elle se propose donc d'étudier l'usage et la perception (tant en termes de qualité que de notoriété) des pages entreprises et personnelles sur la plateforme Linkedin.

En sciences de la gestion, Chebli Youness (2016)[75] quant à lui, s'intéresse à la *corporate reputation* et analyse les retombées (confiance, engagement, risque, etc.) de l'e-réputation pour le client en utilisant les plateformes de e-commerce. Le droit n'est pas en reste dans la littérature en e-réputation. Jean-Charles Ize (2018)[76] nous propose dans ses travaux une analyse des différentes règles juridiques à prendre en compte afin de résoudre les problèmes de réputation en ligne. Même si, nous le voyons, certains auteurs se sont intéressés à la question avec une approche pluridimensionnelle (sciences de la gestion, droit, SIC, marketing, etc.), force est de constater que peu a été fait sur la notion et de manière plus précise sur la mesure de l'e-réputation. Dans cette partie, nous proposons d'analyser les différents concepts clés qui concourent à la construction des modèles théoriques de l'e-réputation, et nous intéresserons plus particulièrement aux notions de trace numérique, d'identités numériques, de confiance et d'influence numérique.

---

[74] Fueyo, C. (2015) *E-réputation corporate : influence de la voix digitale des employés via l'e-contenu de leur « Profil » sur les réseaux sociaux professionnels ; application au secteur automobile. École Doctorale Sciences de Gestion TSM (Toulouse). Accessible sur* http://publications.ut-capitole.fr/19361/1/FueyoCeline2015.pdf

[75] Chebli, Y., *L'e-réputation du point de vue client : modèle intégrateur et échelle de mesure,* Université Grenoble Alpes, 2016. Accessible sur https://www.theses.fr/2016GREAG016

[76] Ize J. C., *Les droits relatifs à l'e réputation*, Université d'Aix Marseille, 2018. Accessible en ligne sur https://www.theses.fr/2018AIXM0029

1. La trace numérique comme source de l'e-réputation

Parler de trace numérique revient à s'intéresser en amont à la notion de trace, qui peut se décliner sous diverses formes et pour des objets différents : récit, déroulé, scénario, compte rendu, enregistrement, etc. (Milles, 2013, p.8).

Alexandre Serres (2002)[77] nous propose un retour rapide à travers les différentes époques du concept de trace : d'abord perçue comme empreinte ou suite d'empreintes matérielles ou morales, elle équivaut ici *à « ce qui a été laissé »* par un homme ou un animal vers 1120 ou une chose (en 1690), à une « manière d'agir (vers 1190), à un « exemple à suivre » (1530).

Un deuxième sens, toujours selon lui, fera son apparition vers 1250 sous le sens cette fois de « marque » puis « d'égratignure ». On parlera entre autres « d'impression » vers 1538, et notamment de la notion de mémoire en 1679, renvoyant de facto vers la troisième définition, celle de « petite quantité » en 1847. La dernière signification pour Serres, apparaîtra notamment avec le graphisme en 1439 encore sous diverses déclinaisons : « faire une trace pour raturer, effacer », « écrire » avec la notion de trait de quelque chose apparue au XVIe, «marquer, représenter via des lignes au XVIIe puis marquer le contour ».

Cette première assertion permet à Serres de distinguer de facto plusieurs significations du concept de trace : comme empreinte relative au psychisme, à la mémoire d'un objet, la trace comme indice ou quantité "infinitésimale", la trace comme mémoire collective faisant référence au concept de redocumentarisation (Ertzscheid, 2013 ; Merzeau, 2009) , la trace comme écriture au sens d'une représentation de la parole et enfin, la trace comme caractère moral au sens "*d'exemple à suivre*". Afin d'établir une relation entre ces différentes approches et le concept de trace numérique, il convient à partir de là, de questionner les constituants clés intervenant dans le passage de la trace à la trace numérique, et nous nous aiderons pour ce faire de quelques définitions du concept de trace numérique perçue comme « *trace d'une activité humaine dans un environnement numérique* »[78] ou « *mode inédit de*

---

[77] Serres A.. *Quelle(s) problématique(s) de la trace ?* 2002. sic_00001397. Accessible en ligne https://archivesic.ccsd.cnrs.fr/sic_00001397/document

[78] LAFLAQUIÈRE J., « Trace numérique : de l'inscription de connaissances à la mémoire active », Atelier Trace. IC2011, Lyon, [s.n.], 2011, p. 1.

présence et d'efficacité, lié aux caractéristiques techniques et sociales des réseaux[79] » ou encore et pour paraphraser Collomb (2016) « *il y a trace numérique dans la mesure où toute action dans un environnement informatisé est aussi une interaction avec des machines* »[80]. Passer donc de la trace à la trace numérique suppose une certaine forme « d'interaction homme-machine »[81] et plus généralement l'existence d'environnements numériques comme interface, environnements qui ont favorisé l'apparition de certaines propriétés du numérique inhérentes aux traces numériques : la matérialité, l'intentionnalité, la temporalité, calculabilité, l'interprétativité, la réflexivité et la mémorisation (Collomb, 2016 ; Serres, 2012 ; Merzeau, 2009).

- Par matérialité, on part du postulat comme le note Serres que le signe distinctif de la trace numérique est son « support » et que sans écran ni dispositif technique, celle-ci ne serait ni visible, ni décodable, ni lisible. Cette assertion, qui vient en contrepoint des visions communes du virtuel et de la dématérialisation du numérique, comme il le souligne, crée du lien entre traces numériques et non numériques, les dernières étant également gravées dans une matière;

- La deuxième propriété, l'intentionnalité, rejette le caractère spontané et non intentionnel de la trace numérique (en partie comme le note Merzeau, 2009) en ce sens qu'elle est « construite » à travers le temps et est vouée à être enregistrée à partir des interactions des usagers avec une pléthore d'outils et systèmes. Les traces numériques sont donc provoquées, même si les interactions peuvent relever de la spontanéité;

---

[79] Merzeau Louise, « Du signe à la trace : l'information sur mesure », *Hermès, La Revue*, 2009/1 (n° 53), p. 21-29. URL : https://www.cairn.info/revue-hermes-la-revue-2009-1-page-21.htm , p.2

[80] Collomb, Cléo (2016), Un concept technologique de trace numérique, Thèse en philosophie, Université de Technologie de Compiègne, p.15

[81] Ibid.

- Le concept de temporalité suppose lui, que toute trace numérique ou non, étant liée à un événement ou une action passée, est *de facto* inscrite dans le temps. La trace numérique est ici perçue comme un document numérique en référence à la vision de Pédauque et Salaün (2006, p. 186)[82] et Melot[83], qui est composée « *d'unités isolables, agençables et calculables* » renvoyant donc la trace numérique ici au caractère de discrétisation qu'on retrouve chez Bachimont à travers l'analyse de Collomb (2016);

- La quatrième propriété, la calculabilité qui selon Serres éloigne le plus les traces numériques de celles analogiques, renvoie simplement à la finalité, à l'utilité faite des données collectées soit l'analyse et la prédiction. Elle induit et est à l'origine d'une forme de « redocumentarisation » (Pédauque, Salaün et Melot , 2006 ; Ertzscheid, 2013 ; Merzeau, 2009);

- L'interprétabilité (au sens de Serres) ou manipulabilité (chez Collomb) part du postulat que la calculabilité ou l'autonomie des algorithmes utilisés, n'exclut nullement le caractère interprétatif, car, présentées sous forme brute, les données numériques ne sauraient être comprises. De plus, l'interprétation étant subjective, un cadrage préalable s'avère nécessaire;

- Une autre propriété inhérente aux traces numériques est la réflexivité qui renvoie au caractère intrinsèque même des données numériques (quantité de contenus collectés, diversité des sources, multiplicité des formats, etc.);

- Et enfin, la mémorisation, qui comme le souligne Serres n'est pas un phénomène nouveau pour la trace, mais revêt à travers le numérique une nouvelle forme.

---

[82] La Redocumentarisation du Monde par le collectif Roger T. Pédauque (Cépaduès Éditions, 2007) et J.-M. Salaün, « La redocumentarisation, un défi pour les sciences de l'information », Études de Communication, n° 30, « Entre information et communication. Les nouveaux espaces du document », Université Lille III, 2007

[83] M. Melot, préface à Roger T. Pédauque, *Le Document à la lumière du numérique*, Caen, C&F Éditions, 2006, p. 14.

De ce rapide panorama, plusieurs constats peuvent être faits. D'une part, on se rend compte que les traces numériques dont Merzeau nous propose une typologie comme relevant autant des traces déclaratives (*ce que je dis de moi*), comportementales (*non intentionnelle grâce notamment aux cookies*), documentaires (*ce que je produis comme documents*) que de celles d'autrui (*le ouï-dire responsable notamment de notre e-réputation*), ne sont pas un *phénomène nouveau* comme on pourrait être tenté de le penser au premier abord, mais une mutation de la *trace analogique* (essentiellement sur son statut et sa nature) remettant en question certains des modèles conceptuels préétablis.

Le concept de trace analogique nous l'avons vu, fait appel à certaines terminologies relatives en l'occurrence au paradigme indiciaire de Carlo Ginzburg (2010)[84] avec la notion de petite quantité, au caractère historique donc relevant du passé, d'une certaine forme de témoignage au sens de Marc Bloch (1974)[85] et par extrapolation de sa localisation spatio-temporelle et comme empreinte ou marque convoquant l'idée de ce qui est donné à voir dans sa version brute et non comme construit. D'emblée, on peut noter, à la suite de Serres que dans le passage de la trace à la trace numérique, on est passé « *de l'ordre du donné à celui du construit, de la rareté à la surabondance, de l'indice au symbole, de la singularité à la calculabilité, de la localisation à la délocalisation, de la fixité à la volatilité, de la marque du passé à l'enregistrement du présent, de la conservation au dynamisme des traces...*[86]». Nous retiendrons ici à l'instar de toutes les propriétés inhérentes à la trace numérique (*relative à l'interaction homme-machine*) qu'elle est un construit, « *une séquence d'événements ordonnés (temporellement ou non)*[87]» au sens de Deransart.

---

[84] Carlo Ginzburg, « Traces. Racines d'un paradigme indiciaire », In *Mythes, emblèmes, traces : morphologie et histoire*, nouvelle édition augmentée, Verdier-Poche, Paris, 2010, pp. 218-294.

[85] Marc Bloch, *Apologie pour l'histoire ou Métier d'historien*, Armand Colin, Paris, 1974, p. 56.

[86] Serres Alexandre, « Problématiques de la trace à l'heure du numérique », *Sens-Dessous*, 2012/1 (N° 10), p. 84-94. DOI : 10.3917/sdes.010.0084. URL : https://www.cairn.info/revue-sens-dessous-2012-1-page-84.htm , p.27

[87] Pierre Deransart, « Traces, traces numériques, connaissances et cognition. Essai d'approche transversale », In *Contribution atelier IC2011*, 2011, p. 4. Disponible sur : http://pauillac.inria.fr/~deransar/ICAtelierIC1/introIC1-PD.pdf.

Ce qui fait sens si on part du postulat que les espaces numériques ont aboli ou du moins réduit les notions de distances autant informationnelles (« *la probabilité qu'un internaute se trouve confronté à une certaine information (un document, une idée, etc.), ce qui dépend de l'usager, de l'information et de l'ensemble des couches médiatrices*[88] ») que sociales (« *la probabilité que deux personnes communiquent ou interagissent* ») au sens de Rieder. Ceci implique de facto que tout *événement* (au sens d'action ou d'activité) ou *expérience* dans les environnements numériques se mue en donnée et laisse indéniablement des empreintes, que *tout fait trace* donc, au sens de Merzeau qui se réapproprie la maxime de l'école de Palo Alto « désormais, *on ne peut pas ne pas laisser de trace*[89] ».

Ces données qui oscillent entre requêtes, publications, échanges notamment via les réseaux de discussion et les plateformes de partage (*sharing*) peuvent être enregistrées, conservées, archivées, effacées, ou disparaître. La notion d' « effacement » est à appréhender au sens de Derrida comme n'étant pas définitive, car comme le souligne Sybille Krämer (2012)[90] *la trace une fois effacée laisserait donc place à la trace de cet effacement*, et donc ce concept est à distinguer très nettement de celui d' « effaçage » qui est volontaire, définitif et total. Ces traces qui sont disséminées à travers les espaces numériques, une fois collectées, stockées et traitées via notamment les algorithmes (ce qui soulève des questionnements notamment sur la *gouvernance* de ces données, le caractère déontologique du traitement des traces numériques par les algorithmes et des règles de conduite à suivre dans les pratiques informationnelles dans les espaces numériques au sens de Rieder, les données étant utilisées non à des finalités d'aide comme on serait tenté de le penser, mais plutôt dans une optique de surveillance[91])

---

[88] La notion de couches médiatrices joue ici un rôle central dans la mesure où le web étant un construit de pages dynamiques et non statiques, de liens et d'une pléthore de systèmes, la « distance utilisateur-information » se voit modifiée, distance qui pour Rieder se résume en partie aux nombres de clicks séparant les deux mais aussi et surtout dépend de plusieurs facteurs spatio temporels.

[89] Merzeau Louise, « *Du signe à la trace : l'information sur mesure* », *Hermès, La Revue*, 2009/1 (n° 53), p. 21-29. URL : https://www.cairn.info/revue-hermes-la-revue-2009-1-page-21.htm , p.4

[90] Sybille Krämer, « *Qu'est-ce donc qu'une trace, et quelle est sa fonction épistémologique ? État des lieux* », *Trivium* [En ligne], 10 | 2012, mis en ligne le 30 mars 2012, consulté le 24 mai 2017. URL : http://journals.openedition.org/trivium/4171

[91] La notion de surveillance de Rieder perçue au sens de Foucault (2004) comme pratique à visée « disciplinaire » (ce qui pour lui signifie le faite de vouloir « transformer un individu (potentiellement) désobéissant en « corps dociles »), est source de discrimination dans le sens où par la captation de nos gestes, idées, valeurs, etc., les algorithmes utilisés par les différentes plateformes nous cloisonnent dans des catégories ou cases préétablies, nous enjoignant à rester enfermé dans certains carcans et nous privant d'une certaine forme de « liberté ».

constituent donc la première couche médiane servant de base à la construction de l'e-réputation. On est donc tenté de s'intéresser aux autres *attributs* nécessaires dans le processus de formation de l'e-réputation.

2. La construction des identités numériques

De par les interactions que les individus entretiennent à travers l'utilisation notamment les services proposés par le web social, ils produisent de manière systématique, nous l'avons vu, des traces numériques. Ces dernières constituent des signes distinctifs rattachés à un ou plusieurs identifiants, permettant de caractériser un individu. La notion d'identification voir d'authentification au sens large (il est question autant des adresses IP, identifiants : ID/mot de passe, etc.) est ici essentielle, car c'est à majorité par elles que toutes les traces numériques laissées par un individu renvoient à la même entité (F. Georges, 2009).

Perçues comme « *la collection des traces (écrits, contenus audios ou vidéos, messages sur des forums, identifiants de connexion, etc.) que nous laissons derrière nous, consciemment ou inconsciemment, au fil de nos navigations sur le réseau et le reflet de cet ensemble de traces, tel qu'il apparaît « remixé » par les moteurs de recherche* », O. Ertzscheid (2013[92]) ou encore comme « *composée de signes à l'écran qui manifestent l'utilisateur* » F. Georges (2009), les identités numériques qui sont un concept à problématique polyfacettes[93] au sens de Nathalie Pinède (2018[94]) sont la résultante nous l'avons vu d'un calcul ou codage des activités des internautes à travers l'utilisation des services proposés par le web social (parfois à leur insu).

---

[92] ERTZSCHEID, Olivier. *Chapitre 1 - Les logiques identitaires* In : *Qu'est-ce que l'identité numérique ? Enjeux, outils, méthodologies* [en ligne]. Marseille : OpenEdition Press, 2013 (généré le 04 octobre 2016). Disponible sur Internet : <http://books.openedition.org/oep/405>. ISBN : 9782821813380. DOI : https://doi.org/10.4000/books.oep.405. p.6

[93] Dans la mesure où elles englobent une pluralité de champs et sont autant accolées aux concepts de présence, de territoire que d'attributs identitaires entre autres.

[94] Pinède Nathalie, « *Du site web aux identités numériques organisationnelles. Proposition d'un modèle d'analyse* », *Questions de communication*, 2018/2 (n° 34), p. 75-94. URL : https://www-cairn-info.ezproxy.u-bordeaux-montaigne.fr/revue-questions-de-communication-2018-2-page-75.htm

Cette problématique polyfacettes met en exergue dans un premier temps la question des identités numériques à la croisée de celles de "*territoire numérique ou de marque* (Cardon et al, 2015)" et de "présence numérique (Merzeau, 2009[95] ; Cordier Anne, 2010[96])", avec la prise en compte des différents "attributs identitaires" qui entrent en jeu (Georges, 2009 ; Ertzscheid, 2009[97] ; Coutant et Stenger, 2011[98]; Alloing et Pierre, 2012).

Quelles articulations existent donc entre les questions d'identités, territoire et présence numériques ? Les travaux de Merzeau (2010[99]) nous offrent une première piste de réponse. Elle nous propose cette approche de la présence numérique : « *Ni somme ni statut, la présence se déploie dans le temps : elle est irréversible et imprévisible, c'est-à-dire fondamentalement sociale, quand bien même les traces par lesquelles elle se manifeste sont traitées par des machines*[100] ». À la différence des identités numériques qui sont la somme de nos données, publications, contacts, etc., et qu'on subit quelque peu, la présence elle, qui n'induit aucun calcul se veut une représentation de nous à l'*infomonde* et donc imprévisible, représentation sur laquelle on doit avoir un droit regard.

Substituer la présence à l'identité suppose donc une certaine forme de démocratie co créative de notre présence avec en prime le droit de disposer, agencer, s'approprier, se réapproprier et gérer nos principales traces, notre historique dans les environnements numériques, avec notamment la mise en place comme le suggère Merzeau de *conventions collectives*. Cette vision quelque peu utopiste révèle la complexité des identités numériques comme agrégat

---

95 Merzeau L., « *Présence numérique : les médiations de l'identité* », *Les Enjeux de l'information et de la communication*, 2009/1 (Volume 2009), p. 79-91. DOI : 10.3917/enic.009.0079. URL : https://www.cairn-int.info/revue-les-enjeux-de-l-information-et-de-la-communication-2009-1-page-79.htm

96 CORDIER, A., « *Une présence numérique à multiples facettes*». *Documentaliste-Sciences de l'information*, vol.47, n°1, février 2010, p.54

97 Ertzscheid, O., « *L'homme est un document comme les autres : du World Wide Web au World Life Web* », *Hermes*, 53 (2009) p. 33-40. URL : http://archivesic.ccsd.cnrs.fr/sic_00377457/fr/

98 Coutant, A., et Stenger T., « Production et gestion d'attributs identitaires ». *Les Cahiers du numérique* Vol. 7, no 1 (1 juillet 2011): 61-74. URL : http://www.cairn.info/resume.php?ID_ARTICLE=LCN_071_0061.

99 Merzeau L., *La présence plutôt que l'identité*. Documentaliste - Sciences de l'Information, ADBS, 2010, 47 (1), pp.32-33. halshs-00489655v2. URL : https://halshs.archives-ouvertes.fr/halshs-00489655/document

100 Ibid, p. 32

autant du « *visible et de l'invisible, de l'intentionnel et de l'involontaire* » (Nathalie Pinède, 2018[101]) et permet de distinguer les identités numériques au sens de l'individu de celles qualifiées d'identités numériques organisationnelles au sens de Pinède.

Au sens de l'individu, la mise en évidence de cette pluralité d'identités numériques se trouve matérialisée par Fanny Georges (2009[102]) qui en distingue trois (3) principales sortes : celles déclaratives, agissantes et calculées. L'identité déclarative se dit de l'ensemble des "signes" utilisés par l'internaute dans sa présentation, celle agissante utilise les différentes activités de ce dernier et enfin l'identité calculée qui est une agrégation des variables quantifiées produites d'un calcul par le système (dans l'optique de prédiction).

Cette approche de l'identité numérique perçue comme « égocentrée » au sens de Alexandre Coutant et Thomas Stenger (2011)[103], est à distinguer, ils le notent, de celle dite "régalienne" prenant en compte les « identifiants socles » (nom, prénom, date de naissance, entre autres) ou signes distinctifs facilitant une identification, une reconnaissance, une forme de singularité. Dans la même lignée que celle relative à l'individu, l'identité numérique organisationnelle elle, procède d'une computation des traces numériques de celle-ci, mais *a contrario*, ne dénote pas d'un choix personnel de mise en avant de soi comme cela peut être le cas pour l'individu, mais plutôt d'un objectif de visibilité numérique comme une extension de sa présence sociale et à visée stratégique pour l'organisation avec des enjeux et acteurs différents de ceux de l'individu. On est donc tenté de questionner à ce stade l'élaboration des identités numériques pour les individus et les organisations.

La construction de l'identité numérique au niveau individuel part, nous l'avons mentionné, de la problématique du lien social, d'une quête de socialisation qui serait la transposition dans le monde numérique de besoins traditionnels pouvant se traduire par le choix d'un identifiant (comme besoin sécuritaire), l'utilisation des plateformes communautaires de partage,

---

[101] Pinède Nathalie, « *Du site web aux identités numériques organisationnelles. Proposition d'un modèle d'analyse* », *Questions de communication*, 2018/2 (n° 34), p. 75-94. URL : https://www-cairn-int-info.ezproxy.u-bordeaux-montaigne.fr/revue-questions-de-communication-2018-2-page-75.htm , p.14

[102] Georges F., Op.Cit, p.167

[103] Coutant, Alexandre, et Thomas Stenger. « *Production et gestion d'attributs identitaires* ». *Les Cahiers du numérique* Vol. 7, no 1 (1 juillet 2011): 61-74. http://www.cairn.info/resume.php?ID_ARTICLE=LCN_071_0061.

discussion, rencontres, etc. (relatif aux besoins sociaux ou d'appartenance), la construction et gestion de son e-réputation (relative aux besoins d'estime) et enfin, afin de matérialiser le besoin d'accomplissement de soi l'adéquation entre la réputation perçue et l'identité voulue avec en jeu les questions de notoriété, célébrité, etc.(Jacques Perriault, 2009[104] ; Olivier Ertzscheid, 2013 : 33). Ce besoin de visibilité se traduit donc par une exposition non contrôlée et un affichage de soi à travers les divers attributs identitaires qui oscillent entre adresse IP, *cookies*, courrier électronique, nom, prénom, pseudo, coordonnées (personnelles, administratives, bancaires, professionnelles, sociales), photos, avatars, logos, tags, liens, vidéos, articles, commentaires de forums, données géolocalisées, etc. (Ertzscheid, 2013 : 6) qui se voient corrélés à la myriade d'indices ou signes infinitésimaux autant furtifs, involontaires qu'insignifiants (Merzeau, 2010) issus des pratiques informationnelles et perçues comme sources premières de la construction de l'identité numérique.

En effet, loin d'être construites à partir des traces déclaratives (qui du fait qu'elles sont volontaires sont embellies, déguisées parfois travesties), les identités numériques sont à majorité composées des traces agissantes et calculées (Georges, 2009) et *« s'inscrivent dans et résultent d'un système d'interactions complexes, hyper dynamiques, imbrication sans cesse renouvelée de signes délibérés, de comportements, de traces calculées et reconfigurées, où actions individuelles, sociales, artefactuelles et algorithmiques redessinent en continu les contours d'une « identité », d'une « présence » dans l'univers numérique» (Pinède, 2018 : 20)* . Si cette configuration des traces est essentielle, c'est qu'elles participent (les traces) d'une forme de (re)documentarisation (Salaün, 2007; Ertzscheid, 2009[105]; 2013[106]) de l'être humain perçu comme « objet documentaire » n'ayant que peu de maîtrise sur « *ses traces numériques* » et sa visibilité.

---

[104] Perriault J., « *Traces numériques personnelles, incertitude et lien social* », *Hermès, La Revue*, 2009/1 (n° 53), p. 13-20. URL : https://www.cairn.info/revue-hermes-la-revue-2009-1-page-13.htm

[105] Ertzscheid O., « *L'homme, un document comme les autres* », *Hermès, La Revue*, 2009/1 (n° 53), p. 33-40. URL : https://www.cairn-int.info/revue-hermes-la-revue-2009-1-page-33.htm

[106] ERTZSCHEID, O., *Qu'est-ce que l'identité numérique ? Enjeux, outils, méthodologies*. Nouvelle édition [en ligne]. Marseille : OpenEdition Press, 2013 (généré le 04 octobre 2016). Disponible sur Internet : <http://books.openedition.org/oep/332>. ISBN : 9782821813380. DOI : https://doi.org/10.4000/books.oep.332.

Si on part de notre postulat de départ selon lequel l'identité numérique est un agrégat des traces numériques autant profilaires, navigationnelles, que déclaratives disséminées sur la toile, et qu'elle relève autant de l'individu lui-même que d'autrui, toutes les interactions, requêtes, avis, évaluations, logs, etc. aussi furtifs ou infimes soient-ils apparaissent computés, enregistrés, (re)diffusés, remixés et parfois reformulés. Grâce au caractère indexable de toutes ces données et au principe de *searchability* (favorisant la recherche d'un individu, Boyd, 2007), l'individu se mue en document numérique, documenté continuellement au fil de ses navigations et qui ne fait que croître.

Cette économie sociale informationnelle et des documents nous renvoie à l'idée de réputation comme construite par autrui et désormais, par les diverses plateformes numériques utilisées.
Les identités numériques individuelles s'élaborent donc à la croisée de deux axes avec des stratégies identitaires diverses et variées alliant visibilité partielle/ totale, tentatives d'anonymat ou parfois hypervisibilité (avec l'utilisation des avatars, des pseudonymes, etc.), et supposent de fait une dualité entre l'être et l'agir ; entre "*ce que je suis ou ce que je veux être*" *a priori* au sein de ces espaces numériques.

Au sens de l'organisation, l'identité ou la présence numérique se construit d'une part au travers des plateformes corporate de production (sites web, blogs) et les réseaux sociaux numériques des institutions avec le site web comme attribut privilégié (Alloing, 2016; Pinède, 2018) et est par la suite exportée et « *remixée* » via notamment les espaces d'interactions et d'échanges (forums entre autres) de leurs divers publics. Il est important de noter ici le caractère complexe des identités numériques organisationnelles comme agrégat à la fois de l'image de l'organisation, de ses diverses marques (dans certains cas), mais aussi et surtout de la pluralité des individus qui composent l'organisation et la représente, l'image des ces derniers pouvant influer sur celle de l'organisation.

3. La construction de la confiance numérique

3.1. La confiance : un concept abstrait, complexe et multidimensionnel…

Considéré comme un concept important pour les échanges en général et les relations commerciales en particulier entre particuliers, mais également auprès des entreprises en *b2b* (Clark and Lee 1999 ; Dellarocas 2003 ; Grandison & Sloman 2000 ; McKnight & Chervany 2001 ; Sabater and Sierra 2005), le concept de confiance en dépit de son importance avérée, se révèle être une notion à la fois complexe (McKnight and Chervany 2001), abstraite (Schwieters 2015), dynamique (Botsman, 2015, 2016) et confuse (Shapiro 1987) tant sa définition n'est pas stabilisée.

Le Centre National de Ressources Textuelles et Lexicales définit la confiance comme *« la croyance spontanée ou acquise en la valeur morale, affective, professionnelle... d'une autre personne, qui fait qu'on est incapable d'imaginer de sa part tromperie, trahison ou incompétence »* [107]. Tiré du latin confidentia qui lui-même est dérivé du verbe confidere (confier), la notion de confiance signifie donc "*avoir foi en quelque chose ou en quelqu'un*", *avoir une certitude de la loyauté d'autrui*, *accorder du crédit à quelqu'un ou quelque chose*, etc. Cette première proposition se voit renforcée par celle de Gambetta (2000) dont la définition fait partie des plus reprises dans la littérature et qui la perçoit comme suit : « *Trust (or, symmetrically, distrust) is a particular level of the subjective probability with which an agent assesses that another agent or group of agents will perform a particular action, both before he can monitor such action (or independently of his capacity ever to be able to monitor it) and in a context in which it affects his own action* »[108].

---

[107] Définition disponible sur https://www.cnrtl.fr/definition/confiance

[108] Gambetta, D., *Trust: Making and Breaking Cooperative Relations, chapter Can We Trust Trust?*, pages 213 – 237. Department of Sociology, University of Oxford, 2000. Traduit par nos soins : " La confiance (ou symétriquement la méfiance) est un état particulier de la probabilité subjective selon laquelle un agent évalue qu'un autre agent ou groupe d'agents effectuera une action particulière, à la fois avant qu'il ne puisse contrôler cette action (ou indépendamment de sa capacité à être apte à le surveiller) et dans un contexte où il cela affecte sa propre action"

La confiance qui ici est une forme de transaction, met en scène la probabilité qu'un agent effectue une action tel qu'initialement prévu par un autre agent ou groupe d'agents, et ce, en prenant en compte un contexte singulier. Il est intéressant de noter l'apparition de notions telles celles de dépendance à un contexte (Sabater et Sierra, 2005, Grandison & Sloman, 2000) et de prédiction. À partir de cette proposition de définition, d'autres auteurs présenteront des variations en y intégrant notamment les concepts de risque et d'incertitude comme *output* négatifs (Luhmann, 2000 ; Castelfranchi & Falcone, 2001 ; Laeequddin et al. 2010 ; McKnight and Chervany 2001) ; de croyances et dépendance au sens de Josang (1996) ; de délégation, sécurité et honnêteté (Castelfranchi & Falcone, 2001) ; d'aptitude/ compétence et d'intégrité pour McKnight et al. (2002)[109] et Marsh (1994)[110].

En dépit de ces diverses approches, s'il est une chose sur laquelle tous s'accordent, c'est que la confiance est un concept multidimensionnel (Kini and Choobineh 1998; McKnight & Chervany 2001; Sabater and Sierra 2005) et qu'elle est la résultante d'une relation entre ceux que les auteurs qualifient de "*truste*r" et de "*trustee"* (figure 1) avec pour *truster* une entité en quête d'un service (au sens général du terme) et qui de fait établit une relation de confiance avec le *trustee* qui est supposé proposer ou être à l'origine du service en service en question (Laeequddin et al. 2010).

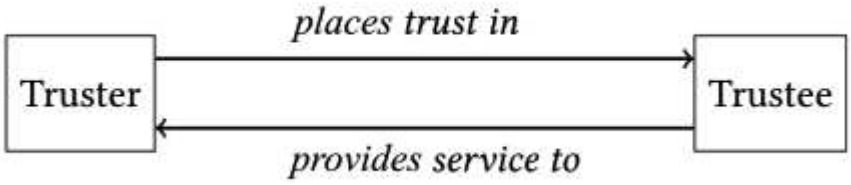

Figure n°3 : La relation entre le truster et le trustee, source Grandison & Sloman (2000)

[109] D. H. McKnight, V. Choudhury, and C. Kacmar. *Developing and validat- ing trust measures for e-commerce: An integrative typology*. Information Systems Research, 13(3):334 – 359, September 2002

[110] S. P. Marsh. *Formalising Trust as a Computational Concept*. PhD thesis, Department of Mathematics and Computer Science, University of Stirling, Scotland, UK, 1994

Il est intéressant de noter ici à la suite de Grandison & Sloman (2000) que cette relation qui s'établit entre le *truster* et le *trustee* n'a ni spécialement besoin d'être interpersonnelle, mais peut prendre en compte une pluralité d'agents selon la situation, ni encore d'être symétrique. Si la confiance résulte de la relation entre deux ou plusieurs entités et qu'elle englobe autant les concepts de risque, certitude/ incertitude, délégation, compétence, etc., on serait tenté de se questionner sur ces divers termes. Quelles articulations possibles existent donc entre ces diverses approches et le concept de confiance?

3.2. La confiance comme concept composite : entre risque, certitude et délégation

Le challenge que posent les nouveaux environnements et plus spécifiquement ceux qualifiés de systèmes multi agents, est d'arriver à trouver les informations et partenaires fiables. Par système multi agents, on entend tout simplement ceux composés d'agents autonomes appartenant à divers groupes avec divers centres intérêts qui peuvent intégrer et quitter le système à tout moment et qui interagissent dans l'optique d'atteindre leurs objectifs, objectifs qui passent entre autres par des stratégies de coopération, coordination, négociation et résolution de conflits (Wooldridge, 2009[111]).

Afin de réduire l'incertitude dans les interactions, certains auteurs (Marsh, 1994 ; Ramchurn et al., 2004 ; Rehak et al., 2005 ; Wang et al., 2008 ; Burnett et al., 2011 ; Venanzi et al., 2011) proposent des mécanismes (à partir de méthodes probabilistes, de l'apprentissage bayésien et de la théorie des jeux) de mesure des attitudes et compétences des agents, mécanismes basés essentiellement sur le degré de fiabilité de ces derniers. Plus un agent est honnête dans ses interactions avec les autres agents, plus ces derniers seront enclins à le "*recommander*" réduisant pour les nouveaux agents le risque d'entrer en relation avec lui et *a contrario*, moins il le sera (honnête), plus la relation présentera de risque pour ses potentiels partenaires.

---

[111] Michael J. Wooldridge: *An Introduction to MultiAgent Systems*, Second Edition. Wiley 2009, ISBN 978-0-470-51946-2, pp. I-XXII, 1-461

On le voit bien, confiance, certitude et risque sont des concepts imbriqués, et la confiance est nécessaire, voire indispensable, dans des situations et environnements qu'on qualifierait de risqués et où les agents ont quelque chose à perdre (Aljazzaf et al., 2010[112]). La gestion des risques est essentielle dans les modèles de confiance, mais aussi de réputation afin d'appréhender les environnements changeants (Lu et al., 2009[113]). L'évaluation du facteur risque en amont des interactions avec les agents d'un système devient alors un enjeu majeur et plus encore dans les situations où s'invite la délégation de tâches.

Les notions de confiance et délégation dans les systèmes ouverts ont fait l'objet de nombreuses études (Carter & Ghorbani 2003[114] ; Neville & Pitt 2004[115] ; Venanzi & al. 2011[116]; Piunti & al. 2012[117], etc.) et ont été introduits par Castelfranchi et Falcone (1998)[118] pour qui les deux notions sont intrinsèquement liées et pour qui la confiance est nécessaire voire indispensable dans la délégation d'une *tâche*. Pour faire confiance selon eux, il est nécessaire d'avoir des objectifs et convictions accolés à des tâches et d'être en mesure de déléguer ces tâches à un agent (celui en qui on est censé accorder sa confiance). Partant de ce constat, on est tenté de se questionner sur les mécanismes cognitifs et/ ou convictions intrinsèques garantissant une forme de contrat implicite sur la fiabilité de l'agent à qui la

---

[112] Aljazzaf, Zainab & Perry, Mark & Capretz, Miriam. (2010). *Online Trust: Definition and Principles*. Proceedings - 5th International Multi-Conference on Computing in the Global Information Technology, ICCGI 2010. 163 - 168. 10.1109/ICCGI.2010.17. Accessible en ligne https://www.researchgate.net/publication/224193095_Online_Trust_Definition_and_Principles

[113] Lu, Y., Zhao, L., & Wang, B. (2010). *From virtual community members to C2C e-commerce buyers: trust in virtual communities and its effect on consumers' purchase intention*. Electronic Commerce Research and Applications, 9(4), 346e360.

[114] J. Carter and A. A. Ghorbani. 2003. *Value centric trust in multiagent systems*. In Proceedings of IEEE/WIC International Conference on Web Intelligence (WI 2003)

[115] B. Neville and J. Pitt. 2004. *A computational framework for social agents in agent mediated e-commerce*. In Proceedings of the Seventh International Workshop on Trust in Agent Societies, R. Falcone, S. Barber, L. Korba, and M. Singh, Eds., New York, July 2004, 83–91.

[116] M. Venanzi, M. Piunti, R. Falcone, and C. Castelfranchi. 2011. *Facing openness with socio cognitive trust and categories*. In Proceedings the 22nd International Joint Conference on Artificial Intelligence (IJCAI), Barcelona, Spain, 16-22 Jul 2011. AAAI Press, pp. 400–405

[117] M. Piunti, M. Venanzi, C. Castelfranchi, and R. Falcone. 2012. *Multimodal trust formation with uninformed cognitive maps (UnCM)*. In Proceedings the 11th Conference on Autonomous Agents and MultiAgent Systems 3, pp. 1241–1242

[118] C. Castelfranchi, Cristiano, et Rino Falcone. « *Principles of trust for MAS: Cognitive anatomy, social importance, and quantification* ». In *Proceedings International Conference on Multi Agent Systems (Cat. No. 98EX160)*, 72-79. IEEE, 1998.

tâche est déléguée. Castelfranchi et Falcone identifient donc trois mécanismes à prendre en compte dans une optique de délégation soit la croyance en matière de compétence (croire en l'aptitude d'un agent à effectuer une tâche), la croyance en matière de dépendance (choisir de s'en remettre à un autre agent pour effectuer une tâche donnée) et enfin celle en matière de disposition[119] qui sous-tend une croyance infaillible dans le fait que l'autre agent accomplira la tâche et qu'il le fera tel qu'elle doit être faite.

Dès lors que ces trois conditions sont remplies, s'établit une certaine forme de contrat (implicite) dans lequel l'agent décide de ne pas renoncer à ses objectifs, de ne pas les atteindre personnellement, de n'avoir d'autre alternative que l'agent à qui il délèguera la tâche et qui lui permettra d'atteindre son objectif. On assistera dès lors à l'apparition de mécanismes visant à choisir le partenaire adéquat dans une situation de délégation que certains auteurs se sont donc attelés à formaliser.

C'est notamment le cas de N. Griffiths (2006[120]) qui propose une grille avec d'une part les compétences ou appétences de tous les agents et d'autre part les différentes tâches à exécuter. L'objectif est donc, à partir des préférences et compétences renseignées par chaque agent et qui sont évaluées, de choisir le partenaire adéquat. On parlera plutôt de coopération entre deux agents au sens de Grandison et Sloman (2000[121]), ce qui implique une relation gagnant-gagnant et donc une certaine forme de compensation à l'agent à qui la tâche est déléguée.

---

[119] Il est essentiel ici de prendre en compte la volonté et l'intention de l'agent *à* accomplir cette tâche mais aussi et surtout la persistance, la stabilité de ce dernier à aller au bout de sa démarche

[120] N. Griffiths. 2005. *Task delegation using experience-based multi-dimensional trust*. In Proceedings of the 4th International Joint Conference on Autonomous Agents and MultiAgent Systems (AAMAS'05), ACM Press, pp. 489–496

[121] T. Grandison and M. Sloman. 2000. *A survey of trust in internet applications*. IEEE Communications Surveys and Tutorials 4, 4, 2–16.

Cette compensation, qui dépend de la finalité de la tâche effectuée, doit être décidée en amont par l'agent qui souhaite déléguer une tâche (Burnett et al., 2011[122] ; Chopra et al., 2011[123]) et vise généralement à attribuer une polarité (digne de confiance ou non, honnête, fiable ou non, etc.) à l'agent qui effectuera la tâche, mention qui, influera sur sa réputation.

Le concept de délégation avec notamment le premier mécanisme de Castelfranchi et Falcone, soit la croyance en matière de compétence, implique indéniablement les facteurs de *compétence* et de *contexte* que s'attachent à souligner certains auteurs (Grandison et Sloman, 2000; Sabater & Sierra, 2005, Zhong & al., 2015). En effet, et comme le soulignent Grandison & Sloman " *A trustor trusts a trustee with respect to its ability to perform a specific action or provide a specific service within a context*[124]*"*. Afin d'illustrer ce fait, Sabater & Sierra prennent l'exemple d'un médecin qui, de par son expertise, sera perçu comme quelqu'un de confiance et de référence dans le cadre d'une prescription ou d'un traitement médical, mais *a contrario*, à qui on ne fera pas appel pour le choix d'une bouteille de vin (ce domaine ne faisant pas partie de ses compétences). Zhong et al. (2015) vont même au-delà de cette perception et proposent de distinguer la compétence (ce qu'ils qualifient de *trust in expertise*) de l'intégrité (ce qu'ils appellent le *trust in goodwill*[125]), avec le *trust integrity* comme facteur ne dépendant pas du contexte.

Le concept de confiance recouvre donc une sémantique plurielle, et si on s'arrêtait là, on serait tenté de postuler qu'il existe plusieurs types de confiance dans une situation d'interaction donnée dont Grandison et Sloman (2000)[126] nous proposent une spécification

---

[122] C. Burnett, T. J. Norman, and K. Sycara. 2011. *Trust decision-making in multi-agent systems*. In Proceedings of the Twenty Second International Joint Conference on Artificial Intelligence. 115–120.
Ils identifient cinq mécanismes au choix dont la délégation simple, la délégation surveillée, celle non surveillée, la délégation avec incitement ayant une répercussion sur la réputation et enfin le retrait de la délégation

[123] A. M. Chopra, E. Paja, and P. Giorgini. 2011. *Sociotechnical trust: An architectural approach*. In Proceedings Conference Conceptual Modeling, LNCS 6998, pp. 104–117.
Les auteurs distinguent quatre principales stratégies : la délégation, la compensation, l'annulation et la renégociation

[124] Grandison, T., & Sloman, M. (2000). *A survey of trust in internet applications*. IEEE Communications Surveys & Tutorials, 3(4), 2–16. doi:10.1109/comst.2000.5340804; p.3. Traduction par nos soins : " *Une entité en quête d'un service (au sens général du terme) fait confiance à une autre sur sa capacité, son aptitude à effectuer une action spécifique ou à fournir un service spécifique dans un contexte donné*".

[125] traduit par nos soins :" la bonne volonté"

[126] Grandison, T., & Sloman, M. (2000). *A survey of trust in internet applications*. IEEE Communications Surveys & Tutorials, 3(4), 2–16. doi:10.1109/comst.2000.5340804. https://scihub.wikicn.top/10.1109/COMST.2000.5340804

sémantique à cinq (5) niveaux : le *service provision trust*, le *resource access trust*, le *delegation trust*, le *certification trust* et l'*infrastructure trust* à partir desquels Josang et al. (2007, p.11)[127] nous proposent le schéma ci-dessous :

- La notion de ***service provision trust*** qualifiée de *business trust* (terme employé par The Liberty Alliance Project[128]) décrit la confiance accordée à un fournisseur de services ou ressources dans un contexte commercial d'incertitude où un utilisateur cherche à identifier et éviter les fournisseurs malveillants. Ce type de confiance nécessite dans certain cas, une forme de contrat spécifiant les exigences de la prestation de service[129];
- Le ***resource access trust*** traité notamment dans les travaux de Grandison & Sloman (2000) fait référence aux ressources proposées par les fournisseurs et notamment à la garantie de leur accessibilité aux utilisateurs[130];
- Par ***delegation trust*** qui induit la notion de délégation, il s'agit de la confiance placée en un individu qui agit et prend des décisions au nom d'un autre individu (avec notamment sa capacité à restituer dans leur conformité les actions qui lui sont déléguées)[131];
- Le concept d'***identity trust*** (ou certification trust pour Grandison et Sloman ou encore authentication trust pour Liberty Alliance, 2003) qui a principalement fait l'objet d'études dans les travaux en sécurité de l'information (avec des schémas

---

[127] Josang A., R. Ismail, and C. Boyd. A survey of trust and reputation systems for online service provision. Decision Support Systems, 43(2):618 - 644, March 2007 . Accessible en ligne sur https://www.oasis-open.org/committees/download.php/28303/JIB2007-DSS-Survey.pdf

[128] Accessible sur http://www.projectliberty.org

[129] Traduit par nos soins de l'anglais " Provisiontrust describes the relying party's trust in a service or resource provider. It is relevant when the relying party is a user seeking protection from malicious or unreliable service providers.The Liberty Alliance Project uses the term"business trust" to describe mutual trust between companies emerging from contract agreements that regulate interactions between them, and this can be interpreted as provision trust. For example when a contract specifies quality requirements for the delivery of services, then this business trust would be provision trust in our terminology."

[130] Traduit par nos soins " describes trust in principals for the purpose of accessing resources owned by or under the responsibility of the relying party. This relates to the access control paradigm which is a central element in computer security. A good overview of access trust systems can be found in Grandison & Sloman (2000)"

[131] Traduit par nos soins " describes trust in an agent (the delegate) that acts and makes decision on behalf of the relying party. Grandison & Sloman point out that acting on one's behalf can be considered to a special form of service provision"

d'authentification tels que X.509 et PGP et dont on retrouve un aperçu dans Reiter & Stubblebine (1997), désigne la croyance en la véracité de l'identité d'un utilisateur sur les plateformes numériques;

- Et enfin le ***context trust*** (ou *infrastructure trust* ou encore *system trust* pour McKnight & Chervany, 1996) qui implique un système de sécurité général visant à garantir la sécurité de toutes les parties prenantes en cas de problème et qui va du développement d'infrastructures solides de détection de fraude entre autres, aux assurances, etc.[132]

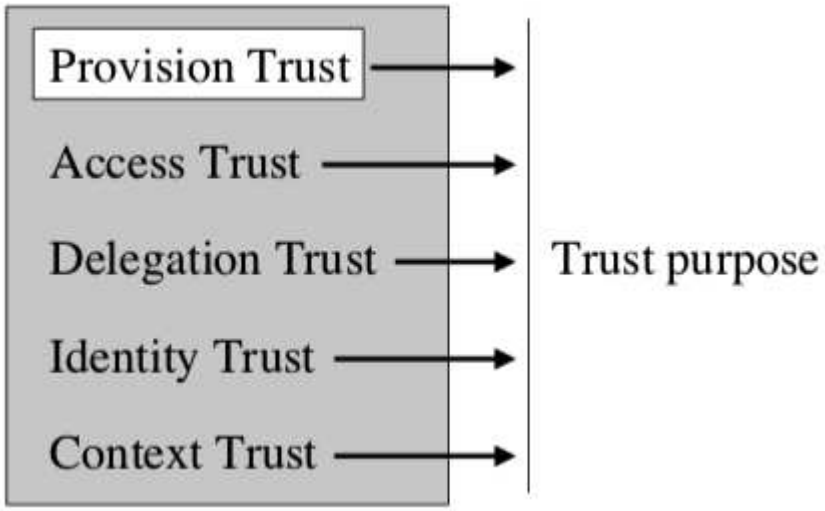

Figure n°4 : Les cinq typologies de confiance, source Josang al. (2017) à partir des travaux de Grandison et Sloman

Le concept de confiance, terme générique qui définit la portée spécifique d'une relation peut donc être utilisé pour exprimer divers attributs dans la relation à autrui. La notion d'identité (avec l'*identity trust*) est ici essentielle si tant est qu'elle soit avérée et vérifiée, car c'est en partie par elle que se construit la relation de confiance.

[132] Traduit par nos soins " describes the extent to which the relying party believes that the necessary systems and institutions are in place in order to support the transaction and provide a safety net in case something should go wrong. Factors for this type of trust can for example be critical infrastructures, insurance, legal system, law enforcement and stability of society in general."

Ces diverses observations nous permettent de recouper le concept de confiance sous deux principales typologies : la confiance initiale (subjective) et celle construite (objective).

- Par subjectif, le Larousse en ligne entend ce « *qui relève du sujet défini comme être pensant, comme conscience individuelle, par opposition à objectif* »[133] auquel est accolée la notion de variation selon la personnalité de chacun. Cette première proposition de définition se voit renforcée par celles que propose le CNRTL[134] et dont nous retiendrons qu'elle se construit *sans influence externe* et qu'elle est dénuée d'objectivité et de scientificité. Penser la confiance de ce point de vue en la dénuant de toute exigence autant méthodologique que de précision revient à s'intéresser de très près à l'affect et aux états mentaux mobilisés dans une situation de confiance.

  Ce type de confiance se manifeste lors de ce qu'on nommera *première interaction* ou *interaction initiale* entre deux agents dans l'optique où il n'existe aucun autre antécédent (interaction au sens large : avis, commentaire, note, etc.) pouvant servir de base d'évaluation de la fiabilité ou non d'un agent. En effet, lorsqu'un individu *A* décide d'interagir avec un individu *B* lors d'une transaction (au sens général du terme : question, transaction marchande, etc.), plusieurs questionnements relatifs notamment à l'objet de la transaction, à la compétence de l'agent à proposer le produit ou service en question, à son honnêteté lors de la transaction, etc. se posent. Dans l'optique où aucune donnée n'est disponible sur les antécédents de l'agent *B,* l'individu *A* se voit dans l'obligation de se référer à une forme d'*intuition* sur la base des données mises à disposition (données personnelles, description d'un produit ou service avec coûts, délais de livraison, etc. s'il en propose, référence, etc.) par le système au sein duquel se trouvent les individus et renseignées par *b*.

- Le second état, la confiance qu'on qualifiera d'objective, loin d'être antagoniste à la première comme on serait tenté de le penser, se veut complémentaire dans l'optique où ce type de confiance implique implicitement qu'il ait eu au préalable une première interaction. Cette seconde typologie, un peu plus complexe fait appel à une certaine

[133] Accessible en ligne https://www.larousse.fr/dictionnaires/francais/subjectif/75038

[134] Accessible en ligne https://www.cnrtl.fr/definition/subjectif

forme de scientificité dans la mesure où la décision d'interagir ou de faire confiance à un agent lors d'une transaction repose sur le calcul, calcul portant tant sur les *compétences* de l'agent qui propose le produit ou service, que sur son *honnêteté* et sa faculté à agir de la manière la plus *fortuite* possible afin de satisfaire les besoins d'un autre agent ce que McKnight et al. (2002)[135] qualifient de « *trusting beliefs* » et qu'ils distinguent du « *trusting intention* ». La confiance ici se construit à travers le temps, la fréquence des échanges, les routines, et est basée sur une forme de *contrat social* et la mise en place d'intermédiaire par le système de *délégation* avec pour visée première la recherche d'une certaine forme de *sécurité* (Grandison et Sloman , 2001 ; Josang et al. , 2007[136]) afin de réduire les éventuels effets négatifs liés au caractère incertain des interactions via les espaces numériques.

Partant du constat que la confiance est un construit (à la fois subjectif lors d'une première interaction puis objectivement en agrégeant toutes les données) au fil du temps, des échanges et routines, nous nous intéresserons ici aux différents mécanismes concourant à sa construction.

3.3. Les interactions au coeur de la construction de la confiance

Penser la confiance comme « construit » soulève dans une première optique la question de l'interaction, mais aussi et surtout celle des données, des sources d'informations utilisées qui dans les systèmes peuvent être divers et variés dépendant tant de l'architecture même des mécanismes que des agents en présence. L'une des sources d'information les plus courantes qu'on retrouve de manière plus ou moins systématique (cumulée à plusieurs autres dans la plupart des cas) est celle qu'on qualifiera *d'interaction directe*. Par interaction directe, on entend l'échange (au sens large : transaction, requête, etc.) direct entre deux agents *X* et *Y* où

---

[135] D. H. McKnight, V. Choudhury, and C. Kacmar. *Developing and validat- ing trust measures for e-commerce: An integrative typology*. Information Systems Research, 13(3):334 – 359, September 2002

[136] A. Josang, R. Ismail, and C. Boyd. *A survey of trust and reputation systems for online service provision*. Decision Support Systems, 43(2):618 – 644, March 2007

X à la possibilité, une fois la transaction terminée, d'évaluer *Y* sur la base de sa propre expérience avec lui. Marsh (1994) qui fut le premier à intégrer cette source d'information et qui d'ailleurs n'utilise que celle-ci dans son modèle de confiance introduit la notion de réciprocité qui influe sur les valeurs attribuées à la confiance et qui, ici, sous-entend que "*si par exemple, un agent X dans le passé a aidé un agent Y et que ce dernier ne le faisait pas en retour, cela réduira la confiance de l'agent X envers le second et inversement*", les obligeant *de facto* à garantir une fiabilité et honnêteté dans leurs échanges.

Cette propension à n'utiliser que les interactions directes ne se retrouve que chez très peu d'auteurs. C'est le cas de Griffiths (2005) pour qui dans une optique de délégation de tâche une interaction directe est indispensable, et de Ghanea-Hercock (2007) dont le mécanisme vise à sélectionner les partenaires adéquats en vue de former une coalition. Cela peut s'expliquer par le caractère même des systèmes ouverts qui rend quasi impossible d'établir des relations avec tous les agents du système. Même si l'interaction directe, nous l'avons noté, est considérée comme la plus pertinente des sources d'information, il est cependant intéressant de noter à la suite de Sabater et Sierra (2005) qu'il est nécessaire de disposer de données plutôt exhaustives (ce qui nécessite plusieurs interactions avec le même agent) afin de garantir une fiabilité de la valeur de confiance d'où le recours à d'autres sources d'information dans la construction de la confiance d'un agent.

Au compte de ces dernières, on retrouve *l'observation directe* qui ne nécessite pas d'interaction au sens premier entre les agents comme cela peut être le cas pour le premier type sus-cité, mais qui se fonde plutôt sur l'observation des interactions passées et présentes des agents afin d'évaluer leur comportement (Sabater et Sierra , 2005). L'observation ici portera tant sur des caractéristiques visibles et facilement accessibles telles que le nombre de retours, le pourcentage d'avis positifs/ négatifs, les actions passées, la finalité des interactions, etc. (Carter and Ghorbani, 2003 ; Rettinger et al., 2008 ; Sierra and Debenham, 2005 ; Klejnowski et al., 2010) que sur celles plus techniques faisant appel parfois à des techniques de machines learning comme l'analyse du langage communicationnel utilisé par l'agent sur lequel porte l'évaluation (Serrano et al, 2012).

Outre ces sources, l'attribution d'un score de confiance peut prendre en compte les *témoignages*. Les informations provenant de témoins sont utilisées lorsque les agents ne disposent d'aucune information directe et doivent donc se référer aux autres agents du système, ce qui implique indéniablement que ces derniers aient à minima une expérience directe avec l'agent en question. Qualifiées d'information indirecte ou plus communément de e-réputation (Sabater et Sierra , 2005), les informations provenant de tierces parties ou témoignages sont calculés à partir de l'agrégation des opinions des agents ayant déjà interagi avec l'agent dont on cherche à évaluer la confiance et peuvent soit prendre la forme de simples réponses à une requête sans sélection au préalable des meilleures sources (comme cela peut être le cas dans les modèles de Josang & Ismail, 2002 ou Teacy et al., 2006) ou *a contrario* prendre la forme de recommandations provenant uniquement des évaluations d'agents ayant déjà interagi avec la cible et dont la transaction est vérifiée (Yu and Singh, 2003). Il est intéressant de noter la mise en place dans le second cas d'une forme de boucle de réciprocité où les agents qui recommandent sont à leur tour évalués sur la fiabilité et l'honnêteté de leur avis (Derbas et al., 2004 ; Wang et Zhang, 2005 ; Abdul-Rahman et Hailes , 2005), mesure qui vise à réduire les informations frauduleuses au sein des systèmes ouverts.

À l'issue de chaque transaction, chaque agent dispose d'un système de ranking (classement) propre à chaque mécanisme pour chaque partenaire avec lequel il interagit, ranking qui prend en compte plusieurs attributs classés selon leur importance pour l'agent qui évalue. Ces attributs ou filtres concernent tant le produit ou service et peuvent porter sur sa qualité, les délais de livraison, le prix, etc. (Sabater et Sierra, 2002 ; N. Griffiths, 2006 ; Koster et al., 2011 ; Su et al., 2013) dans le cadre d'une transaction notamment via les plateformes de e-commerce par exemple, que sur les agents eux-mêmes : disponibilité, réactivité, fiabilité, utilité (Matt et al., 2010 ; Klabi et al., 2012) dans le cadre d'une réponse à une requête dans un forum de discussion par exemple. Ils offrent la possibilité dans certains cas de hiérarchiser les données selon la date (récente ou non).
Le résultat de ces différentes transactions permet donc d'attribuer un degré de confiance représenté par une ou plusieurs valeurs pouvant osciller entre plusieurs niveaux : très fiable, fiable, pas digne de confiance et pas du tout fiable (Abdul-Rahman et Hailes, 2005).

Afin d'illustrer cette mesure, Josang et al. (2007[137]) nous proposent un schéma directeur de la mesure de la confiance selon quatre (4) principales dimensions : subjectivité (qui implique un jugement subjectif d'un individu), objectivité (qui sous-entend que la mesure a été déterminée en évaluant objectivement la confiance par rapport à des critères formels), spécificité (lié à un seul aspect : exemple réactivité, délais de livraison,etc.) et généralité (impliquant la moyenne de tous les aspects sur lesquels doit porter la mesure de la confiance).

	Specific, vector based	**General, synthesised**
Subjective	Survey questionnaires	eBay, voting
Objective	Product tests	Synthesised general score from product tests, D&B rating

Figure n°5 : Les dimensions de mesure de la confiance, source : classification of trust measure, josang et al. (2007)

- Sur l'axe subjectif et spécifique, la mesure de la confiance qui peut utiliser des questionnaires individuels permet de collecter l'opinion d'un individu sur un aspect particulier de l'objet sur lequel porte la mesure avec pour résultats une échelle variable entre 1 et 5 représentant respectivement l'avis selon qu'il est "désastreux=1", "mauvais=2", "moyen=3", "bon=4" ou "excellent=5";

- L'axe subjectif et général mesure la confiance d'un point de vue subjectif de l'objet sur lequel porte l'évaluation dans sa globalité (à titre d'exemple, si un client reçoit un article acheté sur un site de e-commerce en retard par rapport aux délais prévisionnels ou abimés, il donnera indéniablement une note négative au vendeur, cependant ces aspects (retard dans les délais et article abîmé peuvent être causés par le transporteur et non le vendeur ce qui peut biaiser les données);

[137] Ibid, p.12

- L'axe objectif et spécifique est utilisé dans le cas de tests avérés sur l'objet de mesure (qualité, performance, etc.) et à l'avantage d'être basé sur des données vérifiables et accessibles par tous (on peut prendre l'exemple d'un classement d'une organisation par rapport à ses concurrents qui peut être basé sur son CA vérifiable ou son profit, etc.);

- L'axe objectif et général est donc l'agrégation de toutes les valeurs objectives et spécifiques en un score général de confiance (le système de *ranking* de Dunn & Bradstreet's[138] avec leur indice D&B direct pour la conformité et à son nuage de données en est exemple concret, car il fournit des données commerciales et financières, à la fois analytiques et synthétiques, sur les entreprises visant tant à identifier, vérifier, filtrer et surveiller les agents).

4. La mesure de l'influence numérique

Étymologiquement, le terme influence vient du latin « influentia » lui-même dérivé du verbe « influere » (influer) et signifie à travers les époques une *« sorte d'écoulement, de flux censé provenir des astres et agir sur les hommes et les choses »*[139] et relatif à l'influence divine dans les années 1240 ; au XIIIe siècle *« l'action (généralement lente et continue) d'un agent physique (sur quelqu'un, quelque chose), suscitant des modifications d'ordre matériel » ; « l'autorité ou le prestige»* dans les années 1780 *;* en 1793 *« l'autorité politique ou intellectuelle acquise par un pays, une civilisation... à une époque donnée »*[140] et dans les années 1840 *« l'effet produit à distance»* notamment relatif à l'influence électrique ou physique *»* [141].

---

[138] Système de ranking accessible en ligne https://www.dnb.com/ca-fr/

[139] Définition du CNRTL. En ligne https://www.cnrtl.fr/definition/influence

[140] Ibid.

[141] Ibid.

Nous nous intéresserons dans cette partie à l'influence numérique telle que perçue et appliquée à la mesure de la réputation et notamment au processus d'identification et de cartographie des acteurs influents. L'essor des médias sociaux a réinventé la manière dont les individus interagissent et communiquent. Les plateformes que proposent ces différents médias offrent désormais la possibilité, non plus essentiellement de communiquer ou d'échanger, mais aussi, et surtout de donner son opinion sur une entité, ses produits et/ou services, d'influencer *in fine* le comportement des autres individus.

L'internaute dans ces espaces numériques endosse désormais une certaine posture de "célébrité" sur la base de la confiance et de l'engagement des diverses entités de sa communauté, eu égard à l'ensemble de ces (inter) actions menées de manière régulière (Childers et al., 2018). Il est important de noter ici que cette posture d'influence ne se résume pas uniquement, au nombre de followers d'un internaute ou encore au nombre d'engagements obtenus, mais plutôt au nombre d'utilisateurs pertinents[142] qui verront et réagiront aux différentes publications[143] ou actions.

Dans la littérature, les études sur l'introduction des stratégies d'influence sont très récentes. À partir d'études sur Facebook (Aswani et al., 2017b ; Cavalli et al., 2011 ; Arora et al., 2019[144]), Twitter (Aswani et al., 2017a ; Bokunewics et Shulman, 2017 ; Arora et al., 2019), Instagram (Marijke et al., 2017 ; Muqaddam et Jin, 2018 ; Sokolova et Kefi, 2019 ; Arora et al., 2019) Youtube (Lund et Hodzic, 2018 ; Sokolova et Kefi, 2019) ; Github (Bana et Arora, 2018) et sur plusieurs autres plateformes tout aussi populaires comme Linkedin, Snapchat, les blogs (Goodman et al., 2014) ou encore les sites de e-commerce (Boss et al., 2018), les auteurs sus-cités ont tenté de proposer quelques pistes afin d'exploiter au mieux ce nouvel

---

[142] Et cette notion de pertinence est essentielle car étant au centre de la posture et de la typologie de son profil

[143] Extrait de la définition " *the ability to drive action. It is NOT about how many followers you have, nor it is how many engagements you get. The expected number of relevant users that will see and react to your post defines how influential you are"* proposée sur la plateforme Klear https://klear.com/what-is-influence ; https://klear.com

[144] Arora, A., Bansal, S., Kandpal, C., Aswani, R., & Dwivedi, Y. (2019). *Measuring social media influencer index- insights from facebook, Twitter and Instagram. Journal of Retailing and Consumer Services, 49, 86–101.* doi:10.1016/j.jretconser.2019.03.012

avantage qu'offre la détection et la cartographie des profils d'influence dans les environnements numériques.

Lagree et al., (2017) se sont intéressés à la plateforme de microblogging Twitter et, à partir de données réelles collectées via la plateforme, ont tenté de modéliser un mécanisme de diffusion visant à appréhender le concept d'influence. Mallipeddi et al. (2018) s'en sont servis afin de modéliser un mécanisme heuristique[145] visant à la fois à identifier les influenceurs sur Twitter, mais aussi et surtout à établir une stratégie de publications sur la plateforme. La mesure de l'influence d'un internaute suppose la prise en compte d'un constituant clé, le score ou l'index d'influence, ce que Byrne et al. (2017), Mittal et al., 2017 et plus récemment Arora et al. (2019) et Quan et al. (2020) ont tenté de modéliser.

Dans leur article[146], Arora et al., à partir de l'utilisation d'algorithmes d'apprentissage machine[147] (*machine learning*) sous-jacents, proposent un modèle de mesure de l'index d'influence d'un internaute adapté à trois principales plateformes (Facebook, Twitter et Instagram) et identifient les différents attributs (*features*) à prendre en compte dans la mesure de l'influence. L'étude de Quan et al. (2020), qui est encore plus complète utilise plutôt les approches de *deep learning*[148] et s'intéresse d'une part à l'identification des thématiques dans les publications via notamment les microtextes, et d'autre part à la détection des différentes typologies d'influenceurs via les catégories dans lesquelles ils peuvent être classifiés. Avant d'aborder la mesure de l'influence en tant que méthode sous-jacente de l'offre de mesure en e-réputation, il est intéressant de se questionner sur ce qu'est l'influence sur les plateformes socionumériques et comment sont identifiés les profils influents au sein de ces divers espaces numériques.

---

[145] Qu'ils ont nommé le "*polynomial time heuristic model*"

[146] Arora, A., Bansal, S., Kandpal, C., Aswani, R., & Dwivedi, Y. (2019). *Measuring social media influencer index- insights from facebook, Twitter and Instagram. Journal of Retailing and Consumer Services, 49, 86–101.* doi:10.1016/j.jretconser.2019.03.012

[147] Ils ont utilisé entre autres : les modèles de Ordinary least squares (OLS), le K-NN regression, le SVR (Support Vector Regression) et le Lasso regression

[148] Avec notamment l'utilisation de techniques de word embedding, le variational autoencoder et le Latent Dirichlet Allocation

Il est important de noter, nos travaux s'inscrivant dans les SIC, que l'origine du concept d'influenceur remonte à la théorie du *two-step-flow* (Katz, 1957) qui stipule que les messages qui sont diffusés à travers les médias de masse aux leaders d'opinion sont par la suite transmis par ces derniers au reste de la population la moins active. En effet, la théorie du two-step flow utilise la communication interpersonnelle (one to one) afin d'influencer les publics inactifs ou peu actifs (Lazarsfeld, Berelson & Gaudet, 1944). Adaptée aux médias sociaux, la théorie du *two-step flow* a mué vers une forme de *multi-step-flow*, partant du principe que sur le réseau internet, la circulation de l'information suppose diverses directions et interactions (Burt, 1999; Harik, 1971).

Les utilisateurs des plateformes socionumériques ont recours à une multitude de sources d'informations sur le réseau internet (incluant les leaders d'opinion). Par définition, les *social media influencers (SMI)* ou *social influencer* ou encore *key opinion leader (KOL)*[149] peuvent être perçus comme "*des utilisateurs des médias sociaux qui ont acquis une crédibilité dans un secteur spécifique grâce à leurs activités dans ce média et, qui, de par leur accès à un large public de consommateurs qui suivent régulièrement leurs activités sur les médias sociaux sont généralement perçus comme dignes de confiance par les autres utilisateurs"*[150] *(*Stubb et al., 2019).

En utilisant les plateformes socionumériques, ils partagent des informations au sein de leur réseau avec leurs *followers* (Freberg, Graham, McGaughey, & Freber, 2011) avec pour objectif d'influencer le comportement de leur audience. Perçus comme des leaders d'opinion sur les plateformes numériques, les *social influencers* sont caractérisés par leur crédibilité, leur forte activité sur les différentes plateformes, leur visibilité et leur audience (Ioanid, Militaru, & Mihai, 2015). D'entrée de jeu, on constate que la notion d'influence implique plusieurs concepts clés dont ceux de spécificité d'un secteur[151], d'audience et notamment de

---

[149] Terme employé par Quan et al. (2020) dans leur article sur la détection des influenceurs à partir de leur catégorisation

[150] Stubb, C., Nyström, A.-G., & Colliander, J. (2019). *Influencer marketing. Journal of Communication Management.*doi:10.1108/jcom-11-2018-0119 . Traduit de l'anglais " SMIs are users in social media who have established credibility in a specific industry through their activity in this media (Brown and Hayes, 2008). They have access to a large audience of consumers who follow their social media activity regularly and are generally perceived as trustworthy by other users (Keller and Berry, 2003)". p109 https://sci-hub.tw/https://doi.org/10.1108/JCOM-11-2018-0119

[151] Il est ici question des catégories dans lesquelles les influenceurs peuvent être classés et sur lesquelles nous reviendrons par la suite

la taille de celle-ci avec l'introduction notamment du concept de micro influence, mais aussi d'engagement autour des différentes publications, et surtout de crédibilité qui induit une relation de confiance avec leur audience (Swant, 2016).

Se positionnant comme " expert " dans un ou plusieurs domaines spécifiques, le KOL joue un rôle clé dans la diffusion des informations sur le réseau, selon sa position dans le graphe et sa typologie. Comment dès lors détecter et cartographier ces profils d'influence?

La mesure de l'influence d'un internaute suppose la prise en compte de plusieurs attributs relatifs à l'engagement, à la portée (au *reach*), à la polarité (*sentiment score*), au réseau (*network*), et nécessite l'utilisation de plusieurs techniques de *machine et deep learning* permettant, à partir du volume et de la récurrence de leurs interactions sur les différents médias sociaux, de les catégoriser. Le processus de catégorisation nécessite, dans certains cas (comme nous le verrons dans les modèles théoriques de la réputation) l'attribution d'un indice d'influence qui se mesure au prorata d'un ensemble d'indicateurs : taux de *reach*, d'engagement, de pertinence, de *relevance*, de polarité, de présence et de publication (Arora et al., 2019 ; Quan et al., 2020).

4.1. La détection de profils d'influence

L'une des problématiques à laquelle on se heurte dans l'identification de profils d'influence dans les espaces socionumériques provient de la nature même du concept d'influence. Nous identifions deux principales approches dans l'identification automatique des *social influencers* sur les plateformes numériques : l'approche par contenu et celle basée sur la structure (More et Lingam, 2017[152]). L'approche par contenu, qui utilise les techniques d'apprentissage machine, permet de détecter de manière automatique ou semi-automatique les KOL (key opinion leaders), et part du postulat que les médias sociaux pouvant être considérés comme des graphes complexes, la relation de co-auteur peut être utilisée dans la détection des profils, le contenu du message étant essentiel afin de proposer une catégorisation (Ritam et al., 2018).

[152] More, J.S. and Lingam, C. (2017), "A SI model for social media influencer maximization", Applied Computing and Informatics, Vol. 15 No. 2, pp. 102-108.

La seconde approche, celle structurelle qui est la plus utilisée dans le calcul du score d'influence et le classement des profils (More et Lingam, 2016 ; Arora et al., 2019) prend en compte les relations et liens entre les entités afin de les classifier. En outre, et pour identifier les différents sujets ou centres d'intérêt qui les caractérisent, il est possible soit d'utiliser la méthode manuelle (Bokunewics et Shulman, 2017) ce qui est souvent le cas notamment pour les personnes publiques ou célébrités ou via l'utilisation du crowdsourcing (Tanase et al., 2018).

Dans leurs travaux sur l'identification de KOL sur les médias sociaux, Sebastian et al. (2017[153]) se sont intéressés à la détection des véritables profils d'influence (induisant l'identification de faux profils ou fake ou *bots*) sur les forums de discussion où les utilisateurs publient autant des requêtes, des projets, des réponses que tout autre type de contribution. Afin d'y parvenir, ils ont utilisé deux méthodes d'analyse sémantique qu'ils ont combinées[154] afin d'identifier les similarités dans les documents et filtrer les interactions non pertinentes ou qui relèvent du bruit. Afin d'identifier le coeur du réseau qui, pour eux, est composé des utilisateurs les plus actifs qu'ils qualifient d'experts, ils ont appliqué l'algorithme ADA (*authoritative node detection)* dont l'objectif est de détecter les pages web faisant autorité dans l'agrégation de contenus en réponse à une requête d'un ou plusieurs utilisateurs.

À partir de la constance et la récurrence de leurs activités sur les forums et grâce à l'utilisation de l'algorithme ADA, Sebastian et al. (2017) ont identifié trois constituants clés dans la détection des profils d'influenceurs : *creator*, *last reply* et *all previous*. Le premier constituant part du constat que lorsqu'un utilisateur crée un fil de discussion, chaque réponse lui sera automatiquement liée. Par conséquent, les arcs[155] ne seront définis qu'entre l'utilisateur générant une réponse et l'utilisateur responsable de la publication d'origine dans le fil (cette représentation graphique a une faible densité dans la représentation graphique).

Cet état de fait se voit corrélé au second concept qui stipule que chaque réponse d'un fil de discussion sera considérée comme une réponse au dernier message uniquement.

---

[153] Sebastian, R.S.A., Aguilera, F., Nun~ez-Gonzalez, J.D. and Gran~a, M. (2017), "Semantically enhanced network analysis for influencer identification in online social networks", Neurocomputing, Vol. 326, pp. 71-81

[154] Soit l'analyse de Dirichlet latente ou *Latent Dirichlet Allocation* (Blei et al., 2003) et les méthodes de *Fuzzy concept analysis* (FCA)

[155] Qui permettent dans le graphe de relier les utilisateurs entre eux

Par conséquent, les arcs ne sont définis qu'entre l'utilisateur générant une réponse et l'utilisateur responsable du dernier message dans le thread (cette représentation graphique a une densité moyenne). Enfin, la troisième composante qui est la plus représentative stipule que chaque réponse d'un fil sera une réponse à tous les messages qui sont déjà dans un fil spécifique. Par conséquent, les arcs sont définis entre tous les utilisateurs générant une réponse et les utilisateurs responsables de tous les messages précédents dans le fil. Le nombre d'interactions entre l'initiateur de la publication et la dernière interaction permet donc d'identifier les membres influents qui peuvent être regroupés sous plusieurs typologies.

Un des éléments clés à ce stade et non des moindres, est la détection de faux profils disséminés sur les plateformes socionumériques par les usagers. Afin de garantir une certaine crédibilité des sources, ces faux profils présentent certaines caractéristiques et peuvent être détectés grâce à des mécanismes utilisant des méthodes de *machine learning*[156]. Afin de les identifier, Quan et al. (2020)[157] en proposent cinq principales caractéristiques qu'ils présentent comme suit : « *(1) interactions are not stable, (2) comments seem to be spammy and fake, (3) followers are fake or of low quality, (4) the posts contents are shared from unknown/suspicious links, (5) the text of the post is usually not properly formatted and includes many spelling/typo errors* »[158].

---

[156] Le modèle linéaire SVM de Bennett et Demiriz, 1999 en est un exemple

[157] Quan, T.-T., Mai, D.-T., & Tran, T.-D. (2020). *CID: Categorical Influencer Detection on microtext-based social media. Online Information Review, ahead-of-print(ahead-of-print).* doi:10.1108/oir-02-2019-0062

[158] Ibid. Traduction par nos soins des cinq principaux éléments qui permettent selon Quan et al., de détecter les faux profils d'influence : le manque de stabilité dans les interactions, la qualité des commentaires, le profil des followers, les partages provenant de liens inconnus ou suspects et le soin apporté à la rédaction des publications (coquilles, erreurs de typographie, etc.)

4.2. La classification des influenceurs selon leur typologie

Classifier les influenceurs suppose d'une part de les classer selon leur typologie et d'autre part d'identifier les catégories auxquelles ils appartiennent. À partir donc de la collecte de données sur leurs différents profils (les posts des individus sur leurs pages personnelles, les commentaires d'usagers en réponse aux posts sus-cités, les amis[159] , les followers, soit le nombre de personnes qui suivent un utilisateur[160]), il est possible de classer les influenceurs selon trois principales catégories : **les *hub*s, les autorités et les *spreaders***.

- Les ***hub***s se disent des utilisateurs dont le degré d'activité (il s'agit d'ici plus particulièrement de l'aptitude à produire du contenu à l'écrit) auprès des utilisateurs importants du réseau varie entre 0 et 1 (0 étant le seuil le plus faible et 1 celui le plus élevé). Les utilisateurs ayant un index de hub élevé sont donc ici appelés «rédacteurs», car ils produisent le plus de contenus envoyés aux utilisateurs importants du réseau ;

- L'**indice d'autorité** décrit le degré de citation d'un utilisateur par les autres utilisateurs ayant un degré de hub élevé (les meilleurs auteurs donc), et il peut également osciller dans la plupart des cas entre 0 et 1 (1 étant le niveau d'autorité le plus élevé), les valeurs proches de 1 relevant des profils propres aux autorités. De par leur statut donc, les autorités reçoivent un grand nombre d'informations et, pour cette raison, sont considérées comme importantes aux yeux des autres utilisateurs du réseau. Le plus souvent perçus comme les utilisateurs au plus grand nombre d'abonnés sur les plateformes social media, les autorités ont un rôle crucial et leurs avis font foi, et ce positivement ou négativement et influencent fortement l'opinion des autres utilisateurs ;

Il est important de noter ici que les deux premiers concepts *hub* et autorité sont intrinsèquement liés et les index de hub et d'autorité relèvent des concepts d'activité et de passivité d'un utilisateur au sein d'un réseau. En effet, les notions d'utilisateur actif et passif

[159] Il est question en particulier du nombre d'amis

[160] Se référer à l'article de Quan et al., 2020

sont calculées en comptant le nombre de messages envoyés et reçus (Wasserman et Faust, 1994) par ces derniers. Les deux index, *hub* et autorité, se renforcent mutuellement dans l'optique où un utilisateur avec un index de hub élevé, renvoie vers de nombreuses «autorités», et où une "autorité" est une page web contenant des liens provenant de nombreux « *hubs*".

- Par ***spreaders* ou encore acteurs centraux**, on entend les utilisateurs des media sociaux qui ont un fort potentiel de diffusion de l'information au sein du réseau ce qui leur confère un rôle stratégique auprès des organisations. L'indice de centralité permet donc de détecter les utilisateurs qui sont le plus souvent cités, qui ont de nombreuses connexions et, surtout, qui occupent une place centrale au sein du réseau, car étant les plus proches des organisations cibles et ayant un meilleur accès aux informations. La détection des diffuseurs au sein du réseau d'utilisateurs des médias sociaux pousse donc les organisations à adopter une stratégie de relations publiques «active» caractérisée par l'implication active des *spreaders* eux-mêmes dans la transmission des messages sur les médias sociaux.

4.3. L'attribution d'un score d'influence

La mesure de l'influence numérique peut prendre la forme d'indice ou score (notamment utilisés, nous le verrons dans les modèles théoriques de l'e-réputation). Cet indice suppose la prise en compte de plusieurs indicateurs : le *reach score*, le *resonance score*, le *relevance score* et le *sentiment score* (Boulier et Lohard, 2015 ; Arora et al., 2019 ; Quan et al., 2020) :

- Par *reach score*, on entend le nombre de personnes influencées par le message d'un internaute (sont prises en compte les réactions : like, réponse, partage, etc.);

- Le *resonance score* mesure la capacité de l'influenceur à générer ou non du trafic (interactions) autour d'un sujet défini. Pour y parvenir il est essentiel d'une part de

mesurer le *post-resonance score* et d'autre part l'*engagement score*. Tandis que le post *resonance score* au sens de Quan et al., fait référence à la méthode utilisée pour quantifier la pertinence d'un message et d'une thématique, l'*engagement score*[161] lui mesure l'impact d'une publication sur les habitudes de la communauté à partir du ratio entre la moyenne de l'*interactive score*[162] sur celle du nombre total de followers. Le *resonance score* est donc égal à la somme du *post resonance score* et de l'*engagement score*;

- Le *relevance score* indique la pertinence, la fiabilité d'un influenceur par rapport à une thématique spécifique. La prise en compte de ce score part du postulat qu'un influenceur sur une thématique particulière, ne publie pas uniquement du contenu sur celle-ci, mais à souvent recours à d'autres thématiques souvent affiliées ou pas dans ses posts. Il s'agit donc à partir de l'extraction des posts et commentaires, d'identifier les thématiques qui ressortent le plus dans les posts. Prenons l'exemple d'un influenceur dans le domaine de la mode, mais qui publie également des posts sur le lifestyle, les voyages, et le sport. Afin d'évaluer son *relevance score* dans la mode et à partir de l'analyse de ses différents posts, il s'agit d'attribuer pour chaque thématique sus-citée un indice selon leur importance et récurrence (au prorata du nombre de posts par thématique sur la totalité des posts). Si par exemple les posts par thématiques sont classés comme suit : mode (45%), lifestyle (25%), voyages (20%) et sport (10%), son relevance score dans la mode est de 0,45;

- Et enfin le *sentiment score* qui évalue l'opinion générale (positive, négative) des utilisateurs sur l'influenceur. À partir de méthodes de *deep learning*, le *sentiment score* d'un influenceur[163], se mesure au prorata des avis positifs moins ceux négatifs sur le total des avis (positifs et négatifs) multiplié par 100.

---

[161] La formule de l'engagement score se présente comme suit : *engagement_score = avg_interactive_score / total_followers*

[162] La moyenne de l'*interactive score* se calculant comme suit : *avg_interactive_score = (total_likes + total_shares + total_comments)/total_posts.*

[163] Utilise comme formule : *sentiment_score = (total_positive-total_negative) /(total_positive+total_negative) X 100*

Conclusion du chapitre 2

Dans ce chapitre, nous nous sommes intéressée aux éléments notionnels clés pris en compte dans la construction des modèles théoriques, et qui sont utilisés dans l'évaluation de l'e-réputation. S'il est un constat clé à retenir, c'est que l'e-réputation est un objet complexe qui s'articule autour d'une pluralité de notions connexes : la trace numérique, les identités numériques, la notion de confiance et celle d'influence.

En effet, l'e-réputation se construit à plusieurs niveaux. À partir des diverses expériences (au sens d'action ou activité) au sein des environnements numériques, les internautes produisent des données disséminées sur la toile, qui, prises individuellement, sont des séquences d'événements (Deransart, 2011) qui laissent de manière indéniable des traces. Ces dernières, dès lors qu'elles sont collectées et grâce notamment au phénomène de redocumentarisation (Pédauque et Salaün, 2007 ; Melot, 2006, Merzeau, 2009, Ertzscheid, 2013), sont indexées, mémorisées, codées, calculées, et contribuent à façonner une forme d'identité numérique continuellement documentée au fil des (inter)actions. Ce concept d'identité numérique a notamment été formalisé au sein de modèles tels que celui de Carter & al. (2002), Jurca & Faltings (2003), Gupta & Somani (2004), Tian & al. (2009) et Wang & al. (2013).

La troisième notion clé prise en compte dans la structure des modèles théoriques de l'e-réputation et à laquelle nous nous sommes intéressée, est celle d'influence numérique. La notion d'influence comme métrique de l'e-réputation apparaît dans plusieurs modèles notamment PAGERANK (Page & Brin, 1998), CUTEAID (Chen & al., 2006), PATROL (Tajeddine & al., 2007) et MORE (Osman & al., 2015). Le concept d'influence au sein des espaces numériques tel que conçu par la littérature scientifique remonte à la théorie du Two-step flow (Katz et Lazarsfeld, 1957) et utilise deux principales approches. Si l'approche par contenus prend en compte la relation de co-auteur dans les contenus diffusés, l'approche structurelle, elle, utilise les liens et notamment leur position au sein d'un graphe afin de catégoriser les profils d'influence. Trois principales catégories de profils ressortent de notre analyse : les *hub*s ou *rédacteurs* qui produisent le plus de contenus, les autorités (qui sont les plus cités et dont les avis font foi) et les *spreaders* en charge de la diffusion des informations.

Leur classification dans les modèles théoriques utilise un ensemble d'indicateurs (le *reach score*, le *resonance score*, le *relevance score* et le *sentiment score*) et peut, dans certaines approches, prendre la forme d'indice ou de score d'influence.

Le dernier concept associé à l'e-réputation dans la littérature scientifique est la confiance, avec pour objectif de réduire les incertitudes dans les interactions et de garantir l'intégrité, la fiabilité des données de réputation. Plusieurs modèles ont donc intégré dans leur structure des mécanismes d'évaluation de l'indice de fiabilité, indice qui peut se décliner selon plusieurs niveaux, allant de "*très fiable* à *pas digne de confiance*" (Abdul-Rahman et Hailes, 2005). La confiance est perçue comme un *output* de l'e-réputation, réputation qui sera, selon sa polarité (positive, négative, etc.), gage de confiance ou de méfiance pour l'entité évaluatrice et/ou évaluée.

Ce premier chapitre nous a permis d'identifier les métriques clés sur lesquelles s'appuie l'e-réputation dans la littérature scientifique. Dans le prochain chapitre, nous nous intéresserons aux approches méthodologiques mobilisées dans la mesure de l'e-réputation, et plus particulièrement à la fouille d'opinion, à l'analyse de sentiment et à la visualisation des données issues de l'analyse.

CHAPITRE 3 : DES MEDIAS SOCIAUX À LEUR ANALYSE

Introduction du chapitre 3

L'essor du web social dans les années 2000 semble avoir introduit de nouveaux dispositifs tant en termes d'interaction que de gestion des différents contenus mis à disposition. Le réseau internet et ses multiples canaux (il est ici question des différentes plateformes numériques) qui sont des supports pertinents dans l'établissement des relations entre les organisations et leurs divers publics (Waters et al., 2009) ont ouvert la voie à une certaine forme de personnalisation des contenus visant à faire correspondre les attentes des publics et celles des organisations.

Les médias sociaux peuvent être perçus dans une acception générale comme des plateformes à usage pluriel dont la fonction majeure est centrée sur le contenu et donc le principe de l'*UGC*[164]. Ils permettent entre autres de publier du contenu de tout type (articles, photographies, des vidéos, des fichiers PDF) ; de l'organiser à sa guise ; d'échanger ; de faire de la veille ou de la curation de contenu ; de co-créer, etc. (Proux et al., 2012; Milette, 2013; Coutant et Stenger, 2013; Dutot, 2013).

Parler de médias sociaux revient pour Alexandre Coutant (2013)[165] à les corréler à la notion d'innovation, innovation qu'il perçoit au sens d'Alter (2000). À partir d'une cartographie socio technique réalisée grâce à des entretiens auprès d'entrepreneurs , il distingue, avec Stenger (2013), quatre axes au sein de la nébuleuse des médias sociaux[166] : les réseaux socionumériques qui se retrouvent dans l'axe (soi - amitié), les sites de réseautage qui eux font partie de l'axe (soi- intérêt), les forums privés au sein de l'axe (amitié- contenu) et enfin au sein de l'axe (intérêt - contenu) le micro blogging, les forums publics, les réseaux de loisir créatif, les communautés virtuelles et les plateformes de signet. Dans la même optique, Fred Cavazza (qui nous propose sur son blog et ce depuis 2007, un panorama des médias sociaux) à partir de sa dernière cartographie (2020)[167] distingue six typologies de plateformes relevant

---

[164] User Generated Content ou encore contenu généré par les utilisateurs

[165] Coutant, Alexandre. « Quelle place pour l'innovation dans les médias sociaux ? » *Communication et organisation*, nº 43 (1 juin 2013): 123-34. https://doi.org/10.4000/communicationorganisation.4175.

[166] Cartographie sociotechnique de Coutant et Stenger (2013) http://communicationorganisation.revues.org/docannexe/image/4175/img-1.jpg

[167]Cartographie des medias sociaux selon Fred Cavazza. Accessible en ligne sur https://fredcavazza.net/2020/04/21/panorama-des-medias-sociaux-2020/

des médias sociaux : la publication (dans laquelle on retrouve les plateformes de blog, les wikis, les services hybrides de publication/partage, les services de gestion des podcasts) ; les réseaux de partage (vidéo, de documents, de photos, de données, de diffusion live, les applications de partage pour ados) ; les systèmes de messagerie mobile et traditionnelle ; les plateformes conversationnelles, les systèmes de gestion de commentaires et débats, les plateformes de questions-réponses, les espaces virtuels de discussion ; les messageries professionnelles et plateformes de collaboration et les réseaux sociaux professionnels, de niche, les sites de rencontre et d'organisation d'événements.



Figure n°6 : Panorama 2020 des médias sociaux, source Fred Cavazza

Ce qui de prime abord surprend c'est notamment le noyau formé par les réseaux sociaux (dont Facebook et les autres), constat qui questionne la position hiérarchique de l'ensemble de ces réseaux dans les questions de réputation notamment, mais aussi et surtout les critères de classification.

De manière générale, nous le verrons, la mesure de la réputation s'appuie essentiellement sur trois des catégories présentées : les réseaux de publication, les plateformes de discussion et les réseaux sociaux.

Il est important de souligner que nous avons choisi dans ce chapitre de ne pas proposer une description détaillée des médias et réseaux sociaux, mais plutôt d'axer notre approche sur la place qu'occupe désormais le consommateur (et dans une vision plus large l'internaute) dans la création de l'objet réputation et les approches mises en place afin de capter et analyser l'ensemble des métadonnées produites via notamment l'analyse des réseaux sociaux.

1. Le web 2.0 et les plateformes collaboratives : l'avènement du "*consom'acteur*"

Afin d'appréhender le rôle des consommateurs au sein des plateformes socionumériques et dans les organisations de manière générale, il nous semble intéressant de questionner le comportement de ces derniers basé sur les modalités d'engagement ou des types de relations sociales sous-jacentes (Cardon, 2008) dans une économie d'abondance. Avant qu'internet n'offre la possibilité de rassembler des individus de diverses zones géographiques et ne leur permette d'accéder à une quantité inestimable de données, l'accès à l'information passait par les médias de masse et utilisait les leaders d'opinion comme relais afin d'influencer le comportement des consommateurs (Burt, 1999 ; Watts & Dodds, 2007). L'intention d'achat a donc très longtemps été influencée par les amis ou les proches (Aquino, 2013), mais avec l'essor des médias sociaux, les consommateurs ont acquis un pouvoir d'influence et leurs recommandations font désormais foi dans le processus de décision auprès de leurs pairs.

Ce comportement est lié peu ou prou, et c'est là notre postulat, à deux phénomènes clés. Le premier est relatif à la structure même de ces espaces numériques qui ont aboli les barrières (tant physique que d'accès à l'information) et le second, au développement des théories de l'influence sociale appliquée au web.

La structure des réseaux socionumériques peut être décrite à l'aide de divers concepts clés qui sont : la force des liens, l'orientation de ceux-ci, la connexité, la groupabilité et les problèmes liés à la culture (Rissoan, 2014)[168] qui questionnent certaines problématiques :

- Sur les liens, Granovetter (1973, 1983) qu'il cite, part du postulat que la force des relations crée le réseau et distingue deux types de liens : les liens forts qu'on retrouve entre les membres d'une famille, entre amis ou proches collègues et les liens faibles entre de simples connaissances. Le calcul de la force des liens prend en compte cinq éléments : la fréquence des relations (qui sous-tend que les liens se renforcent plus on est en contact avec une personne), l'intimité (le caractère confidentiel ou intime des informations échangées renforce la relation), l'intensité émotionnelle (il s'agit ici du ton utilisé dans l'information qui est délivrée : humour, sympathie, tristesse, etc.), la notion de réciprocité des services (autant on pourrait avoir besoin d'un service, autant les autres pourraient solliciter notre aide) et la notion de multiplexité qui évoque la notion de clusters (pour Rissoan, appartenir à un mouvement ou un groupe avec lequel on partage les mêmes valeurs, idées et état d'esprit, c'est appartenir à un cluster qui est à distinguer d'un hub dans lequel on ne partage pas forcément de valeurs, et plus un individu appartient à différents clusters, plus sa multiplexité est importante et plus il servira de relais);

- Le deuxième concept inhérent aux réseaux sociaux est la circulation de l'information qui soit peut être unilatérale ou bilatérale. Il reprend la thèse de Vincent Lemieux (1982) pour qui les échanges entre deux acteurs peuvent aller dans un seul sens donc unilatéral (total ou partiel : qui consiste à transmettre une information sans en attendre de retour) ou dans les deux sens, et qualifié de relation bilatérale. Que l'orientation ou la circulation de l'information soit unilatérale ou bilatérale, il est important de prendre en compte le nombre de contacts ou relations qui est un facteur déterminant sur la portée de l'information qu'on véhicule;

[168] Rissoan Romain (2014), Réseaux sociaux : comprendre et maîtriser ces nouveaux outils de communication, 4e éd, p.42

- Par connexité, il faut entendre le nombre et la qualité des relations dont dispose un acteur au sein de son réseau. Plus un acteur aura de contacts, plus sa connexité sera forte et vice versa. Cependant, et comme le souligne Rissoan, il existe une énorme différence entre qualité et quantité des contacts, et il distingue à partir des travaux de Pascal Brassier (Boostez vos ventes 2010) plusieurs critères d'évaluation de la qualité de son réseau regroupés en trois catégories : ce qu'il qualifie de caractéristiques du capital social ou des contacts (la capacité à travailler ensemble, la multiplicité, le potentiel individuel de chaque contact : leader ou passeurs d'informations), le dynamisme du capital social (l'interaction, le partage, la confiance, le temps passé à échanger) et les trous structuraux (concept créé par Burt et qui désigne la faculté à être un relais d'informations ou blogueur influent);

- La groupabilité part de la maxime « les amis de mes amis sont mes amis » et sous-tend que les mêmes liens rassemblent les mêmes personnes. Il s'agit donc ici d'accroître son réseau en utilisant celui déjà existant, notamment le carnet d'adresse de nos contacts actuels en supposant de partager les mêmes centres d'intérêt, idées, etc. Il est donc ici question de recrutement et d'effet de rebond, de résonance de l'information qui amène parfois au buzz.

De par leurs activités dans les environnements numériques, nous l'avons vu, plusieurs catégories d'individus émergent dont les acteurs centraux (dont nous avons vu qu'ils peuvent soit être des *hub*, autorité, *spreaders*, etc.) qui ont acquis une certaine forme de crédibilité et dont le rôle est de servir de lien ou relais entre les autres membres de la communauté et dynamiser le réseau. Les concepts de liens à eux seuls suffiraient à justifier cette posture. En effet, dans la représentation graphique des réseaux sociaux, les utilisateurs nous l'avons vu, sont représentés sous forme de nœuds et les relations qu'ils entretiennent entre eux sous forme d'arcs. On peut donc se demander à partir de là, comment sont reliés les membres entre eux dans la construction du graphe ? Comme c'est le cas dans la vie courante, lorsqu'on se retrouve confronté à des individus "inconnus", l'intégration dans un ou plusieurs groupes s'impose et nécessite un ou des intérêts communs et c'est là qu'interviennent les principes de groupalité, culture et connexité inhérents aux plateformes socionumériques.

Le processus de décision des consommateurs étant influencé tant par des facteurs sociaux, personnels que psychologiques (Brosekhan, Velayutham, & Phil, 2014; Dahl, 2013) ces derniers se voient attirés vers des catégories d'individus qu'ils perçoivent soit comme similaires à eux, ou comme des experts dans leurs domaines respectifs, et en qui ils ont confiance. Trois phénomènes permettent d'expliquer l'influence sociale : l'internalisation, l'identification et le désir de conformité (Kelman, 1961). La stratégie d'internalisation part d'un manque d'informations qui pousse les individus à recourir à d'autres sources qualifiées d'experts en raison de leur fiabilité et cette relation se poursuit tant que la crédibilité de la source est avérée. Le consommateur se réfère donc à l'avis et aux comportements de l'expert dans son processus de décision. Le deuxième phénomène, l'identification, part d'un besoin d'approbation implicite. En effet, cela suppose qu'un individu admire une qualité particulière chez un autre individu, ce qui le poussera implicitement à s'identifier à lui, à l'imiter dans certains cas. On retrouve cette technique avec notamment les célébrités et leurs fans qui suivent tous leurs faits et gestes et achètent par exemple certains produits qu'ils ont ou dont ils font la promotion (Zipporah & Mberia, 2014). Le dernier critère, la conformité qui rejoint en certains points l'identification, découle d'un besoin d'appartenance et fait référence à la réaction d'un individu à une demande singulière afin de satisfaire les attentes d'un groupe particulier et d'obtenir des éloges ou d'éviter une punition (Cialdini et Goldstein, 2004; Kelman, 1961).

Ces trois phénomènes décrits dans la théorie de la comparaison sociale ou *social comparison theory* (SCT) au sens de Festinger (1954), supposent que « *les gens ont un désir naturel à se comparer aux autres afin de voir comment leurs opinions et leurs capacités sont perçues »,* désir de comparaison qui vise tant une auto-évaluation précise (Taylor, Wayment & Carrillo, 1996), une approbation personnelle (Wills, 1981) qu'une auto-amélioration (Taylor & Lobel, 1989). Ce phénomène d'influence sociale appliqué aux plateformes socionumériques a donc donné naissance aux "consom'acteurs" perçus aujourd'hui comme parties prenantes intégrantes des stratégies de marketing d'influence des marques, et à l'origine des données réputationnelles sur ces dernières. Quelles sont alors les différentes approches mobilisées dans la captation de l'ensemble de ces signaux? Sur quel(s) attribut(s) portent-elles?

2. L'émergence de la fouille d'opinion et du sentiment analysis dans l'analyse des réseaux sociaux (ARS)

Le marché de la veille en e-réputation ou fouille d'opinion ou encore *opinion mining* (Das and Chen, 2001 ; Morinaga et al., 2002 ; Pang, Lee and Vaithyanathan, 2002 ; Tong, 2001; Turney, 2002 ; Wiebe, 2000 ; Dave, Lawrence and Pennock, 2003 ; Liu et al., 2007 ; Liu, 2012) a connu un boom depuis les années 2000 avec l'essor des médias sociaux et occupe aujourd'hui une place centrale dans les stratégies de marketing digital. Processus de vigilance qui se veut permanent, la fouille d'opinion vise d'une part à surveiller l'ensemble des mentions sur une entité et son environnement, et d'autre part la détection en amont des signaux faibles (Boutin al al., 2008[169]) afin de prévenir les risques liés à l'image d'un individu ou d'une organisation, et de déployer les actions correctives nécessaires.

Applicable à tous les domaines d'activité[170] (mode, voyage, divertissement, finance, santé, électronique, etc.), la fouille d'opinion implique de prime abord un cadrage tant sur le type d'informations ou de phénomène recherché[171], la visée de la recherche (recherche action ou recherche à but explicatif d'un phénomène) que des ressources à mobiliser (prévoir en amont le binôme homme-machine, les données une fois collectées et analysées nécessitant une interprétation d'un expert dans le domaine). Boullier et Lohard (2012) distinguent à la suite de Camille Alloing (2011) et de l'agence Blueboat qui en a proposé une cartographie, une myriade d'acteurs évoluant dans le domaine de la veille spécialisée en e-réputation et allant de services internes aux entreprises (où elle est ici à l'état rudimentaire les services ne disposant pas des outils adéquats), aux laboratoires de recherches (à majorité rattachés à des universités et évoluant dans le domaine du TAL) en passant par les prestataires de service (répartis en 9 catégories : les agences spécialisées dans la protection de la réputation/

---

[169] Boutin, Eric, Pei Liu, et Lysiane Buisson. « Veille d'image sur Internet : enjeux, méthodes, limites ». *Communication et organisation*, nº 34 (1 décembre 2008): 98-114. https://doi.org/10.4000/communicationorganisation.611.

[170] Confère schéma du cabinet Beacon https://books.openedition.org/oep/docannexe/image/204/img-1.png

[171] Qui s'effectue en trois étapes : en suivant l'évolution des opinions sur une personne, un domaine, un produit/service, etc., en mettant en place des alertes ponctuelles sur des événements ou sujets et en explorant les réseaux où se créent et développent les opinions : différents canaux, liens entre eux, thématiques, etc.

juridique; les *web agencies* ; celles dans le conseil, la communication, l'influence et la stratégie ; celles dans les relations publiques / presse ; les prestataires outils/solutions logicielles ; les *pure players* ; les agences SEM/SEO ; les agences en veille, audit, analyse ; celles en marketing digital/ webmarketing[172]) et éditeurs de logiciels ou solutions. Les solutions ou suites logicielles que proposent ces différents prestataires de services en e-réputation peuvent être gratuites[173] et offrent une utilisation limitée d'un aspect particulier de la fouille d'opinion ou payantes.

Ces dernières sont pour la plupart réservées aux grandes entreprises avec une offre plus ou moins exhaustive. Nous nous intéresserons ici, et comme c'est souvent le cas dans la littérature, au deuxième type de solutions qui proposent une offre plus complète de la fouille de données. Le processus de la fouille de données suppose plusieurs phases : le paramétrage, la collecte automatique des données, l'archivage, leur traitement et leur analyse (Kotras et al., 2015) que nous décrirons par la suite.

2.1. Le paramétrage

La première étape qui jalonnera la suite du travail consiste au paramétrage général (*sourcing*, requêtage, etc.) de la solution. Le *sourcing* consiste à identifier les différentes sources à surveiller afin de collecter les informations nécessaires à l'analyse, choix des sources qui s'avère essentiel pour la suite de la stratégie. Les médias socionumériques, nous l'avons vu, regroupent une pléthore de sources de données disséminées sur la toile et au compte desquelles on retrouve : les sites institutionnels et commerciaux[174], les portails[175], les sites de presse généraliste ou spécialisée, les sites d'avis de consommateurs, les sites d'études

---

[172] Acessible en ligne sur https://books.openedition.org/oep/docannexe/image/204/img-2.jpg

[173] Au compte des solutions gratuites d'opinion mining : Samepoint (reputation management search engine) ;SocialMention (real-time social media search & analysis); BoardTracker (forum discussion search); Twilert (Twitter alerts); Comment (suivi de commentaires), etc.

[174] Qui même s'ils ne constituent pas la source première, permettent néanmoins via notamment les blogs, articles rédigés et commentaires de collecter des données

[175] De par la multitude de ressources vers lesquelles ils renvoient et la pléthore de secteurs qu'ils recouvrent

communautaires de partage de contenus[176], les bases de données[177], les blogs, les forums[178], les réseaux sociaux, le micro blogging, les sites de recommandations et de votes pour des contenus en ligne[179], les newsgroups[180] et les wikis[181] (Boullier et Lohard, 2012).

À partir de là, il est important de choisir les différents mots clés et requêtes à utiliser afin de collecter les opinions, mots clés choisis qui se doivent d'être des plus précis, concis et pertinents possible (il est d'ailleurs intéressant de noter que certaines plateformes offrent la possibilité comme le souligne Kotras (2015) de tester des requêtes[182] ou de générer des nuages de mots clés[183]). Une fois les mots clés choisis, il s'agit de les regrouper en catégories ou par thématiques, afin d'y intégrer l'ensemble des verbatims qui seront par la suite recueillis, et de mettre en place des alertes en prenant soin d'utiliser des mots clés pertinents afin "d'éliminer les bruits"[184] et silences[185]. Comme nous le voyons, à cette étape, un travail de minutie est nécessaire et une fois le cadrage du *sourcing* établi, il est important d'effectuer un travail de qualification et de catégorisation des sources.

Qualifier des sources revient à cette étape du *sourcing*, et de manière générale, à mesurer l'audience des différentes sources et leur attribuer un score d'influence[186]. Il est important de noter ici que plusieurs sources peuvent être utilisées de concert dans la collecte des données afin d'obtenir plus d'exhaustivité et représentativité et certains auteurs l'ont expérimenté dans

---

176 Notamment les sites de partage de vidéos tels Youtube qui renferment des commentaires associés aux vidéos et qui peuvent être analysés via des outils spécifiques de "speech to text"

177 Ces sources étant pour la quasi totalité fermées, elles sont utiles pour Boullier et Lohard pour de la veille et non pour de l'analyse de la tonalité des opinions

178 Qui est une des sources les plus utilisées et notamment par un des prestataires de service dans le domaine Synthesio et notamment sur l'exemple de leur étude pour Orange

179 Exemple de Reddit, Digg, StumbleUpon qui proposent des classements des sites les plus recommandés par les internautes

180 Qui sont des groupes de discussion thématique

181 Systèmes de publication collaborative permettant à tout utilisateur de consulter ou modifier le contenu

182 tant sur la volumétrie que l'échantillon de verbatim potentiel à recueillir, exemple de Meltwater

183 Exemple de la solution Synthesio

184 Ce qu'ils qualifient de "désambiguïsation"

185 Ici perçus comme les "synonymes, fautes d'orthographe courantes"

186 Se référer à l' influence et e réputation au chapitre 2

leurs travaux. C'est notamment le cas de Boutin et al. (2008) qui ont utilisé dans leurs travaux sur la fouille d'opinion appliquée au secteur bancaire et plus précisément à la caisse d'épargne les sites de recommandations (et plus précisément la plateforme Ciao[187]) et les blogs (avec notamment le moteur de recherche de blogs technorati.com[188]), ou encore de Kotras (2015) qui a plutôt utilisé les réseaux sociaux des organisations (à partir de la solution *Buzz* de Meltwater) qui ont au préalable été connectés aux sites institutionnels afin d'identifier les communautés, etc. Une fois le cadrage de la recherche posé, commence alors l'étape de collecte, suivie par le traitement des données.

2.2. La collecte et le traitement des données : préparation du corpus et archivage

La deuxième phase, la collecte des données[189] utilise un "*crawler*" et s'effectue de manière automatique une fois les mots clés définis et le paramétrage effectué. Aussi appelé robot d'indexation, le *crawler* est un logiciel paramétrable (fréquence de passage sur les sites, profondeur des liens à explorer, etc.) qui permet d'explorer ou visiter des pages avec pour objectif soit de découvrir de nouvelles pages et favoriser leur indexation, soit faciliter la maintenance d'un site web par l'analyse des liens hypertextes ou encore de parcourir les pages web afin d'identifier et collecter des informations précises. On distingue deux types de *crawler* : ceux qui parcourent tout le web de manière régulière et indexent à partir de mots clés et via les liens hypertextes qui relient entre elles les différentes pages (exemple de GoogleBot, MSNBot, Exalead) et ceux modulaires qui proposent un « *custom web crawling* » et permettent de choisir les sites à crawler.

---

[187] Analyse des signaux émis et reçus. Application d'algorithmes d'analyse textuelle, polarité d'un texte à l'aide d'un dico de polarité, analyse quanti du signal émis, identification des différents profils clients, collaborateurs, concurrents

[188] Qui aujourd'hui est devenu https://www.synacor.com et répertorie plus de 15 millions de blogs

[189] il est intéressant ici de distinguer ici les données au sens de matériaux linguistiques qui peuvent prendre huit principales formes : un jugement, une évaluation, une opinion, un avis, un sentiment, un goût, un récit d'expérience et un récit de pratique (Boullier et Lohard, 2012)

Le paramétrage est ici important eu égard à la masse d'informations. Il s'agit donc en amont d'identifier les bouquets de sources à *crawler*[190], le type de sources à crawler[191], la granularité et la régularité du crawl[192] et d'identifier à la fois les mots clés et groupes de mots clés à surveiller. Une fois que les données des sites issus du crawl[193] du web et qui traitent du sujet ou de la marque dont on effectue une analyse ont été récupérées[194], plusieurs actions s'avèrent nécessaires : le nettoyage des pages, la classification des contenus dans des BDD, l'indexation des contenus via des métadonnées, la définition des unités servant de base au travail et l'attribution d'un indicateur d'influence ou de visibilité du contenu.

Le nettoyage des pages *crawlées* vise à filtrer les données "inutiles" et susceptibles de "biaiser l'analyse"[195], un paramètre inclus chez certains prestataires[196]. Boullier et Lohard sur la base du rapport Cleaneval[197] identifient les différentes problématiques qui y sont associées, soit l'identification des données sans intérêt linguistique (en-têtes, pieds de pages, etc.), l'identification des paragraphes et la standardisation des contenus pour une analyse linguistique, etc.

---

[190] Les auteurs ont ici pris l'exemple pour un client de surveiller les sites que ce dernier a l'habitude de visiter ou consulter

[191] Des réseaux sociaux aux blogs en passant par la presse, les portails, etc.

[192] Collecte en temps réel, de manière journalière, etc.

[193] Pour les réseaux sociaux et sites de microblogging, on utilise les API de ces derniers ou des moteurs de recherche dédiés afin d'extraire le contenu qui sera analysé

[194] Ici se sont plutôt des urls qui renvoient vers pages web

[195] C'est notamment le cas des publicités, liens vers d'autres pages, liens sponsorisés, etc.

[196] Exemple du crawler utilisé par l'agence Synthesio

[197] Rapport Cleaneval accessible en ligne https://cleaneval.sigwac.org.uk/lrec08-cleaneval.pdf

Dans le rapport Cleaneval, le nettoyage des pages prend en compte la préparation des données[198] et le *scoring*[199]. Une fois les pages nettoyées, elles sont archivées comme le notent Boullier et Lohard dans des bases de données puis catégorisées (sous forme d'étiquettes de manière automatique ou semi-automatique) afin de pouvoir les retrouver. Catégoriser et archiver les contenus nécessite de définir une unité d'analyse qui peut être soit d'archiver toute la page nettoyée ou de choisir le verbatim seul[200] et un formatage commun.

La classification des sources dans une base de données s'effectue selon les objectifs prédéfinis et prend en compte plusieurs critères : par site, thématique, produit, type de sources, etc., et le corpus peut être découpé en commentaire, tweet, verbatim, etc., puis indexé[201] avant d'être archivé dans des bases de données sous forme de métadonnées[202]. La réussite de la stratégie de veille réside, comme le souligne Kotras, dans le caractère flexible et modulable des solutions utilisées permettant de corréler et/ou dissocier des variables et ne les cloisonnant pas à un usage restrictif, préétabli. La phase ultime du processus consistera à analyser l'ensemble du corpus, afin de détecter des tendances ou saillances, à partir desquelles seront déclinées les différentes actions relatives à la gestion de la réputation dans son entièreté.

---

198 Qui comprend la sélection des données (ici à partir de requêtes dans Google et avec des mots clés prédéfinis, ils ont sélectionné les pages web uniquement en html en identifiant des langues ici l'anglais et le chinois, l'exclusion des pages trop longues, trop courtes ou dont le corpus n'est pas centré sur le texte et avec l'utilisation des méthodes de markup language), les directives d'annotation (supprimer tout le code HTML / JavaScript et autres : en-têtes, mentions de copyright, listes de liens, mentions répétées sur la plupart des pages d'un site, etc.; et ajouter un codage de base de la structure de la page en utilisant un ensemble minimal de symboles pour marquer le début des en-têtes, des paragraphes et des éléments de liste https://cleaneval.sigwac.org.uk/annotation_guidelines.html) et la Configuration de l'annotation et le format du document de sortie (avec l'utilisation des méthodes de traduction assistée par ordinateur tels Trados et Wordfast afin de convertir les pages : UTF-8 et enca en supprimant les balisages html, javascript et autres et ajouter manuellement des tags headers (l), paragraphs (p) et lists (l) exemple figure 1 page 3.

199 Qui consiste en l'attribution de métriques de scoring (ici l'écart entre deux fichiers nettoyés et la granularité entre deux différents types de fichiers nettoyés) et à former le scorer de manière à ce qu'il puisse prédire le comportement humain

200 En s'assurant toutefois de préserver son intégrité via la sauvegarde du contexte dans lequel il a été énoncé et du lien vers lequel il renvoie

201 Qui s'effectue de manière générale sous forme de tags et ne reposant sur aucun cadre de référence ni dictionnaire mais sur des critères subjectifs d'où le néologisme de "folksonomie" utilisé pour ce système

202 Source, date, auteur, etc., d'autres métadonnées pouvant être ajoutées à celles existantes

2.3. La phase d'analyse : le dilemme de l'analyse de sentiment

L'analyse du set de données collectées est une étape cruciale et peut prendre, selon les objectifs et besoins, plusieurs formes : l'analyse de volumétrie[203], les mots fréquents, l'analyse de la tonalité, le suivi dans le temps, la distribution des résultats[204], l'évolution comparée des résultats, les entités nommées et expressions significatives associées, les co-occurrences de termes[205], l'identification de corrélations, la détection de communautés[206], etc. (Boulier et Lohard, 2012, p.70). La phase d'analyse la plus "complexe" est l'analyse de sentiments ou *sentiment analysis* à laquelle nous nous intéresserons plus particulièrement.

Elle s'intéresse à l'analyse des opinions, sentiments, évaluations, appréciations, attitudes et émotions des individus envers diverses entités allant des produits, services, des organisations, des individus, aux problèmes, événements, sujets, et leurs attributs (Bing Lui, 2012) afin de les classer selon qu'ils sont positifs ou négatifs (Pang, Lee et Vaithyanathan, 2002; Turney, 2002), polarité qui s'effectue à l'aide d'une liste de mots et phrases répertoriés dans un *sentiment lexicon* ou *opinion lexicon*. Pour se faire, trois principaux éléments sont pris en compte : le document (on postule que chaque document exprime une opinion sur un objet singulier que ce soit un produit, service ou entité et qu'il contient de fait une opinion à classer), la phrase (il est question d'attribuer pour chacune une polarité selon qu'elle se veut positive, négative ou neutre en distinguant les phrases objectives souvent relatives à des informations factuelles de celles subjectives qui expriment des opinions et points de vue subjectifs) et l'entité et tous ses aspects (qui est une méthode holistique, mais plus complexe que les deux premières se basant sur le fait qu'afin d'attribuer une polarité à une opinion, il est essentiel d'identifier tous les aspects sur lesquels se fonde l'avis celui-ci pouvant être

---

[203] Relatif à l'aspect quantitatif

[204] Classification par source, langue, sujet, etc.

[205] "clusterisation"

[206] Qui vise d'une part à identifier les différents membres de la communauté (fans, followers, etc.) et à détecter les différents profils qui s'expriment sur l'organisation (client satisfait, client mécontent, concurrent, intervenant ponctuel, régulier, hostile, favorable, acteur qui commente, commentaire confirmatoire ou infirmatoire, etc.,) et d'autre part les profils influents (qu'ils aient une influence positive ou négative sur la marque) afin de mettre en place pour chacun des profils la stratégie adéquate (Boutin et al., 2008)

positif sur un des attributs et négatif sur l'autre[207]). Ces deux éléments se résument dans les deux typologies d'opinions proposées par Jindal and Liu (2006b) soit les *regular opinions* (exprimant tout simplement un avis sur une entité ou ses attributs) et les *comparative opinions* (qui tendent à comparer soit plusieurs entités soit plusieurs attributs de la même entité). Détecter la polarité ou tonalité dans des documents revient à : identifier si le document comporte ou non une opinion, à en définir le type d'opinion soit générale ou relative à un attribut particulier, à déterminer sa polarité (positive, négative ou neutre) et à associer pour chaque opinion, une cible.

Plusieurs approches sont prises en compte dans l'analyse de sentiment : la sélection et classification d'un matériau subjectif, les approches lexicales et les approches par apprentissage (Boullier et Lohard, 2012). Dans la première approche, l'analyse de sentiment s'effectue à l'aide de textes subjectifs qu'on retrouve pour la plupart sur les sites d'avis de consommateurs sous forme de commentaire ou critiques, ou *a contrario*, à l'aide de signes distinctifs relevant de la subjectivité et qualifiés comme tel par certains chercheurs. C'est notamment le cas des émoticônes[208]. Les deuxièmes approches, celles lexicales qui sont les plus utilisées sur le marché, utilisent des dictionnaires de mots subjectifs[209].
Dans ces dictionnaires, les termes sont répartis en catégories (multiples et variées selon les dictionnaires : positif, négatif, hostile, autoritaire, passif, etc.) afin de classer les textes, et ont cependant très largement été critiqués[210]. Les dernières approches, celles par apprentissage, utilisent les machines dans la détection d'expressions subjectives, de modèles ou de motifs, en effectuant au préalable un pré test sur un corpus.

---

[207] L'exemple le plus courant est celui d'un téléphone portable qu'on peut aimer pour la qualité des photos réalisées mais qui à une autonomie de la batterie courte, ou encore un restaurant qu'on peut apprécier pour la qualité des produits contrebalancé par un service qui laisse à désirer

[208] Les auteurs prennent l'exemple des tweet contenant des émoticônes qui peuvent être collectés via l'API de twitter

[209] On peut citer comme exemple : le General Inquirer, Sentiwordnet, Opinion Finder, NTU Sentiment Dictionary (NTUSD), WordNet Affect, etc.

[210] Ils ne prennent pas en compte le contexte de l'énoncé, les expressions ambiguës sont parfois supprimées, ils ne prennent pas en compte la négation et les figures de rhétorique (sarcasme, ironie, etc.)

Cette approche utilise, pour Boullier et Lohard à la suite de Pang et Lee, plusieurs caractéristiques (*features*)[211], et est construite autour d'une pléthore de méthodes.[212]

Sur le *sentiment analysis*, très critiqué et controversé à travers la littérature, Kotras propose d'une part non pas une analyse des scores bruts au temps "t", mais plutôt une analyse de leurs variations, car comme le souligne un des interrogés dans le cadre de son analyse « *tout le monde sait que ce n'est jamais à 100 % exact. Mais si, pour une raison ou pour une autre, on a un pic qui passe des 5 % habituels à 40 % de négatifs, ça veut dire quelque chose. Ça vaut au moins le coup de vérifier* »[213] et d'autre part, dans le cas où l'organisation serait soumise à une récurrence de résultats inadaptés de désactiver l'analyse de sentiment afin d'éviter ou réduire les biais. Chaque terme étant associé à une polarité, il prend l'exemple d'une entreprise qui parce qu'elle traite de sujets anxiogènes aura une réputation négative qui cependant est le reflet de son cœur d'activité et non de son image. Certains logiciels proposent pour pallier cela de modifier manuellement le codage des verbatims[214], ce qui nécessite une bonne maîtrise de ces solutions et algorithmes et d'autres, moins restrictifs, proposent une interface plus ergonomique et quelque peu plus modulable[215]. Les solutions de veille en e-réputation comme nous le voyons sont construites sur une structure "modulaire" visant à favoriser la perméabilité du processus. Leur utilisation nécessite cependant une certaine expertise tant dans le paramétrage en amont, que dans l'interprétation des résultats issus de l'analyse et présentés sous forme de visualisation dans les tableaux de bord de la réputation.

---

[211] Au sein desquels on peut retrouver ces *features* : Term presence *vs* Frequency, Term-based Features Beyond Term Unigrams, Part of Speech, <#GRAS#>Négation, Topic-oriented features</#GRAS#> et <*#ITALIQUES#*></#ITALIQUES#>Syntaxe

[212] Dont notamment Naive Bayes, Maximum Entropy, Support Vector Machine, la méthode Isotonic Conditional random Fields, etc., qui ont notamment été utilisés sur un corpus de critiques de cinéma

[213] Propos recueillis par lui auprès de Travis, *sales manager* pour Sysomos

[214] Exemple de Synthesio et Radarly

[215] C'est notamment le cas de Sysomos

3. La visualisation des données : le boom des cartographies dans l'analyse des réseaux sociaux

Tableau[216], l'un des acteurs majeurs[217] de la visualisation de données la définit comme *« la représentation graphique d'informations et de données. À l'aide d'éléments visuels comme les graphiques et les cartes, une visualisation de données permet de voir et de comprendre des tendances ou des valeurs inhabituelles dans les données, de manière très accessible »*[218]. Appliquée à la quasi-totalité des secteurs d'activité (marketing, finance, sport, santé, transport, etc.), la *datavisualisation* peut prendre plusieurs formes allant des plus courantes (diagrammes, cartes, graphiques, tableaux de bord, infographies, etc.) aux visualisations quelles que peu spécifiques (heatmap, diagramme de Gantt, matrices, nuages de mots, boîtes à moustaches, chronologie, etc.).

S'intéresser à la cartographie des réseaux numériques revient à prendre en compte plusieurs terminologies clés qu'elle questionne soit les notions de topologies et topographies, de « communauté » parfois remplacée par « territoire » (notamment pour Cardon et al., 2014[219]), de frontière, de noyau central, de liens hypertextes, de cluster, de périphérie, d'échelle, etc.
Sur la question de communauté et notamment son essor dans les espaces du web, elle nous renvoie à Howard Rheingold (1993) avec son approche de *communauté virtuelle*. Le centre National de ressources textuelles et lexicales nous propose plusieurs définitions du concept de communautés dont notamment « *l'état ou le caractère de ce qui est commun à plusieurs personnes*», « *un ensemble de personnes* » , « *groupe de personnes ayant un lien en*

---

[216]Un des acteurs majeurs du marché de la visualisation des données numériques. Accessible sur https://www.tableau.com/fr-fr

[217] Il en existe plusieurs des acteurs selon les besoins en termes en cartographie et au compte desquels on retrouve entre autres Qlikview, Power BI, Tableau, Google data studio, Raw, Sigma JS, D3.js, Datawrapper, Toucan toco, etc.

[218] Définition du concept de visualisation de données proposé par Tableau. Accessible en ligne https://www.tableau.com/fr-fr/learn/articles/data-visualization

[219] Cardon Dominique, Fouetillou Guilhem, Roth Camille, « Topographie de la renommée en ligne. Un modèle structurel des communautés thématiques du web français et allemand », *Réseaux*, 2014/6 (n° 188), p. 85-120. DOI : 10.3917/res.188.0085. URL : https://www.cairn-int.info/revue-reseaux-2014-6-page-85.htm

commun »[220], définitions rapprochées en termes de synonymie à plus d'une trentaine de concepts dont ceux d'association, accord, société, groupe, harmonie, réunion, similitude, unité, conformité, affinité, ressemblance, etc.,[221] et qu'il distingue (les antonymes[222] donc) des questions de différence, disparité, isolement et opposition.

On peut postuler à partir de là que la question de la communauté suppose un collectif réuni autour d'un ou plusieurs attributs ou thématiques communs, collectif au sens d'individus, mais aussi et surtout perçu comme agrégats de sites ou de blogs sur le web (Cardon, 2014) relié soit par une thématique[223] ou par des liens hypertextuels[224]. En effet, tandis que l'approche sémantique regroupe les collectifs selon leurs principales thématiques, celle topologique, qui renvoie à la question des hyperliens, consiste à prendre en compte la connectivité et proximité avec notamment au centre la force des liens hypertextuels (liens sortants et entrants) à l'origine de leur position géographique (avec la position de centralité au coeur des préoccupations).

Les liens entrants, qui par ailleurs jouent un important rôle dans la réputation au sein d'un réseau, permettent d'identifier quatre profils : la *célébrité* (ce qui suppose pour ce dernier que plusieurs liens entrants provenant à la fois de son réseau et des autres communautés pointent vers lui), la *notoriété* (ce qui sous-entend que l'internaute qui figure dans cette catégorie suscite plus d'intérêt dans son propre réseau, mais pas à l'extérieur de celui-ci), la *popularité* (l'internaute qui figure dans cette catégorie suscite plus d'intérêt provenant de liens extérieurs à son réseau qu'en interne) ou encore l'*invisibilité* (pour les profils qui ne reçoivent aucun lien vers la source).

---

[220] Définition du concept de communauté proposé par le CNRTL. Accessible en ligne https://www.cnrtl.fr/definition/communauté

[221] Ensemble de notions associées au concept de communauté. Accessible en ligne https://www.cnrtl.fr/synonymie/communauté

[222] Ensemble de notions à distinguer du concept de communauté. Accessible en ligne https://www.cnrtl.fr/antonymie/communauté

[223] Cardon parle ici de communauté sémantique

[224] Et fait ici cas de communautés topologique

Les liens sortants eux, dénotent le degré d'engagement au sein d'un réseau, et permettent également de distinguer quatre catégories de profils : *curieux* (lorsque tous les liens sortants convergent tant vers son réseau qu'à l'extérieur de celui-ci, *extraverti* (avec des liens sortants qui renvoient uniquement vers l'extérieur de sa communauté) *introverti* (avec des liens sortants centralisés au sein de sa communauté) ou *silencieux* (lorsqu'il ne renvoie aucun lien sortant vers d'autres sites de quelque nature que ce soit, Cardon, 2014). L'identification au sein des communautés des diverses typologies de profils à travers notamment les liens entrants et sortants qu'ils produisent permet donc de distinguer le noyau central d'un réseau le plus souvent perçu comme les acteurs jouant le rôle d'autorité (Severo et Venturini, 2016[225]) se citant réciproquement et formant une nébuleuse autour de laquelle gravitent diverses sources périphériques. La mise en relation de cet ensemble d'acteurs permet la représentation du web sous forme de *clusters*.

La représentation de cette myriade d'acteurs qui soulève plusieurs problématiques (liées notamment à la position des acteurs au sein du graphe, mais aussi et surtout questionne la limite et donc la notion de frontière les séparant : qui se mesure souvent à l'aide de la densité de la connexion entre la myriade de graphes) utilise pour ce faire autant la topologie au sens de Connor (2004) comme relevant de l'étude et la représentation « *des relations spatiales, telles que continuité, voisinage, intériorité et extériorité, disjonctions et connexion*s » que la topographie, qui prend en compte les relations spatiales en rapport étroit avec l'espace géographique.

Plusieurs étapes sont nécessaires dans la cartographie des acteurs du web allant de l'identification des acteurs principaux, en passant par le crawl des données via des crawler, afin de tracer des liens entre les différents sites et en faire apparaître d'autres, la classification des sites pour enfin visualiser les réseaux grâce à un outil de manipulation des graphes (Severo et Venturini, 2016). Sur l'identification des principaux acteurs, il s'agit d'utiliser les méthodes de fouille d'opinion[226] puis à partir de robots d'indexation et à l'aide de mots clés

---

[225] Severo Marta, Venturini Tommaso, « Enjeux topologiques et topographiques de la cartographie du Web. Le cas du patrimoine culturel immatériel français », *Réseaux*, 2016/1 (n° 195), p. 85-105. DOI : 10.3917/res.195.0085. URL : https://www.cairn-int.info/revue-reseaux-2016-1-page-85.htm

[226] Se référer au paragraphe précédent sur l'opinion mining et le sentiment analysis

prédéfinis, d'identifier les sources d'information pertinentes afin de détecter les acteurs influents dans le domaine au sein duquel on souhaite établir une cartographie et détecter les liens et les autres communautés connexes. L'étape suivante consiste à classer, par catégorie, les différents acteurs qui émergeront, afin, par la suite, de visualiser les réseaux identifiés grâce à des outils de spatialisation. Le travail de classification des acteurs identifiés selon les différentes catégories ou thématiques auxquelles ils appartiennent et leur position au sein du cluster est essentiel afin de faciliter la visualisation des graphes et la lecture de la carte.

En effet, la composition, position et taille de chaque noeud au sein du graphe est à prendre en compte dans l'optique où « *la position de chaque noeud acquiert du sens par rapport aux positions relatives des autres* »[227], et la perception de nœuds du même type centrée dans une certaine partie du graphique permet de les rassembler et de déceler le centre du réseau et les zones périphériques. Les algorithmes de spatialisation qui fonctionnent donc selon le couple attraction - répulsion, garantissent une proximité des noeuds appartenant à la même communauté et une séparation avec les autres noeuds ne relevant pas du même ensemble, avec la représentation des acteurs sous forme de noeuds et les liens hypertextuels sous forme d'arcs.

Conclusion du chapitre 3

La réputation, nous l'avons vu, se construit sur les plateformes du web social, à partir des in(actions) des internautes. Ce chapitre visait d'une part à présenter les divers media sociaux à l'origine de la création de l'ensemble des traces numériques, mais aussi et surtout le processus théorique de mesure de la réputation. Si la réputation s'appuie sur les données provenant des réseaux de publication, des plateformes de discussion, de partage et des réseaux sociaux[228], le processus de mesure s'effectue en quatre (4) principales étapes : la phase de paramétrage, la collecte et le traitement des données, la phase d'analyse et enfin la visualisation des données.

---

[227] Severo Marta, Venturini Tommaso, « Enjeux topologiques et topographiques de la cartographie du Web. Le cas du patrimoine culturel immatériel français », *Réseaux*, 2016/1 (n° 195), p. 85-105. DOI : 10.3917/res.195.0085. URL : https://www.cairn-int.info/revue-reseaux-2016-1-page-85.htm , p.16

[228] Distinction proposée par Fred Cavazza sur son blog

Conclusion de la partie I

La première partie de nos travaux avait pour principal objectif de proposer un état de l'art des concepts clés mobilisés dans la littérature scientifique sur l'objet e-réputation.
L'une des notions clés à laquelle nous nous sommes intéressée est notamment celle de réputation au sens traditionnel, avec pour objectif, non pas de présenter une énième revue des courants qui l'ont étudiée, mais plutôt, d'identifier l'objet sur lequel porte l'évaluation à travers les divers champs. Selon la discipline, l'évaluation de la réputation porte sur les actions passées, les attributs identitaires, et l'offre (produits et/ ou services) pour les économistes et les gestionnaires, sur les interactions sociales et le contexte dans lequel celles-ci interviennent (pour la sociologie), les états mentaux pour les psychologues, la trace (perçue comme signal) pour la philosophie, et enfin, les opinions et interactions pour les SIC. Qu'elle soit perçue comme résultante d'interactions sociales, comme construit d'opinion ou d'information, comme signal ou encore comme actif intangible d'une entité, la construction de la réputation et donc son évaluation suppose quatre principales composantes : un contexte d'évaluation, deux ou plusieurs agents (le couple émetteurs - destinataires), un objet sur lequel porte la mesure (personne physique, organisation ou offre) et un historique (l'ensemble des données disponibles sur l'objet de mesure).

Nous nous sommes par la suite intéressée à l'objet central de nos travaux, l'e-réputation, en questionnant les éventuelles mutations, notamment des concepts mobilisés dans sa construction, dans le passage de la réputation traditionnelle à l'e-réputation. Nous nous sommes alors appuyée sur quatre principales notions : la trace numérique, les identités numériques, l'influence numérique et la confiance. L'e-réputation d'une entité se construit, il est important de le souligner, par l'agrégation de l'ensemble des traces disséminées au sein des environnements numériques par ses divers publics. Ces traces, qui seront par la suite indexées, enregistrées, calculées, contribuent à façonner une forme d'identité numérique à l'origine de l'e-réputation.

À l'instar de ces deux concepts, l'e-réputation comme objet pluridimensionnel fait également appel, dans ses définitions, aux notions d'influence et de confiance. Sur l'influence numérique, il est intéressant de noter qu'elle apparaît dans la littérature sur les modèles

théoriques de l'e-réputation avec l'un des modèles de référence sur la question, le PAGERANK (Page & Brin., 1998). PAGERANK propose un indice de popularité des pages web (et est notamment l'un des algorithmes utilisés par le moteur de recherche Google), et fonctionne selon le principe des graphes et des liens (entrants et sortants). Trois principales catégories d'acteurs ressortent donc de notre analyse : les *hub*s, les autorités et les *spreaders*.
Le dernier concept sur lequel nous nous sommes appuyée dans cette partie est celui de confiance. S'il est un constat qui peut, d'emblée être fait, c'est que les deux notions (l'e-réputation et la confiance) sont intrinsèquement liées dans la littérature scientifique, avec la confiance perçue comme un *output* de la réputation. Ce constat a donc poussé plusieurs auteurs à intégrer des mécanismes d'évaluation de la confiance numérique au sein de leurs modèles de réputation. C'est notamment le cas des modèles hybrides de confiance et de réputation (REGRET, 2002 ; AFRAS, 2002 ; XREP, 2002 ; EIGENTRUST, 2003 ; TRUMMAR, 2004 ; FIRE, 2006, PATROL, 2007, etc.), qui proposent, à partir des données de réputation, d'attribuer un indice de confiance aux entités.

Ce premier cadrage des éléments notionnels et de contexte de notre objet d'étude posé, le troisième chapitre de cette partie consistait à identifier et définir les techniques et méthodes d'analyse des données de réputation, et plus particulièrement les méthodes allant de la fouille d'opinion ou *opinion mining* à la visualisation des données dans les tableaux de bord de la réputation. Si la mesure de la réputation peut dans certains cas s'effectuer de manière manuelle[229], la plupart des acteurs utilisent une approche automatique[230] tant sur la collecte que l'analyse des données. Mesurer la réputation suppose donc, dans une première approche, un paramétrage c'est-à-dire le choix des diverses sources à surveiller (les forums, les réseaux sociaux, les sites de presse, les bases de données, etc.), la catégorisation des sources, la granularité et la régularité de la collecte, et la mise en place d'un système de requêtage. La phase de collecte des données utilise un crawler du web, et les données issues de la collecte sont alors nettoyées, indexées, catégorisées puis archivées. La phase d'analyse peut prendre plusieurs formes selon les objectifs : co-occurence de termes, détection de communautés,

[229] Pour une minorité des acteurs et le plus souvent en complément de l'approche automatique.

[230] Avec l'utilisation de solutions proposées par les éditeurs de logiciels de réputation

entités nommées, analyse de sentiment, etc., et les résultats sont présentés sous forme de visualisations dans les tableaux de bord de la réputation.

L'objectif central de nos travaux, est, rappelons-le, de déterminer si les professionnels de l'e-réputation se sont inspirés dans leurs approches, de savoirs académiques et plus précisément des modèles théoriques. La deuxième partie de nos travaux permettra de proposer des éléments de réponses.

PARTIE II : DES MODÈLES THÉORIQUES AUX MÉTHODES DE MESURE

Introduction de la partie II

L'objectif principal de nos travaux, rappelons-le, est de déterminer si les pratiques professionnelles de mesure de la réputation se sont inspirées, peu ou prou, de savoirs académiques. Afin d'y parvenir, il est essentiel d'identifier les grandes dimensions de classification des modèles, catégories qui seront par la suite utilisées dans l'approche d'analyse comparative avec les méthodes professionnelles proposées.

Il est important de souligner que plusieurs auteurs de la littérature scientifique, surtout anglo-saxonne, se sont déjà intéressés aux modèles conceptuels de l'e-réputation, en proposant des grandes dimensions de classification de ces derniers. Dans le premier chapitre de cette partie, nous proposerons d'abord une méta analyse sur les grands concepts théoriques des dimensions de la réputation. Ensuite, la deuxième partie de ce chapitre sera consacrée au processus de mesure de l'e-réputation basé sur une approche structurelle et méthodologique des modèles théoriques. Dans la dernière partie, nous proposerons les quatre grandes dimensions de classification des modèles issues de notre analyse et sur lesquelles nous nous appuierons dans le processus de comparaison avec les méthodes professionnelles de mesure. Nous présenterons à l'issue de cette première phase d'analyse et à partir des dimensions et les items qui servent à les étayer, un tableau de classification des modèles conceptuels que vous retrouverez en annexe 1.

Le second chapitre de cette partie vise à présenter un état des lieux du marché de l'e-réputation. Le marché de la réputation, nous le verrons, a connu depuis les années 2009 un essor, avec une offre et des prestataires multiples et variés. Nous tenterons donc de proposer dans un premier temps une réflexion sur les mutations du marché de l'e-réputation. Ensuite, et dans un second temps, nous présenterons une cartographie de l'ensemble des acteurs qui évoluent dans le marché puis, nous identifierons enfin, l'ensemble des services proposés par les acteurs du marché de la réputation.

CHAPITRE 4 : LES MODÈLES CONCEPTUELS DE MESURE DE LA RÉPUTATION EN LIGNE, PROPOSITION D'UNE TAXONOMIE

Introduction du chapitre 4

L'une des problématiques à laquelle nous nous sommes heurtée dans la proposition d'une taxonomie des modèles de réputation est d'arriver à identifier ceux qui relèvent réellement de la mesure de la réputation. En effet, en explorant la littérature sur la thématique "modèle de réputation *or reputation model*", celle-ci ne fait *a priori* aucune distinction entre la confiance et la réputation, et les modèles qui en découlent. En effet, s'il est une question sur laquelle la plupart des auteurs s'accordent, c'est que les concepts de réputation et de confiance sont deux notions intrinsèquement liées, qui font souvent l'objet de confusion (Abdul-Rahman & Hailes 2000 ; Mui & al., 2002 ; Dellarocas 2003; Lu et al. 2009). Yu & al. (2013) proposent d'ailleurs, à défaut d'arriver à déterminer des définitions permettant de distinguer les deux concepts, de ne pas définir la relation les caractérisant et de ne pas les différencier. En approfondissant nos recherches, plusieurs courants de pensée s'entrecroisent.

Pour certains auteurs, la réputation est perçue comme le degré de confiance qu'un individu est capable de générer dans un domaine particulier, et est de fait perçue comme un *output* de la confiance (Zaccharia et al., 1999). Pour Lu et al. (2009), elle est plutôt une confiance non privée ou socialisée, c'est-à-dire une mutation du concept de confiance à l'ère du numérique. Pour plusieurs autres auteurs *en outre*, la réputation est un *input* important ou un facteur contribuant au calcul de la confiance d'une entité (Abdul-Rahman and Hailes, 2000 ; Dellarocas, 2003 ; Kravari and Bassiliades, 2016 ; Pinyol and Sabater-Mir 2013b ; et Ruohomaa et al., 2007). Il est indispensable pour nous de présenter ces diverses approches en amont, notre méthodologie de recherche en dépendant. Afin d'identifier les modèles qui nous serviront dans nos travaux, nous nous sommes appuyée sur la troisième approche.

Afin de faciliter la reproductibilité de nos travaux, nous nous sommes appuyée sur un protocole méthodologique précis. L'objectif principal de notre recherche étant d'identifier dans la recherche scientifique les modèles théoriques de mesure de la réputation, nous nous sommes appuyée sur plusieurs sources de données de la recherche académique : ACM Digital Library, Hal, SCOPUS, Science Direct, et Web of Science, Researchgate, link.springer,

ieeexplore.ieee, soit des bases de données ou sites de ressources choisis pour leur large couverture et la pertinence des travaux sur la thématique de la modélisation de la réputation.
Nous avons par la suite utilisé comme requête “modèle de réputation or *reputation model or reputation assessment model or online reputation model*” qui, contrairement à ce qu’on serait tenté de penser, nous a permis d’obtenir des données assez complètes tout en évitant les résultats indésirables avec notamment l’utilisation des guillemets afin que les sources de recherche choisies ne sélectionnent que les articles contenant la requête sus-citée. Nous avons choisi de nous intéresser aux modèles développés entre 2000 et 2016 (nos travaux de recherche ayant débuté en novembre 2015).

Enfin, nous nous sommes assurée que chaque article fournisse une mise en œuvre ou à minima une simulation indiquant que l'approche a été déployée et/ou expérimentée, afin de garantir que les modèles proposés peuvent être mis en œuvre et, sur la base des tests publiés, leur performance approximative peut être estimée et donc de fournir une garantie supplémentaire de la rigueur des travaux. Des critères d'exclusion supplémentaires n'ont pas été pris en compte et si des raisons quelconques obligeaient à omettre un article, nous les exposerons en annexe.
Cette première phase de collecte de données nous a permis de collecter un certain nombre de documents (189) qu’on peut classifier selon trois principales catégories : les articles relatifs aux modèles de mesure de la réputation, ceux relatifs à la mesure de la confiance et de la réputation (les modèles hybrides) et enfin, ceux relatifs uniquement à la mesure de la confiance. Comme mentionné au début de cette partie, nous partageons le point de vue du courant de pensée selon lequel la confiance et la réputation sont deux concepts distincts, mais complémentaires, avec la confiance comme *output* de la réputation. Dans cette optique, sur les 189 articles collectés, 85 ont été utilisés dans nos travaux : ceux qui traitaient uniquement de la mesure de la réputation, mais aussi ceux sur la mesure de la confiance et de la réputation donc les modèles hybrides. Les 104 documents restants n’ont pas été utilisés pour deux raisons : soit parce qu’ils ne traitaient pas spécifiquement de la mesure de la réputation et n’étaient pas non plus relatifs à des modèles hybrides (mais se référaient uniquement à la mesure de la confiance), soit parce qu’ils étaient un duplicata d’articles retenus ou des

travaux de synthèse de modèles qui figurent dans notre échantillon. La liste des modèles retenus figure en annexe 1.

Les sources académiques les plus exhaustives sur notre thématique restent indéniablement Researchgate, link.springer, ieeexplore.ieee, et si on s'intéresse de plus près aux articles, on constate que la thématique de la mesure de la réputation a été étudiée pour divers usages allant du e-commerce (Abdel-Hafez et al., 2015), à la *supply chain* (Chang et al., 2015), en passant par la sécurité des réseaux (Skoric et al., 2016), l'internet des objets (*IoT),* aux approches plus généralistes notamment en communication(Liu et al., 2013).

À l'issue de la phase de collecte de l'ensemble du corpus et afin d'analyser et de proposer une classification des modèles théoriques de réputation, il nous a semblé nécessaire de nous appuyer sur des matrices dites conceptuelles visant à comparer différents articles sur la base d'un ensemble fixe de concepts (Webster et Watson, 2002). Partant de ce constat, ce premier chapitre s'articule autour de trois parties : dans un premier temps, nous proposerons une revue des diverses approches de classification des modèles conceptuels proposés par les chercheurs qui se sont également intéressés à la question. Nous nous intéresserons par la suite, dans un second temps, au processus de mesure de la réputation tel que conçu par la littérature scientifique. Cette deuxième partie s'appuie sur une analyse structurelle et méthodologique, et vise notamment à déterminer les différents contextes d'utilisation des modèles, les différentes étapes du processus d'évaluation (du *sourcing* à la catégorisation) et les méthodes utilisées dans le calcul de la réputation.

Enfin, nous présenterons dans la troisième partie les quatre grandes dimensions de classification des modèles issues de notre analyse et notre tableau de classification des modèles conceptuels que vous retrouverez en annexe 1. Les dimensions de classification et notamment les attributs qui les composent serviront par la suite (nous y reviendrons dans le chapitre 7) à déterminer si les méthodes professionnelles s'appuient sur des savoirs académiques, et à étayer notre problématique principale de recherche.

1. Revue de la littérature : les approches scientifiques de classification

Plusieurs chercheurs (Sabater et Sierra 2005 ; Koutrouli et Tsalgatidou 2006 ; Marti & Garcia-Molina, 2006; Wang & Vassileva, 2007 ; Hoffman & al., 2009 ; Yuan & al., 2012 ; Vavilis & al., 2014 ; Hendrikx & al., 2015 ; Granatyr & al. 2015) se sont attelés avant nous, à proposer dans leurs travaux, une taxonomie des modèles théoriques de confiance et de réputation, à partir de plusieurs dimensions. Cette partie vise à présenter les différents attributs autour desquels sont construites les catégories de classification proposées dans la littérature scientifique.

1.1. Les dimensions de classification de Sabater et Sierra, 2005

Sabater et Sierra (2005)[231], dont les travaux sur les modèles théoriques de mesure de la réputation et de la confiance, et notamment les dimensions d'analyse, ont servi de base à plusieurs auteurs, proposent dans leur article sept (7) principales dimensions de classification : le type de modèle ; les sources d'information ; le type de visibilité, la granularité des modèles ; la prise en compte des biais ; le type d'information échangée et enfin la fiabilité de la valeur de la réputation ou confiance.

- **Le type de modèle** : Sabater et Sierra distinguent deux typologies de modèles, cognitifs et numériques. Les modèles construits autour d'une approche cognitive sont basés, pour Esfandiari et Chandrasekharan (2001) qu'ils citent, sur des croyances sous-jacentes avec pour composantes majeures les "états mentaux", affectifs et émotions. Les modèles numériques, qui utilisent la théorie des jeux, se fondent sur l'utilisation et l'agrégation numérique des données de réputation ;

[231] Sabater, Jordi, et Carles Sierra. « Review on Computational Trust and Reputation Models ». *Artificial Intelligence Review 24*, nº 1 (septembre 2005): 33-60. https://doi.org/10.1007/s10462-004-0041-5.

- **Les sources d'information :** Sabater et Sierra en distinguent quatre principales : les données issues d'expériences directes qui comprennent à la fois les interactions directes et observées entre plusieurs agents ; les témoignages qu'ils qualifient de bouche-à-oreille ou informations indirectes et qui selon eux sont les plus abondants et complexes à appréhender dû à un degré d'incertitude et de fiabilité élevé ; les données relatives à la récurrence des interactions ; et enfin ce qu'ils qualifient de "prejudice" (les préjugés), l'ensemble des signes permettant de reconnaître, identifier ou catégoriser un individu (CSP, sexe, etc.) ;

- **Le type de visibilité** : la valeur de la réputation ou confiance attribuée à un agent peut pour Sabater et Sierra, soit être subjective ou globale. La visibilité globale est la résultante (somme, moyenne, etc.) du calcul de l'ensemble des opinions des individus ayant interagi avec une entité dans le passé, valeur accessible par tous les membres d'une "communauté" et automatiquement mise à jour à chaque nouvelle interaction. Il s'agit de la valeur finale de la réputation. La visibilité est dite subjective lorsqu'elle est relative à chaque donnée de réputation (avis d'un internaute, commentaire, etc.) prise de manière individuelle. On parlera dès lors dans ce dernier cas, de réputation d'un agent X selon un agent Y ;

- **La granularité des modèles** : par granularité, les auteurs entendent la notion de contexte d'utilisation. Ils distinguent donc les modèles à contexte unique qui pour eux sont conçus pour associer une valeur unique de confiance / réputation et adaptés à un seul contexte d'utilisation ; et ceux qui sont multi contextes c'est-à-dire qui sont appliqués à divers environnements (les approches modulaires, complètes) avec une pluralité de valeurs de confiance et de réputation ;

- **La prise en compte des biais** : un autre critère de classification pris en compte par les auteurs est la notion de fiabilité envers les agents, à travers notamment la mesure de leur aptitude à fournir des informations exactes, biaisées ou fausses. Ils distinguent trois (3) niveaux de classification : le niveau 0 qui prend en compte les modèles dans lesquels les biais ou informations erronées/fausses ne sont pas prises en compte, ni

traitées et où les modèles s'appuient sur un grand nombre d'agents offrant des évaluations honnêtes pour contrer l'effet potentiel des évaluations fournies par des agents malveillants ; le niveau 1 dans lequel les agents peuvent cacher ou biaiser des informations, mais peuvent plus difficilement fournir des informations mensongères ; et enfin le niveau 2 avec des modèles qui admettent que certains agents puissent mentir ou cacher volontairement des informations et disposent donc de mécanismes spécifiques pour traiter les menteurs ;

- **Le type d'informations échangées** : il s'agit ici du type d'informations collectées, choix essentiel dans la conception des modèles. Les auteurs distinguent donc les modèles qui reposent sur des méthodes probabilistes et fonctionnent avec des informations issues de variables quantitatives, et ceux basés sur des mécanismes d'agrégation de variables qualitatives ;

- **La fiabilité de la valeur de la réputation ou de la confiance** : il est ici question de mécanismes utilisés dans la conception des modèles et permettant de garantir la fiabilité des résultats des données de réputation / confiance. Cette mesure, qui pour Sabater et Sierra se présente sous forme de valeur unique, est par la suite intégrée à la valeur de réputation/confiance afin d'obtenir une évaluation finale. Selon les modèles, les éléments pris en compte dans le calcul de cette valeur peuvent varier[232].

[232] Selon les modèles, plusieurs données peuvent être prises en compte : le nombre d'avis / commentaires laissés par la personne qui évalue, les avis laissés par les autres entités, etc.

1.2. Les dimensions de classification de Koutrouli et Tsalgatidou 2006 / 2012

Koutrouli et Tsalgatidou proposent dans leur article une architecture de comparaison des systèmes de confiance *peer to peer* basés sur la réputation. Ils distinguent trois (3) principales composantes : la collecte d'informations ; l'estimation de la réputation et la représentation de la réputation.

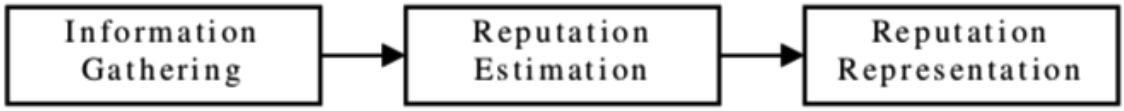

Figure n°7 : Catégories de classification des modèles, source Koutrouli et Tsalgatidou

Sur le premier élément, les **données à collecter**, ils identifient cinq (5) principaux critères :

- Le format des informations : il s'agit ici du type de données à collecter, des sources de l'information et surtout de son lieu de stockage ;
- La fiabilité des données : il s'agit dans ce cas de figure de garantir l'intégrité des informations collectées et stockées localement, en évitant qu'elles ne puissent être manipulées par des agents malicieux ;
- La crédibilité de l'agent qui recommande : identifier et distinguer les agents qui fournissent des évaluations biaisées ou fausses ;
- Le type de données à privilégier : si la mesure de la réputation s'effectue à partir de l'agrégation des informations issues de transactions positives ou négatives, les auteurs proposent de prendre en compte plus les données négatives du fait de leur impact) ;
- De la notion de contexte qui renvoie à celle perçue au sens de Sabater et Sierra (2005).

Le second aspect, **la mesure de la réputation**, prend en compte cinq critères :

- La gestion des nouveaux entrants : il s'agit ici des mécanismes de gestion des nouveaux agents au sein d'un système et qui ne possèdent de fait aucune réputation ;

- La portée des informations de confiance (informations globales et objectives ou locales et subjectives) qui renvoie au concept de visibilité de Sabater et Sierra (2005) ;
- La méthode de calcul : il s'agit des méthodes de mesure de la réputation qui vont des fonctions statistiques simples (par exemple la moyenne) aux méthodes probabilistes en passant par les logiques floues, etc. ;
- La chaîne d'évaluation : il est ici question des différents niveaux composant la chaîne de recommandation indirecte dans un système ;
- La récence des données : la récence des transactions et donc des évaluations devrait primer selon les auteurs dans la collecte des données de mesure de la réputation.

Et enfin, la dernière composante à prendre en compte dans la comparaison des modèles de réputation, la **représentation de la valeur de la réputation** qui prend en compte :

- Le type de valeurs : les valeurs de confiance ou réputation peuvent être, pour Koutrouli et Tsalgatidou, des variables discrètes ou continues ;
- La représentation de la confiance : il est ici question de mettre en place des mécanismes permettant d'isoler des agents malveillants. Dans certains modèles, le degré de confiance / méfiance en une entité est explicitement représenté, soit comme une donnée spécifique des valeurs de réputation, soit en complément.

1.3. Les dimensions de classification de Marti and Garcia-Molina, 2006

Sergio Marti et Hector Garcia-Molina[233] distinguent quant à eux, trois principales composantes de classification des systèmes de réputation : la collecte des informations (*information gathering*) ; le processus d'attribution d'un score de réputation (*reputation scoring and ranking*) et le feedback ou *response* (récompense ou punition).

[233] Sergio Marti and Hector Garcia-Molina (2005), Taxonomy of Trust: Categorizing P2P Reputation Systems. Stanford University. Accessible en ligne sur http://ilpubs.stanford.edu:8090/675/1/2005-11.pdf

- **La collecte d'informations** : première composante pour les auteurs, elle vise à attester de la fiabilité d'un agent lors d'une interaction. Sont ici pris en compte, les éléments relatifs à **l'identité** de l'entité (avec la mise en place de processus d'anonymat et de pseudonymat), à l'ensemble **des informations** collectées (notamment les sources d'information quantitatives ou qualitatives et leur fiabilité), et à **la gestion des nouveaux agents** (avec ici deux principales stratégies : croire en la bonne volonté des nouveaux agents pour les optimistes et contribuer à construire leur réputation, ou, pour les pessimistes, refuser d'interagir avec eux) ;

- **Le *scoring* et le *ranking*** : attribuer pour Marti et Garcia-Molina un score de réputation à un agent à l'issue d'une interaction dépend de plusieurs facteurs. Ils identifient 4 éléments à prendre en compte : **les *inputs*** (les données utilisées pour mesurer la réputation qui sont issues de l'historique des transactions et qui peuvent être positives, négatives ou neutres) ; les ***outputs*** (la valeur finale de la réputation qui peut être une valeur binaire, une échelle continue, etc.), la notion de **temporalité** dans la collecte des données (mises à jour des données et récence à prendre en compte), et le **classement** (ranking qui vise de manière générale à identifier les entités qui ont le score de réputation le plus élevé) ;

- **La participation** : les systèmes de réputation doivent, selon les auteurs, permettre de récompenser ou motiver les agents qui contribuent de manière positive et honnête, ou, sanctionner les autres. Ils identifient donc 2 types de mécanismes qui peuvent être utilisés : la **motivation** (à travers un avantage attribué, réduction des coûts lors d'une prochaine interaction par exemple) et la **punition** (les systèmes de réputation devraient en amont pouvoir identifier les agents malicieux et les punir : les bannir du système étant la punition ultime).

1.4. Les dimensions de classification de Wang and Vassileva, 2007

Wang et Vassileva[234] proposent dans leur article d'analyser les modèles de confiance et de réputation et de les comparer selon leur structure. Ils distinguent donc trois principales dimensions de classification hiérarchiques des modèles : la notion de centralisation et décentralisation ; l'objet sur lequel porte l'évaluation (une personne / un agent versus une ressource) et enfin, la nature de l'opinion collectée (personnalisation versus global).

- **L'approche centralisée versus décentralisée** : comme le soulignent plusieurs auteurs, un système de confiance et de réputation peut être centralisé ou décentralisé. Pour Marti et Garcia-Molina, dans un système centralisé, le nœud central assumera toutes les responsabilités de la gestion de la réputation de tous les membres, tandis que dans un système décentralisé (peer-to-peer par exemple), il n'existe pas de nœud central. Les membres du système doivent coopérer et partager les responsabilités pour gérer la réputation. Ils soulignent en outre que les mécanismes des systèmes centralisés sont moins complexes et plus faciles à mettre en œuvre que ceux des systèmes décentralisés. Mais ils ont besoin de serveurs centraux puissants et fiables pour le stockage de données et leur diffusion ;

- **Le couple personne / agent versus une ressource** : les systèmes de confiance / réputation peuvent également, pour Wang et Vassileva, être classés selon qu'ils permettent d'évaluer la réputation d'une personne ou d'une ressource. Le premier type de modèle qui modélise la réputation d'une personne physique, est souvent assimilé au *personal branding* (aux techniques de mise en avant de son profil sur les différentes plateformes numériques) ; et le second est relatif à la mesure de la réputation d'un produit ou service, mesure qui peut dans certains cas être liée à la réputation de l'agent qui le propose ;

---

[234] Wang, Yao & Vassileva, Julita. (2007). A Review on Trust and Reputation for Web Service Selection. Proceedings - International Conference on Distributed Computing Systems. 25-25. 10.1109/ICDCSW.2007.16. DOI:10.1109/ICDCSW.2007.16

- **Personnalisé versus global** : Wang et Vassileva identifient un troisième critère de classification des systèmes de réputation qui peuvent être personnalisés ou généraux. Dans les systèmes dits généraux, la réputation d'une personne, d'un agent ou d'un produit / service est basée sur les opinions de la population dans leur ensemble, évaluation qui est publique et visible pour tous les membres du système. Il renvoie au concept de visibilité proposé par Sabater et Sierra (2005).

1.5. Les dimensions de classification de Hoffman et al., 2009

Hoffman, Zage et Nita-Rotaru (2009)[235] distinguent trois (3) dimensions de classification des modèles : la collecte (*formulation*), la méthode de calcul (*calculation*)et la diffusion des résultats (*dissemination*).
La première dimension, **la formulation (ou la collecte)**, utilise quatre principaux critères: le type d'informations, les sources de l'information, la métrique de la réputation et la temporalité.

- **Les sources d'information**: les auteurs identifient deux principales sources d'informations brutes utilisées dans la mesure de la réputation. Ce sont : les sources manuelles et celles automatiques.

  Les sources manuelles, qui sont relatives aux informations qualitatives, sont obtenues à partir des commentaires laissés par les agents et le rôle de la composante de reformulation est donc de convertir les métriques qualitatives en quantitatives ou d'associer la métrique quantitative avec des informations qualitatives dans l'intention d'agréger ces informations avec la métrique de réputation.

  Les sources automatiques, elles, sont issues d'observations directes (faites par un tiers et résultant de données directement observées sur une transaction ou interaction) ou indirectes (données indirectement perçues ou déduites d'observations directes. Elles sont souvent utilisées en cas de manque ou d'insuffisance d'informations) ;

---

[235] K. Hoffman, D. Zage, and C. Nita-Rotaru. A survey of attack and defense techniques for reputation systems. *ACM Computing Surveys*, 41(4), December 2009

- **Le type d'information**: il est relatif à la nature des données à collecter qui peuvent soit être positives (donc favoriser l'établissement d'une relation de confiance) ou négatives (induire un manque de confiance) ;
- **La métrique de réputation**: la représentation mathématique ou algorithmique de la réputation est pour Hoffman et al. la composante la plus importante. Ils identifient trois (3) principaux formats de la métrique de réputation : binaire, discrète ou continue. Dans la représentation binaire qui n'utilise que deux caractères (0,1), les données qualitatives sont converties au format numérique pour être calculées. Dans les systèmes utilisant des métriques discrètes, on distingue différents niveaux de réputation prédéfinis et présentés sous forme de nombres entiers (1, 2, etc.). Enfin, les métriques représentées sous forme de variable continue peuvent prendre une infinité de valeurs sur un intervalle donné ;
- Et enfin, les **aspects temporels** qui font référence à la prise en compte de la récence des informations dans le calcul.

La deuxième dimension, celle **de calcul**, est pour Hoffman et al., la partie concrète du système de réputation et peut être perçue comme la résultante de celle de formulation. Elle comprend deux composantes essentielles : la structure de la dimension de calcul (centralisée ou distribuée) et la méthode de calcul (déterministe ou probabiliste).

- **La structure des systèmes dans la dimension de calcul**: le calcul de la réputation peut s'effectuer en prenant en compte une « autorité centralisée » ou *a contrario*, à partir d'une pléthore de participants, et donc s'effectuer de manière distribuée. Les systèmes de réputation centralisés qui sont utilisés entre autres par plusieurs plateformes de e-commerce par exemple, sont d'une utilisation simple et offrent une marge de manœuvre de manipulation des données très faible aux agents malicieux. Cependant, une approche centralisée repose sur l'idée que les membres du système ont une confiance aveugle en l'autorité centralisée qui, en retour, se doit d'être fiable et toujours disponible.

La plupart des systèmes de réputation utilisent par contre, pour Hoffman et al., une approche distribuée qui, comme ils le soulignent, s'avère plus complexe. En effet, la complexité de l'authentification des données et des entités dans les systèmes dépourvus d'autorité centralisée offre aux agents malveillants la possibilité de biaiser les données de calcul de la réputation ;

- **De l'approche utilisée dans la dimension de calcul** : les systèmes de réputation utilisent selon Hoffman et al., une approche déterministe ou probabiliste. Le résultat d'un calcul déterministe peut être déterminé uniquement à partir de la connaissance des données d'entrée ou *inputs*, ce type de système étant statique et réagissant de la même manière face à un phénomène donné. En effet, les calculs déterministes pour les valeurs de réputation globale supposent qu'à un ensemble de phénomènes entrants donné, sont indéniablement associés un ensemble de phénomènes sortants dans le même ordre chronologique.

  Les approches probabilistes mobilisées dans l'évaluation de la réputation elles, *a contrario*, s'appuient notamment sur la théorie des jeux, et supposent qu'un phénomène donné est associé à une ou plusieurs valeurs. Ce comportement aléatoire dépend du hasard, et la valeur attribuée à l'événement se présente sous forme de nombre réel compris entre 0 et 1[236].

La dernière dimension est relative à celle de **diffusion** et vise à rendre accessibles les données calculées aux agents évoluant dans un système, en préservant les informations de l'altération. Souvent liée à la dimension de calcul dans l'implémentation des méthodes, elle prend en compte quatre (4) étapes : la structure de diffusion, l'approche de diffusion, les stratégies de stockage et la redondance[237].

- La **structure de diffusion**: pour Hoffman et al., les mécanismes de diffusion centralisés impliquent une autorité centrale qui stocke et diffuse des valeurs calculées. L'autorité centrale peut en réalité être représentée par des graphes, avec une approche classifiée comme centralisée, car l'ensemble de la structure est contrôlée

[236] Le plus grand nombre représentant la probabilité que le phénomène se produise et inversement.

[237] Traduit de l'anglais : dissemination structure, dissemination approach, Storage strategies et dissemination redundancy

exclusivement par une seule entité. Par exemple, pour qu'eBay puisse évoluer, il doit être implémenté sur des clusters. Toutefois, logiquement, eBay utilise une autorité centralisée pour diffuser les informations de réputation calculées. Dans une diffusion centralisée, l'autorité centrale a plus de pouvoir pour protéger l'intégrité du processus, mais elle devient également un point de faiblesse unique face à certaines attaques externes. Faute de quoi, les dommages causés au système de réputation sont bien plus importants que si le processus était distribué. Dans une diffusion distribuée, chaque participant est responsable d'une partie des valeurs de réputation calculées. La répartition des responsabilités peut être symétrique[238] ou asymétrique. La diffusion distribuée est intrinsèquement plus vulnérable à la manipulation et utilise souvent la redondance des données, des mécanismes cryptographiques ou d'autres mesures pour préserver l'intégrité métrique. Les mécanismes distribués sont pour les auteurs, plus difficiles à mettre en œuvre et à tester correctement, et peuvent donc être plus vulnérables à l'exploitation ;

- **L'approche de diffusion:** Le modèle de communication des mécanismes de diffusion peut être qualifié de déterministe (notamment dans les modèles comme EIGENTRUST, 2003 ; POWERTRUST, 2007 ou PEERTRUST, 2003 par exemple) ou de probabiliste (on peut citer à titre d'exemples les techniques de diffusion basées sur les épidémies[239] et les inondations[240]) ;

- **Les stratégies de stockage:** Le stockage transitoire est défini comme toute mémoire à accès aléatoire non contrôlable, tandis que le stockage permanent est tout stockage dans lequel les données sont préservées même pendant les périodes sans alimentation. En fonction de la volatilité des composants du système et de la complexité des calculs, il peut être avantageux de stocker les valeurs de réputation calculées dans un stockage permanent pour les récupérer ultérieurement. Hoffman et al. notent que le fait qu'un système utilise ou non le stockage permanent est davantage un problème de

[238] Tableaux de hachage distribués, par exemple Ratnasamy et al., 2000

[239] C'est notamment le cas pour les modèles de Eugster et al. (2001) ou Walsh et Sirer (2006)

[240] On peut, par exemple, citer les modèles de Marti et Garcia-Molina (2004) , Lee et al. (2003)

mise en œuvre qu'un composant central d'un système de réputation. Les mécanismes de calcul peuvent nécessiter un stockage permanent pour détecter les changements de comportement lents sur de longues périodes. En outre, le stockage permanent doit être protégé contre les manipulations malveillantes, la corruption des données physiques et la perte de données ;

- **La notion de redondance:** Le degré de redondance intégré dans les mécanismes de diffusion est un compromis entre la résilience à la manipulation et l'efficacité du stockage / de la communication. La redondance peut être utilisée dans bon nombre des composants de la dimension de la diffusion, telle que la messagerie redondante dans les modèles de communication ou les sauvegardes en double des valeurs stockées. La plupart des systèmes de réputation modernes utilisent des protocoles de messagerie avec des messages redondants pour assurer la livraison des messages et fournir une certaine résilience aux nœuds malveillants au prix d'une surcharge de communication. Chaque nœud demande et reçoit plusieurs copies d'une valeur de réputation (à partir de différents emplacements de stockage), ce qui augmente la contrainte de liaison du réseau lorsqu'il est visualisé par rapport à la messagerie monodiffusion traditionnelle à partir d'un seul emplacement de stockage. Les méthodes moins efficaces de mise en œuvre de la redondance (par exemple, la duplication complète des données) sont souvent préférées aux méthodes plus souhaitables en théorie, telles que les codes de Reed-Solomon [Reed et Solomon 1960], car ces méthodes sont souvent plus faciles à mettre en œuvre. Enfin, les systèmes diffèrent par la manière dont ils résolvent la redondance pour produire une métrique de réputation finale. Dans certains modèles, les données sont présentées de manière brute sans faire cas de redondance.

1.6. Les dimensions de classification de Yuan et al., 2012

Dans la même optique que Hoffman et al., Yuan Yao, Sini Ruohomaa et Feng Xu[241] identifient trois principales dimensions de composition des systèmes de réputation : la diffusion d'informations, le calcul et l'évolution[242].

Sur la **diffusion d'informations**, les auteurs notent qu'elle permet aux participants de tirer parti des expériences des autres, même s'ils le soulignent, les informations directes sont plus fiables, mais pas toujours accessibles au moment de la décision et celles de tierces parties parfois inexactes et ayant donc moins de poids dans le processus de décision. Le processus de diffusion peut s'effectuer de deux principales manières : soit la méthode *pull* lorsqu'un utilisateur reçoit une demande de recommandations ou encore la méthode *push* nécessitant un déclencheur interne permettant d'envoyer activement de nouvelles informations. Les différents aspects pris en compte dans cette dimension sont:

- Les **métriques** qui peuvent être discrètes ou continues ;
- La **structure** (centralisée, décentralisée ou hybride) ;
- La notion de **réseau en superposition** dans lequel le réseau aide à gérer les défis liés au stockage et à la diffusion des informations distribuées. Tandis que la plupart des systèmes *peer-to-peer* sur Internet sont non structurés, les auteurs le notent, plusieurs des systèmes qu'ils ont étudiés proposent une approche de stockage des données très structurée et organisée ;
- Les **recommandations**: qui peuvent prendre quatre formes (simple notation d'une transaction ; décompte d'interactions satisfaites ou non ; agrégation de moyennes pondérées ; et enfin les avis génériques ou recommandations standard.

Dans cette dimension, la confidentialité et l'anonymat des utilisateurs est un enjeu majeur, de même que la protection du système face aux attaques extérieures (détecter les avis frauduleux, éviter la suppression d'avis défavorables, etc.).

---

[241] Yuan Yao, Sini Ruohomaa, and Feng Xu, Addressing Common Vulnerabilities of Reputation Systems for Electronic Commerce, Journal of Theoretical and Applied Electronic Commerce Research. ISSN 0718-1876 Electronic Version VOL 7 / ISSUE 1 / APRIL 2012 / 1-20 © 2012 Universidad de Talca - Chile. Accessible en ligne sur https://scielo.conicyt.cl/scielo.php?script=sci_arttext&pid=S0718-18762012000100002

[242] Traduit de l'anglais "dissemination, calculation, evolution"

La deuxième **dimension, celle de calcul,** à partir des informations recueillies et analysées de la première dimension, sert de support d'aide à la décision aux utilisateurs. Dans cette dimension, cinq (5) principaux éléments sont essentiels à la décision d'entrer en relation avec un agent : l'approche choisie ; la notion de confiance / méfiance (les deux approches, pas de méfiance, etc.) ; la confiance en soi (et notamment aux valeurs des données de confiance/ méfiance calculées par l'agent lui-même) ; la notion de contexte (la valeur, le temps de livraison, qualité, prix, etc.) et la crédibilité de la recommandation (sont à prendre en compte la source et le contenu de la recommandation afin d'attester de sa fiabilité). Les auteurs soulignent que l'élément central de cette dimension est l'approche algorithmique choisie qui comprend cinq (5) composantes:

- La **méthode de comptage**: il s'agit de compter les évaluations positives et négatives avec comme valeur de la réputation la somme ou moyenne de toutes les notes ;
- L'**approche probabiliste**: avec notamment la théorie des jeux, via l'identification d'un ensemble de valeurs associées à un phénomène ;
- L'**approche floue**: basée sur la logique floue (*fuzzy logic*), les systèmes qui utilisent cette approche permettent de traiter les incertitudes liées aux transactions et notamment de mesurer la crédibilité des recommandations et la réputation du système [243] et de calculer des scores de confiance locaux ou universels[244] ;
- L**'approche par flux (*flow based*)**: dans laquelle le calcul de la confiance prend en compte les noeuds et notamment la corrélation entre eux et la manière dont se propage la confiance (chaînes en boucle, longues, etc.) ;
- Les **autres approches**: certains systèmes, selon les auteurs, ne prennent en compte que les expériences directes et emploient une approche d'apprentissage par renforcement[245].

Afin de protéger le système, les deux éléments importants dans cette dimension sont la situation dans laquelle la transaction a lieu et la crédibilité des informations partagées.

---

[243] On peut citer à titre d'exemple le modèle REGRET de Sabater et Sierra

[244] Dans le modèle FUZZYTRUST par exemple

[245] C'est notamment le cas du modèle de Tran et Cohen

Et enfin, la troisième catégorie, la **dimension évolution** est relative à la manière dont le système tire parti des nouvelles données acquises (des nouvelles interactions) afin de mettre à jour les données de réputation. Les différents aspects à prendre en compte dans la dimension évolution de la réputation sont : la mise à jour des connaissances locales sur la base de nouvelles entrées ; le déclenchement de mises à jour de la réputation globale (avec une date d'activation du processus de diffusion) sur la base de nouvelles informations ; et enfin la détection et la réaction aux changements de comportement des individus. Certains systèmes comme eBay offrent la possibilité aux utilisateurs de saisir manuellement leurs opinions personnelles dans le système, et d'autres notamment dans les environnements *peer-to-peer*, utilisent un système de détection automatisée des transactions réussies. La fin d'une transaction ou interaction est, de manière générale, le déclencheur de mise à jour des valeurs de la réputation, si tant est qu'il soit spécifié dans les systèmes.
En sus des trois (3) dimensions mentionnées, Yuan et al. identifient six (6) pré-requis de sécurité et robustesse des systèmes de réputation :

- L'**authenticité, l'intégrité et la confidentialité des messages** : pour que la crédibilité de la source d'informations soit prise en compte, elle doit pour Yuan et al. être connue de manière fiable ; les informations ne doivent pas être altérées ; et, pour limiter les représailles et la réciprocité, elles ne doivent pas être divulguées à des acteurs autres que ceux spécifiés dans le système ;
- La **protection contre la falsification et la mise à disponibilité des informations de réputation stockées** : les données de réputation peuvent soit être stockées localement, en externe sur un serveur ou encore réparties sur un réseau de pairs ;
- La **gestion des identités par effet de levier** : il s'agit ici d'empêcher à un acteur spécifique d'accroître son influence dans le système de réputation en créant de nouvelles identités ;
- La **prise en compte du contexte dans le processus de décision** : certains systèmes de réputation sont multicontextes et d'autres peuvent être utilisés dans un unique contexte ;
- L'**évaluation de la crédibilité de la recommandation** : selon les auteurs, dans les systèmes de réputation distribués, les agents sont susceptibles de mentir, fournir des

informations biaisées, produire des interprétations erronées, etc., et les moniteurs automatisés peuvent produire des données incorrectes. Les systèmes doivent donc mettre en place des stratégies permettant de mesurer la crédibilité des recommandations ;

- La **gestion des incohérences dans l'évolution de la réputation** : il s'agit ici, et ce à partir des informations de réputation concernant les interactions passées d'un agent, de prédire le comportement futur de ce dernier tout en évitant les incohérences.

1.7. Les dimensions de classification de Vavilis et al., 2014

À partir des travaux de Witkowski et al. (2011) et Smeltzer et al. (1997), Vavilis et al.[246] s'appuient sur les **trois (3) dimensions** de classification des modèles proposées par Hoffman et al., (2009), c'est-à-dire la formulation, le calcul et la diffusion pour distinguer **treize (13) pré-requis** qu'ils classifient dans les trois dimensions sus-citées, et **dix (10) fonctionnalité**s à prendre en compte dans la mesure de la réputation.

Des pré-requis au bon fonctionnement d'un système de réputation

Vavilis et al. identifient, pour faciliter le fonctionnement des systèmes de réputation, 13 pré-requis qu'ils regroupent en 3 catégories (relatives aux 3 dimensions).

Dans la **dimension de formulation**, les auteurs distinguent 7 pré-requis qui sont relatifs aux informations et méthodes d'agrégation utilisées pour mesurer la valeur de la réputation :

- Les **évaluations** devraient permettre d'identifier le comportement des utilisateurs : cette première assertion souligne le fait que les évaluations ou notations attribuées à un agent doivent permettre de classifier et distinguer les comportements des utilisateurs d'un système. Ce premier élément est intrinsèquement lié au second ;
- Le système de réputation devrait pouvoir **discriminer les notations «incorrectes»** : un des éléments clés pris ici en compte est la qualité et la fiabilité des informations fournies par un utilisateur. Certains utilisateurs peuvent, dans l'optique de "ternir" la réputation d'un agent intentionnellement ou simplement par inadvertance, lui attribuer

[246] Vavilis, Sokratis & Petković, Milan & Zannone, Nicola. (2014). A Reference Model for Reputation Systems. Decision Support Systems. 61. 10.1016/j.dss.2014.02.002.

des retours négatifs. C'est donc le rôle du système, de par les informations préalables fournies par l'utilisateur, de juger de sa fiabilité et de censurer les notations biaisées ou fausses ;

- Une **entité ne devrait pas pouvoir se noter elle-même** : la notion de fiabilité des notations est ici aussi prise en compte et le système de réputation doit jouer un rôle de gouvernance, de supervision des agents évalués en leur interdisant formellement l'auto-évaluation ;
- L'**agrégation des notations devrait être significative** : il s'agit ici du type de données à collecter dans la mesure de la réputation. Plusieurs auteurs ont tenté dans la littérature d'identifier les différentes sources d'informations issues des interactions entre les différents agents ;
- La réputation doit être évaluée à partir d'une **quantité d'informations suffisante** : il est ici question de la quantité de données utilisées dans la mesure de la réputation. Celle-ci se doivent d'être exhaustives afin de garantir la véracité et la fiabilité des notations, ce qui n'est pas toujours le cas dans les systèmes de réputation, les évaluations émanant des acteurs et non du système, et étant souvent sporadiques ;
- Le système de réputation devrait **distinguer les informations de réputation par l'interaction qui les agrège** : les auteurs introduisent ici la notion de contexte. En effet, toute évaluation est relative à une interaction ou un objet spécifique et ne saurait être généralisée à tous les contextes ou cas, ce qui justifie souvent les écarts de *ranking* entre plusieurs utilisateurs ;
- La **réputation devrait refléter et enregistrer l'évolution du comportement des utilisateurs** : est ici mise en exergue la notion de temporalité, le comportement des utilisateurs étant changeant à travers le temps. La réputation d'un agent étant basée sur l'ensemble des interactions passées, le système devrait conserver l'évolution du comportement des utilisateurs afin de garantir leur fiabilité.

Afin de compléter ces premiers éléments, Vavilis et al. identifient 5 autres pré-requis inhérents aux **dimensions de calcul et de distribution**. Ils sont relatifs au traitement équitable des nouveaux utilisateurs et notamment les limites dans la sélection de la valeur de réputation accordée aux nouveaux utilisateurs.

- Les **utilisateurs, sous prétexte de leur nouveauté, ne devraient pas être avantagés** : lorsque de nouveaux utilisateurs rejoignent le système, leur comportement est méconnu des autres, et en règle générale, les systèmes de réputation attribuent une valeur de réputation par défaut aux nouveaux utilisateurs. L'objectif ici est, d'éviter que des agents malveillants ou mal intentionnés ne tirent profit de ce statut de nouveau ;
- Les **utilisateurs, sous prétexte de leur statut de "nouveau" ne devraient pas être pénalisés** : l'idée est la même que dans le critère précédent avec ici la visée que la valeur de la réputation par défaut attribuée ne devrait pas pénaliser les nouveaux agents pour leur statut. Si les nouveaux arrivants sont traités comme des utilisateurs avec une mauvaise réputation, ils peuvent ne jamais être choisis par d'autres utilisateurs et ne pourront donc peut-être pas construire leur réputation ;

Et enfin, les trois derniers critères préalables au bon fonctionnement d'un système de réputation sont relatifs à l'intégrité des données collectées (les valeurs de réputations) :

- Les utilisateurs ne devraient pas pouvoir modifier directement les notations afin de garantir leur authenticité (entre la transmission, le stockage et le traitement des données) ;
- Les utilisateurs ne devraient pas pouvoir modifier directement les valeurs de la réputation ;
- Il est interdit aux utilisateurs de calculer leur propre réputation afin d'éviter que les données ne soient manipulées, falsifiées ou biaisées. Il incombe donc au système sur la base des informations fournies par les utilisateurs, d'attribuer les valeurs de la réputation.

Des fonctionnalités utilisées dans la mesure de la réputation

Outre les trois dimensions et les pré-requis sus-cités et relatifs au bon fonctionnement d'un système de réputation, Vavilis et al. identifient **plusieurs attributs** relatifs autant à la nature des informations (fiables) à utiliser dans la mesure de la réputation qu'aux propriétés des valeurs de réputation et des méthodes d'agrégation de ces données. Ils distinguent donc plusieurs caractéristiques essentiellement liées à la dimension " formulation " :

- La **confiance / méfiance** : cette première fonctionnalité la " confiance et la méfiance" représente les différentes perceptions des agents qui, à partir de leurs interactions, doivent être en mesure de représenter, et ce avec précision, le comportement des agents dans la mesure de la réputation : bonne, mauvaise et neutre ;
- Les **valeurs de réputation** : les valeurs attribuées à la réputation doivent l'être de manière singulière et décorrélée des valeurs des autres utilisateurs afin d'éviter les biais ;
- L'**origine / Cible** : il s'agit d'identifier l'origine des sources de données et les entités productrices de ces informations ;
- La **certitude / l'incertitude** : est relative au degré de (in) certitude accordé aux informations fournies par un agent ;
- La **portée de l'interaction** : fait notamment référence aux différents agents susceptibles (de par une expertise spécifique) de fournir des éléments de réponse ou avis à un agent sur un sujet donné ;
- Les **substituts** : liés notamment au système de requêtes associé aux concepts de similarité dans les mots clés choisis, aux mots clés de longue traîne par exemple;
- La **chaîne d'évaluation** : est relative à la notion de confiance indirecte et plus particulièrement traite des relais (informations indirectes) utilisés dans la mesure de la réputation ;
- Le **couple *functional et referral trust*** : ces deux notions font notamment référence à l'écart entre habilité à fournir un service et celle à recommander. Le *functional trust* est donc relatif à la capacité d'un agent à fournir un produit ou service et le *referral trust* son aptitude à attribuer des recommandations ;

- La **temporalité** : se dit de la temporalité et en l'occurrence du temps (t) durant lequel s'est déroulée l'interaction, et sa prise en compte dans la mesure de la réputation.

1.8. Les dimensions de classification de Hendrikx et al., 2015

Hendrikx et al.[247] distinguent dans les systèmes de réputation **deux (2) principales dimensions** étayées par **quinze (15) caractéristiques**.

Sur les dimensions, Hendrikx et al. distinguent les systèmes de réputation implicites de ceux dits explicites. Par **système de réputation implicite**, ils entendent un système qui, de manière explicite, ne possède pas de mécanisme de réputation, mais dans lequel les membres peuvent toutefois utiliser les données de réputation dans leur prise de décision. Ces systèmes qui sont très peu structurés ou pas du tout, permettent toutefois de garantir la fiabilité des agents même s'ils sont parfois tentés de "trahir" les autres. Au compte de ces derniers, les auteurs citent notamment les réseaux socionumériques comme Facebook ou Linkedin ou encore Google. Les **systèmes de réputation explicites** eux, en outre, utilisent des mécanismes permettant de mesurer le degré et la nature de la confiance attribuée dans les interactions des agents. Ce type de système repose sur une pluralité et diversité de membres, et implique de fréquentes interactions entre ces derniers. Tandis que les systèmes explicites ont comme critères ceux allant de l'historique à l'objet de mesure, ceux implicites sont relatifs aux données allant de la question de l'autorité à la temporalité.

Les auteurs proposent par la suite, une analyse des différentes caractéristiques inhérentes à ces deux dimensions. Au compte de ces dernières (les caractéristiques), on retrouve :

- L'**historique** : la notion "d'historique" renvoie à l'ensemble des informations relatives aux transactions et aux résultats des transactions d'un utilisateur, et sert donc de base afin de prédire l'issue des transactions futures ou actuelles. Cet ensemble d'informations (*feedback*) est souvent représenté sous forme de notation. Les auteurs distinguent deux (2) types d'historiques : l'historique personnel (utilise les

[247] Ferry Hendrikx, Kris Bubendorfer, Ryan Chard. Reputation systems: A survey and taxonomy. J. Parallel Distrib. Comput. 75 (2015) 184–197, http://dx.doi.org/10.1016/j.jpdc.2014.08.004.

informations directes ou observées) et celui global (informations partagées par les autres membres) ;

- Le **contexte** : les auteurs identifient plusieurs types de contextes simple ou multiple ;
- Le **type d'informations collectées** : informations directes ; indirectes et dérivées ;
- La **valeur de la réputation** peut revêtir plusieurs formes : binaire ; discrète ou continue ;
- La **méthode de calcul** : décrit la manière dont le score de réputation est mesuré (la somme ou moyenne, l'approche bayésienne, la logique floue, la SL ou la méthode basée sur les flux) ;
- L'**objet de mesure** : se dit de l'entité sur laquelle porte l'évaluation dans un système de réputation. Il peut s'agir d'un individu ou d'une entité ;
- La **question de l'autorité** : il s'agit de la présence ou non d'une autorité dans le système afin de fournir les informations de réputation aux agents concernés. Certains systèmes nécessitent une présence totale de l'autorité ; d'autres nécessitent une présence partielle et certains n'ont besoin d'aucune autorité afin de rendre disponibles les valeurs de la réputation ;
- Le **type de système** : les systèmes de réputation peuvent être centralisés ou décentralisés ;
- L'**accessibilité des données** : l'interopérabilité est relative à l'accessibilité des agents aux données de réputation dans un système. Les données de réputation peuvent être ouvertes (accès et utilisation) ou fermées (gérée par une entité auprès de laquelle il faut en faire la demande) ;
- Les **mécanismes de contrôle** : il s'agit des techniques de gestion et de contrôle mises en place afin d'assurer la fiabilité des données de réputation. Certains modèles ont donc intégré des règles ou des techniques de motivation / punition ;
- L'**évaluation** peut prendre deux formes selon les modèles partielle (sur la base d'un échantillon des informations disponibles) ou holistique (prenant en compte toutes les interactions d'un agent afin de lui attribuer une valeur de la réputation) ;
- Le **filtrage des données** : ce critère questionne la présence de mécanismes permettant de filtrer, modérer les contenus. Certains systèmes notamment ouverts ne limitent ni

ne filtrent les données collectées et d'autres, proposent un sous-ensemble de critères de publication des informations ;

- La **temporalité** : ce critère part du principe, selon les auteurs, que plus les données de réputation sont anciennes, moins elles sont fiables, et qu'il faudrait donc prioriser les évaluations récentes. Certains modèles ne prennent donc pas en compte la question de temporalité et les informations sur la réputation sont conservées indéfiniment. D'autres les prennent en compte et réduisent la confiance et la granularité des informations de réputation plus anciennes à mesure que le temps passe. Enfin, quelques-uns offrent la possibilité de supprimer les données très anciennes de réputation.

1.9. Les dimensions de classification de Granatyr et al. 2015

Une autre approche qu'on retrouve dans la littérature est celle proposée par Granatyr et al. (2015)[248] à partir des travaux de plusieurs auteurs (Sabater et Sierra, 2005 ; Grandison et Sloman, 2000 ; Ramchurn at al., 2004b ; Lu et al., 2009 ; Pinyol et Sabater, 2013) et qui prend en compte **onze (11)** principales dimensions.

Sur les p**aradigmes utilisés**, les auteurs en distinguent trois à partir des travaux de : l'approche cognitive, celle numérique avec notamment la théorie des jeux et les modèles hybrides (Sabater et Sierra, 2005).

- L'approche **cognitive** est relative aux croyances et états mentaux des agents dans un système. La mesure de la confiance s'effectue à partir du degré de ces croyances et des conséquences mentales de la décision de faire confiance ou non à un autre agent. Dans les modèles théoriques qui sont qualifiés comme tels pour Sabater et Sierra, les machines sont construites à l'image des hommes, utilisant des architectures cognitives humaines qu'ils nomment BDI (*belief, desire and intention*) dans le calcul de la confiance ou de la réputation. Au compte des premiers modèles à prendre en compte

[248] Granatyr, Jones & Botelho, Vanderson & Lessing, Otto & Scalabrin, Edson & Barthès, Jean-Paul & Enembreck, Fabrício. (2015). Trust and Reputation Models for Multi-Agent Systems. ACM Computing Surveys. 48. 10.1145/2816826. DOI:10.1145/2816826

cet aspect dans leur conception, on peut citer ceux proposés par Castelfranchi et Falcone (2001), Carter et Ghorbani (2003) et Zhang & al. (2007) ;

- Le deuxième paradigme mis en évidence part du constat que plusieurs modèles utilisent dans leur conception, des **théories probabilistes et numériques** d'agrégation des données à partir des interactions passées entre les agents d'un système. La plupart des modèles qu'on retrouve dans ce paradigme utilisent comme le soulignent Granatyr et al. (2015) des méthodes statistiques telles les probabilités Bayésiennes (Regan et al. 2006) ou les fonctions de Dempster-Shafer (Yu et Singh 2003, Liu et al. 2012) par exemple. On note également dans ces modèles, l'utilisation des techniques de machine Learning (Tran et Cohen, 2004) ;

- Les approches **hybrides** comme l'indique leur nom, utilisent à la fois les deux types de paradigmes cités plus haut. À titre d'exemples, nous pouvons citer des modèles un peu plus récents tels que ceux de Matt et al. (2010), Koster et al. (2011), Parsons et al. (2011).

Sur la question des **sources d'information**, ils distinguent : les expériences directes, l'observation directe, les témoignages, les informations sociologiques, les données de réputation certifiée et enfin les règles.
Sur la prise en compte des **informations frauduleuses** (*cheating assumptions* au sens de Sabater et Sierra, 2005), ils définissent trois niveaux : le niveau 0, le niveau 1 et le niveau 2.

Sur la notion de **sémantique** (*trust semantics*), les auteurs partent du postulat que la plupart des modèles évaluent la confiance et la réputation en prenant en compte une seule valeur, qui indique si l'agent est fiable ou non, s'il a une bonne réputation ou non. Par ailleurs, la sémantique ou la multidimensionnalité vise à transformer cette valeur unique en une moyenne de plusieurs aspects, en la convertissant en une mesure composite. C'est notamment le cas des modèles qui utilisent des ontologies pour combiner différents aspects en une seule valeur, les modèles basés sur les priorités (coût, rapidité, fiabilité, etc.) qui mesurent la confiance au moyen d'ensembles de priorités ajustables en fonction des préférences de

l'agent, etc. Les auteurs soulignent que pour l'architecture des systèmes multi-agents (MAS), il est important de représenter clairement le mode d'affichage des évaluations, afin que tous les agents puissent interagir de manière cohérente en utilisant le même standard sémantique. Toute distorsion sémantique peut compromettre la fiabilité des résultats des calculs de confiance, et pour résoudre ce problème, l'utilisation des ontologies fournies par les MAS peut constituer une alternative à tous les agents de l'environnement évaluant correctement leurs partenaires.

Sur la question des **préférences** ou filtres (*trust preferences*) il s'agit ici de la possibilité pour les agents de définir des pondérations pour chaque attribut et de les hiérarchiser en fonction des besoins. Il est question comme dans le point précédent de définir un ensemble de priorités (prix du produit, délai de livraison, relation client, etc.) à prendre en compte. Certains modèles proposent donc d'attribuer un système de poids (ou valeur) aux informations collectées en fonction de leur source ; de personnaliser le comportement selon que le risque d'interaction est élevé ou non ; etc. Les auteurs soulignent que les préférences sont souvent une sorte d'informations gérées en interne par l'agent lui-même. Il est donc important que les modèles MAS incluent des mécanismes facilitant les priorités du système et les préférences des agents. Pour résoudre ce problème, les ontologies pourraient être utilisées pour représenter les préférences des agents, il serait donc facile de partager ces concepts avec le reste de la communauté.

À partir de cet état de l'art, et comme on a pu le constater, les modèles conceptuels de réputation et de confiance se basent sur une myriade d'attributs (concepts et critères) pouvant se regrouper (plusieurs auteurs ayant développé les mêmes typologies de classification). Le tableau ci-dessous vise à proposer une synthèse de l'ensemble des approches de classification précédemment étudiées :

AUTEURS	DIMENSIONS (CRITÈRES)
Sabater & Sierra (2005)	Conceptual model, Information sources, Visibility types, Model's granularity, Agent behavior assumptions, Type of exchanged information, Trust / reputation reliability measure
Koutrouli & Tsalgatidou (2006)	Information gathering, Reputation estimation, Reputation representation
Marti & Garcia Molina (2006)	Information gathering, Reputation scoring, Ranking and response
Wang & Vassileva (2007)	Centralized vs decentralized, Person / agent versus ressource, Personalized versus global
Hoffman & al. (2009)	Formulation, Calculation and Dissemination
Yuan & al. (2012)	Dissemination, Calculation and Evolution
Vavilis & al. (2014)	Formulation, Calculation and Dissemination (*Trust/distrust, absolute reputation values, origin/ target, (un)certainty, interaction scope, scope similarity, trust transitivity, functional vs referral, interaction context, timestamp)*
Hendrikx & al. (2015)	Implicit, Explicit (*History, context, collection, representation, aggregation, entities, presence, governance, fabric, interoperability, control, evaluation, data filtering, data aging)*
Granatyr & al. (2015)	Paradigm, Information sources, Cheatings assumptions, Trust semantics, Trust preferences, Delegation trust, Risk measure, Incentive feedback, Initial trust, Open environment, Hard security

Tableau n°2 : Synthèse des dimensions de classification des modèles théoriques

Notre travail de recherche visant à déterminer un lien entre les méthodes d'évaluation des professionnels de la réputation et les modèles théoriques, nous proposons d'identifier dans une première approche, les catégories de classification des modèles théoriques sur lesquelles nous nous appuierons. Afin de définir ces grands axes de comparaison, nous proposons d'abord de présenter l'ensemble des concepts (les 8 présentés dans le tableau ci-dessous) qui seront utilisés dans notre tableau de classification des modèles théoriques (que vous retrouverez en annexe 1).

Indice	**Caractéristique**	**Définition**
D	Date	Il s'agit de la date de publication du modèle
Cit	Citation	Il est question du nombre de citations sur Google Scholar relatif au modèle
Cont	Contexte	Se dit du contexte d'utilisation du modèle qui peut être multiple ou relatif à un contexte singulier
Dbi	Détection de biais	Il est ici question de l'implémentation de mécanismes permettant de détecter les biais dans les données
Mrep	Métriques de la réputation	Est relatif à l'emploi d'une ou de plusieurs métriques de la réputation
Srd	Source des données	Qualifie les typologies de données collectées et la diversité de sources utilisées
Typ	Typologie	Il est ici question du type de modèle qui peut soit être uniquement de mesure de la réputation ou hybride
Mma	Modèle mathématique	Il s'agit de distinguer si la méthode de calcul relève des probabilités, de la *fuzzy logic*, du *flow based* ou de la s*ubjective logic*

Tableau n°3 : Concepts utilisés dans notre tableau de classification des modèles

Nous proposons en annexe n°1 un tableau de classification de l'ensemble des modèles théoriques issus de notre analyse de la littérature scientifique. À partir des critères de classification définis dans le tableau ci-dessus, nous identifierons (nous y reviendrons dans le point 3 de ce chapitre) quatre (4) principales dimensions de classification des modèles théoriques, dimensions qui serviront à comparer les méthodes professionnelles et les approches académiques (nous y reviendrons dans le chapitre 7).

Au vu des premiers constats faits sur et issus des tentatives de classification des chercheurs, et avant de présenter notre approche de classification, il convient de réaliser un travail d'analyse sur le processus de mesure théorique de la réputation tant sur l'approche structurelle utilisée qu'à partir de la méthodologie de mesure.

2. La mesure de la réputation en théorie : approche contextuelle et méthodologique

Le processus d'évaluation de la réputation d'une entité suppose, à partir des diverses approches issues de la littérature que nous avons analysées, plusieurs phases allant de la collecte de l'ensemble du corpus à l'attribution d'une valeur de la réputation. Les modèles conceptuels, nous l'avons vu, pouvant être construits en utilisant une approche centralisée ou distribuée, la mesure suppose ici une double approche structurelle et méthodologique. Nous présenterons donc d'abord dans un premier temps l'approche contextuelle (les contextes d'évaluation) de mesure utilisée, puis, dans un second temps, la méthodologie d'évaluation (de la collecte à l'analyse des données de réputation) qui peut différer selon les contextes au sein desquels s'effectue la mesure.

2.1. Les contextes d'utilisation des modèles

Notre analyse de la littérature nous a permis d'identifier deux principales approches structurelles des modèles de réputation : les modèles adaptés aux structures centralisées et qui sont construits autour de certaines caractéristiques bien distinctes, et ceux utilisés uniquement dans ce qu'on qualifiera de systèmes ouverts ou décentralisés (*decentralized*).
L'approche centralisée dans la mesure de la réputation en ligne fait référence aux systèmes de réputation construits autour d'une autorité centrale avec une utilisation quasi simple, offrant une marge de manœuvre (de manipulation des données) très faible aux agents malicieux. Ce qui suppose une confiance aveugle en l'autorité centrale qui, en retour, se doit d'être toujours disponible et impartiale tant dans la collecte que la mise à disposition des données de réputation.

L'approche centralisée suppose l'existence de plusieurs groupes essentiellement formés par une coalition d'entités ou d'individus reliés entre eux par un intérêt commun ou un sujet spécifique, ce qui pourrait donc garantir que les informations collectées par l'autorité centrale proviennent d'entités avérées. Par conséquent, dans l'optique d'interagir avec un agent dont on ignore la réputation et plutôt que de rechercher manuellement des données relatives à ce dernier, ces *context groups* permettent d'obtenir des informations plus ou moins fiables (ce qui questionne l'impartialité de l'autorité centrale) auprès d'individus ayant déjà interagi (propos à nuancer également eu égard aux problématiques de véracité des avis collectés) avec l'agent en question. Le calcul de la réputation, qui est lié au contexte spécifique autour duquel l'ensemble a été formé, est donc effectué par l'autorité centrale (une tierce partie) à partir des avis reçus des diverses entités. Les modèles construits autour d'une approche centralisée de la mesure de la réputation présentent certaines spécificités pouvant être perçues comme des avantages par certains auteurs (Zacharia & Maes, 2000 ; Bertocco et al., 2008, etc.). À titre d'exemple, la méthode de collecte des données se voit simplifiée dans la mesure où les agents n'utilisent qu'une seule source centrale où sont stockées les données de réputation et la mise à jour de celles-ci est automatique eu égard aux nouveaux avis qui sont fournis par les nouveaux membres du groupe à la demande de l'autorité centrale. Une autre des caractéristiques essentielles à cette typologie de modèles est la pertinence des avis,

pertinence qui suppose une authentification des agents qui proposent une évaluation sur un objet (organisation, produit ou service).

Ces assertions se voient en certains points remises en cause. D'abord, sur la question de l'exhaustivité des données, n'utilisant qu'une seule source, on pourrait se questionner sur la représentativité des valeurs de la réputation sur une entité donnée en prenant en compte le fait que tous les agents qui ont interagi avec l'entité en question ne répondent pas toujours à la demande de l'autorité en charge de la collecte des données, ce qui peut présenter un biais. Un autre élément, et non des moindres, est la question de la pertinence des données. Ce postulat se fonde sur la capacité à identifier les auteurs à l'origine des avis, auteurs qui n'ont aucun intérêt à fournir des données fausses ou biaisées. Dans la mesure où cette première hypothèse serait accréditée, on en reviendrait encore à la crédibilité accordée aux tierces parties (il est ici question de l'autorité centrale) en charge de la collecte des données et à leur impartialité dans la gestion de ces avis. Plusieurs modèles (12) se retrouvent dans cette classification dont essentiellement les modèles adaptés aux plateformes e-commerce / places de marché sont les plus représentatifs (Josang & Ismail, 2002 ; Vogiatzis & al., 2010 ; Fang & al., 2012 ; Ransi & Kobti, 2012; Abdel-Hafez & al., 2015 ; Acampora & al., 2016). On retrouve également les modèles Bertocco & Ferrari (2008) ; Urzica & al., (2011) ; Carter & al, (2002) ; Li & Gui (2009) et dans le domaine de l'*Internet Of Things* (Wang & al, 2016 ; Li & al, 2016).

À l'inverse, l'approche décentralisée ou distribuée qu'on retrouve dans la plupart des systèmes de réputation, se révèle par ailleurs être dans une certaine mesure plus complexe. En effet, elle ne permet pas l'authentification des données et des entités et fait fi d'un acteur central en charge de la collecte des données. De plus, elle utilise dans la mesure de la réputation d'un objet, les données de tous les agents, ce qui fait par conséquent qu'on parlera de réputation d'un agent *A* par rapport à un agent *B*. Le calcul de la réputation est effectué par chaque agent en utilisant à la fois ses propres informations et les données reçues des autres entités ayant interagi avec l'acteur dont on mesure la réputation. L'une des problématiques principales des modèles construits autour de la structure décentralisée réside dans l'authenticité des avis collectés, qui étant laissés par n'importe quel agent du système (incluant les concurrents, les membres de l'entité de mesure elle-même et les tiers n'ayant

jamais interagi avec celle-ci) peuvent de par leur nature être biaisés et non représentatifs de l'entité dont on cherche à évaluer la réputation.

Au compte des modèles qui utilisent cette approche, on retrouve entre autres ceux adaptés aux moteurs de recherche (Page & al., 1998), au partage de documents (Damiani & al, 2002; Wang & vassileva, 2003, Gupta & somani, 2004; Chen & al, 2007; Zhou & al, 2007; Zhao & li, 2008; Tian & yang, 2011, Fang & al., 2011 ; Skoric & al., 2016 ; Lu & al., 2016), à la supply chain (Zhang & al., 2015), au bookmarking et au microblogging (Portmann & al, 2015), aux plateformes de Q&A (Chen & al, 2006), aux plateformes conversationnelles (Serrano & al., 2012), aux wiki (Jung, 2012), aux sites d'avis (Li & al., 2013), etc[249].

À titre d'exemple, EIGENTRUST (Kamvar & al., 2003)[250], système de réputation qui est utilisé pour le partage de fichiers dans les réseaux peer-to-peer utilise une approche décentralisée[251]. En effet, il est construit sur le principe du PAGERANK, et vise à isoler les agents malveillants en permettant aux entités de décider à quels autres agents ils feront confiance lorsqu'il s'agit de télécharger des fichiers dans les réseaux de partage de pair à pair (*Peer to Peer*). Il s'appuie notamment sur un ensemble global / partagé d'avis et de recommandations en matière de réputation, et la question de crédibilité est essentielle et utilisée pour calculer chaque valeur de réputation globale.

---

[249] Au cloud computing (Chandran & al., 2016, Lin & al, 2015), à certaines plateformes de e-commerce (Mui & al, 2002; Carbo & al, 2002; Padovan & al, 2002; Tran & Cohen, 2004, Song & al, 2005; Regan & cohen, 2005; Regan & al., 2006; Zhang & al, 2006, Gutowska, 2008; Madj & Balakrishnan, 2015), au MAS (Sabater & sierra, 2002; Sen & sajja, 2002; Jurca & faltings, 2003; Derbas & al, 2004; Huynh & al, 2004; Teacy & al, 2006; Sabater & al, 2006; Tajeddine & al, 2007; Bertocco & Ferrari, 2008; Tian & al, 2009; Nguyen & al, 2010, Wang & al, 2011; DART, 2012; Pinyol & al., 2012, Khosravifar & al, 2012), aux p2p (Yu & singh, 2002; Kamvar & al, 2003; Xiong & Liu, 2004; Tian & al, 2006; Osman & al., 2015 ;), *Wireless sensor network* (Fang & al., 2015), *Vanet* (Marmol & Perez, 2014), Voip networks (Wang & al., 2013), *Grid computing* (Kussul & al., 2013), *Manet* (Michiardi & al, 2002; Buchegger, 2003; Qureshi & al, 2013, Qureshi & al, 2011, Almulla & al., 2011, Liu & Issarny, 2009), *IOT* (Tuna & al., 2013)

[250] S. D. Kamvar, M. T. Schlosser and H. Garcia-Molina. The EigenTrust Algorithm for Reputation Management in P2P Networks. Technical report, Stanford University, *WWW 2003*, May 20–24, 2003, Budapest, Hungary. Accessible en ligne http://ilpubs.stanford.edu:8090/562/1/2002-56.pdf

[251] Qu'on retrouvera dans notre tableau de classification comme modèle sophistiqué car il intègre des modules évolutifs non pris en compte par Google notamment

2.2. La phase de collecte : des sources de données à leur catégorisation

Si les modèles construits sur une structure centralisée utilisent un acteur central qui à l'issue de chaque interaction va collecter les données de terrain auprès des agents ayant interagi avec l'entité (et de fait des avis réels dans la plupart des cas[252]), l'approche décentralisée elle, ne permet pas d'identifier comme nous l'avons vu, les agents et consiste à collecter l'ensemble du corpus accessible et se rapportant à l'entité. Outre cette distinction, le processus s'appuie de manière générale sur le même principe avec quelques écarts (que nous expliciterons au fur et à mesure) et se décline sous plusieurs étapes.

2.2.a. Le sourcing : des sources et typologies de données collectées

Plusieurs sources de données sont utilisées dans l'approche théorique de mesure de la réputation dans un système[253] et on peut les regrouper selon deux typologies : les sources directes et celles indirectes (de tierces parties, témoignages) et ce qu'on qualifiera de données sociales (utilisées par une minorité de modèles).

L'un des premiers modèles les plus complets à utiliser une combinaison des trois sources est le modèle modulaire de mesure de la réputation appelé REGRET de J. Sabater et C. Sierra (2001, 2002)[254] et qui prend en compte trois (3) dimensions d'informations : la dimension individuelle, la dimension sociale et la dimension ontologique. La dimension **individuelle** de la réputation est relative à la confiance directe issue des résultats des interactions directes entre deux agents (en clair, on parlera de la réputation d'un agent *A* selon un agent *B*).

Cette première dimension se voit complétée par la dimension **sociale**, et elle prend en compte les expériences des autres membres partant du postulat que les individus appartiennent à une

---

[252] Certains modèles ont parfois recours et ce lorsque les avis sont insuffisants, à l'utilisation de données tierces

[253] Par système, nous entendons ici l'espace de production au sens large des données de réputation et les diverses plateformes numériques utilisées (le e-commerce, les médias sociaux, les moteurs de recherche, etc.)

[254] Sabater, Jordi, et Carles Sierra. « REGRET: reputation in gregarious societies ». In *Agents*, 1:194-95, 2001. ; Sabater, Jordi, et Carles Sierra. « Reputation and social network analysis in multi-agent systems ». In *Proceedings of the first international joint conference on Autonomous agents and multiagent systems: Part 1*, 475-82. ACM, 2002.

société et la réputation ne saurait être individuelle, mais plutôt de groupe (collective). Trois types de sources d'informations sont utilisées afin de mesurer la dimension sociale de la réputation. La réputation de témoin qui est calculée à partir des témoignages de l'ensemble des agents ayant interagi avec l'entité, la réputation du voisinage basée sur les relations sociales entre les agents et mesure la réputation de l'agent qui évalue, et enfin la réputation du système calculée à partir de la connaissance des rôles et propriétés générales de l'agent cible.

La dernière dimension, celle **ontologique**, considère la réputation d'un individu comme un concept à multiples facettes s'appuyant sur plusieurs attributs. Pour l'illustrer, les auteurs prennent l'exemple d'une compagnie aérienne de voyage. Pour eux, être réputé comme une bonne compagnie aérienne signifie avoir de bons avions, ne jamais perdre de bagages et servir de la bonne nourriture. En parallèle, avoir de bons avions, sous-tend d'avoir un bon service de maintenance et de renouveler fréquemment la flotte. Mettre en place cette dimension revient donc à prendre en compte et combiner les différents types de réputation pris individuellement pour en obtenir un seul. Processus quelque peu complexe, car comme le soulignent Sabater et Sierra, "chaque individu a généralement une structure ontologique différente pour combiner les réputations, et une manière différente de mesurer l'importance des réputations lorsqu'elles sont combinées".

Dans la même optique, T. D. Huynh, N. R. Jennings et N. Shadbolt (2004, 2006)[255] proposent également une approche modulaire appelée FIRE, qui intègre dans la mesure de la confiance et de la réputation plusieurs sources d'informations selon leur disponibilité. La première source ici prise en compte est celle directe, c'est-à-dire issue du résultat des interactions directes passées entre un agent *A* et un autre *B*. Aux données provenant d'interactions directes entre deux entités, s'ajoutent celles sociales et celles provenant des témoins, construites à partir des témoignages d'autres acteurs ayant interagi avec la même entité (qu'on retrouve dans le modèle REGRET). Et enfin, le dernier type, ce que les auteurs ont appelé la réputation certifiée qui est basée sur les évaluations des références des "*third parties*" (tierces parties) fournies à la personne qui évalue par l'agent cible lui-même (c'est notamment le cas

---

[255] Huynh, Trung Dong, Nicholas R. Jennings, et Nigel R. Shadbolt. « Certified Reputation: How an Agent Can Trust a Stranger ». In *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems - AAMAS '06*, 1217. Hakodate, Japan: ACM Press, 2006. https://doi.org/10.1145/1160633.1160854.

des recommandations ou références produites par l'entité évaluée elle-même). Il est important de souligner à ce stade, comme le notent les auteurs, que le choix des sources d'information dépend de l'objectif spécifique et de la problématique choisis. En effet, pour les auteurs, toutes les typologies d'interactions ou d'informations ne peuvent pas être utilisées dans tous les contextes, et parfois, chaque type d'information peut être utilisé pour mesurer un aspect de la réputation. Pour l'illustrer, Huynh et al. prennent l'exemple des systèmes dynamiques ouverts où la confiance basée sur le rôle d'un agent au sein d'une communauté ne sera pas la source à privilégier étant donné que dans les systèmes dynamiques les rôles sont modifiables et pas très bien définis. Seront plus à prendre en considération dans ce type de systèmes, les informations provenant des interactions directes et les témoignages, l'utilisation de la réputation certifiée nécessitant l'existence d'un mécanisme de sécurité approprié.

À l'image de ces deux premiers, Sen et Sajja (2002) quant à eux dans leur modèle de réputation, prennent également en compte trois principales sources dans la mesure de la réputation en ligne provenant de deux (2) types d'interactions: l'interaction directe et l'interaction observée, sources primaires auxquelles peuvent être corrélés les témoignages (*witness informations*). Pour eux, il existe une pléthore d'agents qui fournissent le même service dont a besoin un utilisateur. Afin de choisir le partenaire le mieux évalué, l'utilisateur peut soit interagir avec les autres agents qui connaissent les performances des services des différents agents évalués et peuvent donc recommander les meilleurs agents, soit observer leurs interactions. Pour Sen et Sajja (à l'image de la perception partagée par la plupart des chercheurs), seule l'interaction directe donne une perception exacte de la réalité, les observations pouvant différer de la réalité.

Il est important de noter qu'à l'exception de certains modèles qui utilisent les trois typologies de sources, certains modèles s'articulent autour des informations issues d'interactions directes qu'ils qualifient comme étant les plus fiables, car comme ils le soulignent, seules ces dernières donnent une perception exacte de la performance des agents. Ce type de données, qui est la résultante d'une transaction entre deux agents *A & B* et où à l'issue de laquelle, l'agent A évalue B à partir de plusieurs critères, est le plus souvent utilisé lors de transactions (marchandes) avec la possibilité de vérifier l'identité de l'agent qui évalue (d'où la fiabilité

accordée à celles-ci). C'est notamment le cas des modèles (Tran & Cohen, 2004 ; Zhang & al., 2006 ; Bertocco & Ferrari, 2008 ; Nguyen & al., 2010 ; Qureshi & al., 2011, etc.).

Aux données issues d'interactions directes dans ce contexte, sont corrélées celles issues d'observation directe. Cette typologie d'informations est essentiellement représentative des modèles construits autour d'une structure centralisée. Le principe de l'observation directe est simple : lors d'une interaction entre deux agents au sein d'un système (au sens d'environnement numérique), une tierce partie (l'autorité centrale dans les modèles centralisés par exemple) peut, durant le processus d'échange, observer l'interaction et fournir à la fin des informations de réputation sur l'entité dont on cherche à évaluer la réputation. Prenons par exemple un internaute qui achète un produit ou service sur un site e-commerce auprès d'un autre agent. Si pendant le processus de l'achat à la réception, l'acheteur s'est entretenu avec le vendeur sur n'importe quelle problématique (non respect des délais de livraison, produit abîmé, retour, etc.), la tierce partie en charge de la gestion du système, peut dans l'attribution d'un indice de réputation au vendeur, utiliser les données issues de son observation directe de la transaction. Dans la littérature scientifique, plusieurs modèles utilisent cette typologie d'informations. C'est notamment le cas des modèles de Sen & sajja, 2002 ; Fullan & al., 2005 ; Almulla & al., 2011 ; Fan & al., 2011 ; Serrano & al., 2012.

Cependant, et comme déjà mentionné, la mesure de la réputation d'un agent suppose la prise en compte d'une myriade d'avis sur ce dernier. Les informations issues de sources directes étant insuffisantes, le recours à d'autres sources dites « indirectes » s'avère essentiel. Plusieurs modèles s'appuient de fait sur d'autres sources qualifiées de *witness informations* dans la littérature c'est-à-dire de témoignages ou recommandations de tiers. Ce type de données très couramment utilisées supposent qu'un agent *A* souhaite interagir avec un autre agent *B* et dans la mesure où *A* ne dispose d'aucune information personnelle (*direct information*) sur *B*, il a donc recourt à *C* (à partir d'une requête), qui par ailleurs est supposé [256] avoir déjà interagi (interaction ici perçue au sens large) avec *B* et le recommande.

Cette typologie de données est typique des modèles de mesure de la réputation dans les systèmes décentralisés (environnements ouverts), les avis étant épars et disséminés sur la toile. Il est important de noter que les *witness informations*, au même titre que les autres

[256] Ce qui n'est pas toujours avéré

sources, s'utilise de manière combinée avec les autres dans la plupart des modèles afin d'obtenir une certaine exhaustivité et représentativité. Les seuls systèmes de réputation à ne prendre en compte que les secondes typologies sont les systèmes de recommandations (*recommender systems*) ou plus simplement les sites d'avis qui agrègent (de manière centralisée donc) les opinions d'agents ayant déjà interagi avec l'acteur dont on mesure la réputation.

2.2.b. La catégorisation : les métriques utilisées dans le classement

Dès lors que les données sont collectées, elles sont filtrées, c'est-à-dire catégorisées selon plusieurs catégories qui reprennent les métriques essentielles de mesure de la réputation : la confiance, l'influence, l'image et l'identité. L'une des métriques (devenue essentielle et indissociable de la mesure de la réputation) qu'on retrouve systématiquement de manière explicite ou implicite dans la totalité des modèles théoriques est la notion de confiance. En effet, les notions de confiance, incertitude et risque sont intrinsèquement liées dans la littérature scientifique, et s'appuient sur le postulat selon lequel « *la confiance n'est requise que dans les situations à risque lorsque les agents ont quelque chose à perdre* » (Gambetta, 2000)[257].

Certains modèles proposent d'implémenter des méthodes permettant de mesurer le degré de confiance d'un agent à partir du degré de risque qu'il peut représenter dans la décision d'interaction avec un autre agent. C'est notamment le cas des modèles hybrides de confiance et de réputation . À côté de ce premier phénomène, la mesure de la réputation s'appuie également sur le concept d'image, qui se voit dans certaines approches substitué à celui d'identité. L'approche par influence ou popularité est également représentée dans la littérature sur les modèles scientifiques, avec notamment le premier modèle de notre classification à l'intégrer dans sa conception : le PAGERANK de Page et Brin (19991). Pour chacune de ces métriques, nous reviendrons plus en substance sur l'approche utilisée dans le processus de mesure.

---

[257] Gambetta, Diego (2000) '*Can We Trust Trust*?', in Gambetta, Diego (ed.) Trust: Making and Breaking Cooperative Relations, electronic edition, Department of Sociology, University of Oxford, chapter 13, pp. 213-237, <http://www.sociology.ox.ac.uk/papers/gambetta213-237.pdf >

Il est important de souligner que la collecte des données peut s'avérer complexe dans certains contextes notamment lorsque les agents n'ont aucun intérêt à coopérer ou à fournir des informations vraies sur leurs interactions, ou encore tout simplement s'ils n'ont pas envie ou ne pensent pas dans l'immédiat à proposer une évaluation. Pour y remédier, certains modèles ont implémenté des mécanismes de motivation / incitations (à laisser des avis) ou *a contrario* visant à punir ceux qui fournissent des informations biaisées ou frauduleuses (*incentives feedback*). Cette propriété vise à mettre en place des stratégies visant d'une part à encourager les avis et d'autre part, à récompenser ou punir les agents selon qu'ils fournissent des recommandations utiles et fiables ou biaisées / fausses. C'est notamment le cas de Jurca et faltings (2003), qui pour faire face à ce problème, ont défini le concept d'agents de courtage, appelés R-agents, qui achètent et vendent des informations de réputation.

En utilisant ce mécanisme, les agents doivent fournir leurs commentaires, car sinon, ils ne seront pas rétribués, et ne pourront de fait pas acheter des informations fiables de réputation (témoignages vérifiés). Epinions[258] disposait également d'un système d'incitation pour les évaluateurs appelé le programme de partage des revenus (*Income Share Program*), par lequel les membres peuvent gagner de l'argent. La part de revenu est automatiquement déterminée en fonction de l'utilisation générale des avis par les consommateurs. Les évaluateurs peuvent potentiellement gagner autant pour aider quelqu'un à prendre une décision d'achat avec un avis positif, que pour aider quelqu'un à éviter un achat avec un avis négatif. Ceci est important afin de ne pas inciter à rédiger des critiques biaisées uniquement dans un but lucratif. Comme indiqué sur les pages FAQ d'Epinions : « Epinions souhaite que vos avis ou commentaires soient les plus honnêtes possibles quitte à ce qu'ils soient négatifs[259]».
Dans la mesure de la réputation, plusieurs *inputs* peuvent, nous l'avons vu, être utilisés allant des retours d'expérience aux témoignages de tierces parties ou encore des données de contexte dans certains cas, et parfois une combinaison de plusieurs d'entre elles. Si la plupart des systèmes de réputation sont construits autour des données provenant de tierces parties et

[258] Qui était un site général d'avis des consommateurs créé en 1999, acquis en 2003 par DealTime, plus tard Shopping.com, et par eBay en 2005. Accessible en ligne https://shopping.com/#top-products-homepage

[259] Traduction de l'anglais " Epinions wants you to be brutally honest in your reviews, even if it means saying negative things"

qualifiées “d’opinion publique”, étant dans une certaine mesure représentatifs d’une forme de “vision commune”, certains systèmes quant à eux, s’appuient sur les informations directes (résultant d'une expérience personnelle), qui sont considérées comme plus fiables. Ces dernières sont cependant moins exhaustives et plus complexes à obtenir. Qu’elles soient issues de sources directes ou non, l’ensemble des informations collectées font, par la suite, l’objet d’une catégorisation. Dès lors qu'elles sont classifiées, les données seront analysées afin de déterminer la réputation d'une entité. Processus d’analyse que nous expliquerons par la suite.

2.3. La phase d’analyse des données : des métriques de réputation à la méthode de calcul utilisée

Dès lors que les données sont collectées, elles sont sauvegardées dans le système de réputation afin de procéder à leur analyse. Si comme nous l’avons vu la réputation est un objet intangible nécessitant une approche pluridimensionnelle, sa mesure s’appuie notamment sur l’évaluation de chacune des métriques la composant, et qui permettra de lui attribuer un résultat. Nous présenterons de fait dans la première section comment ces diverses métriques sont perçues et évaluées, puis par la suite, nous définirons les principales méthodes utilisées dans la mesure.

2.3.a. Des métriques de réputation comme objet de mesure

Dans la mesure de la confiance, deux principaux mécanismes ont été adoptés, celui sécuritaire et de détection de biais dans les données. Sur le premier, qui est d’ordre structurel, il intègre dans les modèles la conception de systèmes informatiques : la « *hard* » et la « *soft* » *security* (Sabater & Sierra, 2005) ; Granatyr et al., 2015, etc.). Protocoles cryptés les plus utilisés dans la plupart des systèmes informatiques pour eux, les modèles qui utilisent des techniques de *hard security* possèdent un système d’identification (avec des identifiants et

justificatifs dans certains cas), d'authentification et de gestion des donnés complexes, garantissant un degré élevé de sécurité des agents d'un système.
Les modèles qui, *a contrario*, intègrent des mécanismes sécuritaires dits « *soft* » dans le calcul de la confiance et de la réputation, se basent sur les observations du comportement des autres agents pour mettre en place une stratégie de protection du système. Ce qui fait de ces approches plus flexibles.

Ces deux typologies se voient complétées par d'autres modèles considérés comme hybrides, qui utilisent des mécanismes plus complets et nécessitent à la fois, afin d'accéder à un système, de posséder une bonne réputation (qui relève de la *soft security*) et de s'identifier, de s'authentifier (*hard security*). Les modèles sophistiqués proposent d'une part des mécanismes de *hard security,* et s'appuient également sur des approches hybrides. Les modèles de Huynh et al. (2006) et Botelho et al. (2009) par exemple utilisent les signatures électroniques pour garantir l'authenticité et la confidentialité des rapports fournis par les agents. La signature électronique qui relève de la *hard security* et la *soft security* est matérialisée par le fait qu'uniquement les personnes choisies par l'agent qui évalue et en qui il a confiance ou qu'il connaît *a minima*, recevront la clé permettant d'accéder aux évaluations. Procédé de signature électronique qui est par ailleurs aujourd'hui repris et adapté à un usage courant, avec des développements sur de nouvelles technologies comme il peut être le cas pour la signature électronique sur Blockchain.[260]

Le deuxième mécanisme utilisé vise la recherche de fiabilité dans les données via l'implémentation des mécanismes algorithmiques (à l'instar de ceux supra cités) visant à réduire les fraudes et biais (*cheating assumptions, cheating behavior ou behavior assumptions*) dans les données collectées qu'on retrouve sous trois principaux niveaux (Sabater et Sierra, 2005) : le niveau 0, le niveau 1 et le niveau 2.

[260] Protocole de signature électronique accessible en ligne sur https://www.blockchain-ez.com/en/la-signature-electronique-sur-blockchain/

- Dans le niveau 0, l'aptitude d'un agent à mentir ou cacher des informations n'est pas prise en compte pour la simple raison que les modèles qui intègrent cette caractéristique, comptent sur le fait qu'il y aurait plus d'agents « honnêtes » qui fourniraient des évaluations fiables, contrecarrant de fait les potentiels effets d'évaluations provenant d'agent malicieux. Les modèles qu'on retrouve dans ce volet (Page & al., 1998 ; Zacharia & Maes, 2000; Mui & al, 2002; Wang & Vassileva, 2003; Wang & al., 2011; Yu & al., 2014, etc.) ne prennent pas en compte les informations « frauduleuses », et aucun mécanisme de filtrage des agents malicieux n'est proposé ;

- Dans le niveau 1, les modèles (Michiardi & Molva, 2002; Chen & al., 2007; Zhang & al, 2008; Urzica & al., 2011, etc.) intègrent la probabilité que des agents puissent cacher ou fournir des informations biaisées, mais rejettent formellement le fait qu'un agent puisse mentir ou attribuer une fausse évaluation volontairement. Dans ce contexte, certaines données, même importantes peuvent être rejetées si elles sont perçues comme biaisées et les informations incomplètes sont ignorées. C'est le cas notamment, à titre d'exemple, du modèle PREP de Haghpanah et desJardins (2012) dans lequel afin d'assurer la fiabilité des données collectées, les auteurs utilisent la méthode bayésienne pour comparer les données qui semblent biaisées aux expériences directes préexistantes ;

- Le dernier niveau, le niveau 2, admet que des agents puissent mentir, et les modèles théoriques construits sur cette base prennent en compte dans leur conception, des mécanismes de prévention des fraudes. Plusieurs modèles proposent soit de comparer les recommandations aux opinions de l'agent en question afin de déceler des similitudes[261], ou encore d'utiliser des clusters pour regrouper les informations.

[261] C'est-à-dire, à partir de l'utilisation d'algorithmes de *ML* et via la collecte de l'ensemble des observations ou recommandations d'un même individu (sur l'ensemble des plateformes numériques) sur plusieurs entités différentes, de déceler d'éventuelles convergences ou divergences dans les propos.

Le modèle LIAR présenté par Muller & Vercouter (2005) s'articule justement autour de la détection de la fraude et de la gestion de la réputation dans les communications. Les auteurs tentent de formaliser les situations "interdites" au regard des informations transmises par les agents et des engagements qu'ils se fixent. Pour ce faire, le modèle définit une procédure visant à détecter les mensonges et biais dans les données. Le modèle de Sen & Sajja (2002), qui prend également en compte les avis des *liars* (menteurs) dont les évaluations sont bien évidemment erronées, propose un mécanisme capable de détecter le pourcentage de personnes fournissant des informations frauduleuses, etc. Le mécanisme utilisé dans ce modèle permet donc de garantir la fiabilité d'un partenaire en calculant le nombre d'agents à interroger, sous-ensemble d'agents choisis de manière aléatoire parmi tous les témoins fiables.

Certains modèles construits autour des mêmes attributs proposent, à partir de fonctions probabilistes, de comparer les informations issues d'interactions passées et récentes d'un agent, afin de déterminer leur fiabilité (TRAVOS de Teacy et al., 2006). Ce premier filtre permet dès lors de classifier les agents à l'origine des informations collectées sous la typologie de "très fiables à fiables"[262], ou sous la catégorie *"pas digne de confiance"* et d'inciter à la méfiance si les données et personnes à l'origine des informations ne sont pas vérifiées.

Outre la question de la confiance, la seconde métrique de réputation évaluée et qui permet de classer les données est relative au concept d'image sur lequel s'appuient plusieurs modèles dont Sabater et al. (2006) avec leur modèle REPAGE, Pinyol et al. (2012) dans leur modèle BDI+REPAGE qui est une déclinaison du précédent, GROUPREP de Tian et al. (2006) et le modèle de Regan & Cohen (2005). Le modèle REPAGE (Sabater-Mir et al. 2006) d'entrée de jeu, dans son architecture, permet aux agents d'un système de faire la distinction entre les questions d'image et de réputation. Même si les deux objets sont des évaluations sociales, l'image relève ici d'une simple croyance évaluative qui indique si un agent cible donné est bon ou mauvais par rapport à un certain contexte (il s'agit ici tout simplement de sa représentation à partir du/des rôles qu'il joue) et la réputation une méta-croyance qui raconte

[262] Si tant est que le mécanisme employé utilise une structure de *hard security* ou hybride, et qu'il a implémenté un mécanisme de détection de biais de type 1 à minima ou 2 pour les systèmes sophistiqués

ce que les autres pensent d'une cible donnée dans un contexte donné. REPAGE organise les évaluations sociales (représentées comme des indicateurs / critères de premier ordre) selon un certain nombre de variations ou échelles, les attributs d'image et de réputation appartenant à l'un des principaux types (image, réputation, part de voix partagée, évaluation partagée) qui contient une évaluation faisant référence à un certain agent dans un rôle spécifique. Le modèle calcule l'image et la réputation grâce à des agrégations à partir de critères élémentaires en utilisant les interactions directes et les communications avec des tiers. Tandis que les expériences directes et les données renseignées sont utilisées dans la mesure de l'image, seules les données indirectes collectées contribuent à l'inférence de la réputation. À titre d'exemple, ils notent qu'un agent *A* peut avoir une très bonne image de l'agent *B* en tant que conducteur de voiture et en même temps reconnaître que *B* a une mauvaise réputation en tant que conducteur de voiture (approche macro de l'ensemble des avis).

Certains modèles ont par ailleurs substitué au concept d'image celui d'identité, partant du postulat que *"la réputation d'un agent est basée sur le degré d'accomplissement des rôles qui lui sont attribués dans la société*" avec une méthode d'évaluation quelque peu différente qui s'appuie sur le contexte de l'évaluation. C'est le cas par exemple des modèles de Carter & al (2002), Jurca & Faltings (2003), Gupta & Somani (2004), Tian & al. (2009) et Wang & al. (2013). Ce dernier, en plus des questions d'identité a intégré la notion d'anonymat et notamment sa gestion. Prenons à titre d'exemple le modèle de Carter & al. qui a été développé au sein d'un système de partage de fichiers (et adaptable à divers environnements de pair à pair). Les auteurs partent du postulat selon lequel chaque individu appartient à une société qu'ils qualifient de "société de partage d'informations"[263] qui édicte elle-même ses propres règles et les rôles à jouer par chaque agent. La mesure de la réputation ici est singulière et donc non universelle, dans cette société de partage d'informations, et pour évaluer la réputation des individus, les auteurs identifient cinq (5) rôles que doivent jouer les agents:

[263] Selon le contexte de leur modèle : partage de fichiers en P2P

- Le premier rôle est celui de *fournisseur d'informations sociales* qui est intrinsèquement lié aux recommandations qu'un agent laisse sur les autres membres de la communauté. L'idée ici est de pouvoir, à partir des recommandations de chaque agent (qui ont chacune un poids), déterminer le degré de satisfaction de ce rôle dans la société dans laquelle il est, en faisant la somme de toutes les mesures dans un intervalle entre 0 et 1 ;

- Autre rôle important, celui d'interactivité qui est corrélé à la fréquence d'utilisation du système par un agent afin d'en garantir son utilité. La mesure de la satisfaction de ce rôle équivaut donc à diviser pour une période donnée, les actions d'un agent par le nombre total d'actions de tous les agents du système ;

- Le rôle de fournisseur de contenus part du postulat que chaque agent d'un système possède une expertise, ou un domaine de compétence particulier qu'il pourrait partager avec les utilisateurs. Satisfaire ce rôle dans une société sous-tend donc de produire des contenus de qualité, selon les centres d'intérêts propres à chaque agent, et que ceux-ci (les contenus) correspondent aux intérêts des autres agents du système ;

- Le rôle administratif de feedback, qui implique que chaque membre évalue la qualité du système dans lequel il évolue à partir de critères prédéfinis : "la facilité d'utilisation, la rapidité, la stabilité et la qualité des informations fournies" ;

- Et enfin, le dernier rôle que doit jouer l'agent et non le moins important, celui de garantie de longévité du système et qui est corrélé à la réputation de manière générale de chacun des agents qui compose le système. Les agents doivent de fait, selon les auteurs, maintenir une bonne réputation, mesurée à partir de la réputation moyenne de chacun, afin d'assurer la survie du système.

L'évaluation de la réputation d'un agent est calculée donc à partir d'une pondération de la valeur de la satisfaction de chacun des rôles définit ci-dessus, valeur qui comme nous l'évoquions est spécifique à chaque société.

La notion d'influence (ou notoriété) comme métrique de la réputation à évaluer a également été introduite dans certains modèles. L'un des premiers modèles à utiliser cette approche est PAGERANK de Page & Brin. (1998). Les premiers moteurs de recherche Web tels qu'Altavista présentaient simplement toutes les pages web qui correspondaient aux mots clés saisis par l'utilisateur, ce qui entraînait souvent la liste de pages trop nombreuses et non pertinentes dans les résultats de recherche. La proposition d'Altavista pour résoudre ce problème était d'offrir des moyens avancés de combiner des mots-clés basés sur la logique binaire. PAGERANK proposé par Brin & Page (1998)[264] représente un moyen de classer les meilleurs résultats de recherche en fonction de la réputation d'une page.

En résumé, le PageRank classe une page en fonction du nombre d'autres pages qui renvoient vers elle. Perçu comme un système de réputation, PAGERANK facilite la collecte d'hyperliens vers une page donnée qui peut être considérée comme une information publique et peut être combinée pour obtenir un score de réputation. Un seul lien hypertexte vers une page Web donnée peut être considéré comme une évaluation positive de cette page web. Le moteur de recherche de Google est basé sur l'algorithme PAGERANK et la popularité croissante de Google au détriment d'Altavista était évidemment due aux résultats de recherche supérieurs fournis par l'algorithme PAGERANK. D'autres approches sont construites sur le même principe : PATROL de Tajeddine & al. (2007), Chen & al. (2006) avec leur modèle CUTEAID, Osman & al (2015) MORE, etc. La mesure de l'ensemble de ces attributs s'effectue, selon les modèles utilisés, à partir de cinq (5) principales méthodes (que nous présenterons par la suite) qui combinées, visent à attribuer une réputation à une entité.

---

[264] Brin Sergey et Page Lawrence (1998). *The PageRank Citation Ranking : Bringing Order to the Web, Computer Science Department, Stanford University*. Consulté le 20 novembre 2016, à l'adresse http://ilpubs.stanford.edu:8090/422/1/1999-66.pdf

2.3.b. Les méthodes de calcul utilisées

Au compte de ces méthodes utilisées dans le calcul de la réputation dans les modèles théoriques, on retrouve : l'approche par somme ou moyenne, les approches bayésiennes, les modèles construits autour des flux (*flow based*), les modèles basés sur la logique floue (*Fuzzy logic*) et enfin les approches liées à la théorie des croyances (*belief theory)*.

- L'approche par moyenne ou somme des données

La première approche utilisée dans le calcul de la réputation (via des scores ou indices par exemple) consiste à établir une équation à partir des notes positives et négatives afin d'obtenir un score final. C'est le principe utilisé notamment dans le modèle de réputation d'eBay[265] qui fait partie des précurseurs en termes d'évaluation de la réputation. La plateforme d'enchères eBay via notamment son *feedback forum*, donne à l'acheteur et au vendeur la possibilité de se noter selon trois attributs positif, négatif ou neutre (c'est-à-dire 1, -1, 0) après une transaction. Les acheteurs et les vendeurs ont également la possibilité de laisser des commentaires tels que "Transaction fluide, merci!" qui sont typiques dans les cas positifs ou "Acheteurs méfiez-vous!" dans le rare cas négatif. À partir donc du *Feedback Forum* eBay recueille toutes les notes et calcule les scores, le score de réputation total en cours de chaque participant étant la somme des notes positives (d'utilisateurs uniques) moins la somme des notes négatives (d'utilisateurs uniques). Afin de fournir des informations sur le comportement plus récent d'un participant, le total des notes positives, négatives et neutres prend en compte une certaine temporalité i) six derniers mois, ii) le mois dernier et iii) 7 derniers jours par exemple.
En général, les notes observées sur eBay sont étonnamment positives et le système de notation se présente comme suit : les acheteurs donnent des évaluations sur les vendeurs 51,7% du temps, et les vendeurs fournissent des évaluations sur les acheteurs 60,6% du temps (Resnick & al., 2002)[266].

---

[265] Site web d'Ebay https://www.ebay.fr

[266] P. Resnick and R. Zeckhauser. Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System. In M.R. Baye, editor, *The Economics of the Internet and E-Commerce*, volume 11 of *Advances in Applied Microeconomics*. Elsevier Science, 2002.

Il est intéressant de noter à la suite de Josang et al. (2007) que de toutes les évaluations fournies, moins de 1% est négatif, moins de 0,5% est neutre et environ 99% est positif. Cela peut s'expliquer par une forte corrélation entre les évaluations des acheteurs et des vendeurs, ce qui suggère qu'il existe un certain degré de réciprocité des évaluations positives et des représailles des évaluations négatives. Ce qui pose une problématique essentielle dans la véracité même des avis. Si on part du postulat que l'obtention de notes honnêtes et équitables est un objectif, il serait peut être plus opportun de ne pas laisser les vendeurs évaluer les acheteurs. On le voit, le système de réputation d'eBay est très primitif et peut être assez trompeur. Avec si peu de notes négatives, un participant avec 100 notes positives et 10 notes négatives devrait intuitivement sembler avoir une moins bonne réputation qu'un participant avec 90 notes positives et aucune négative, mais sur eBay, ils auraient le même score de réputation totale.

Certaines plateformes proposent un schéma légèrement plus avancé qui consiste à calculer le score de réputation comme la moyenne de toutes les évaluations, et ce principe est utilisé dans les systèmes de réputation de nombreux sites web commerciaux d'avis ou marchands, tels que Epinions[267] et Amazon. À l'origine, sur Epinions, les avis sur les produits et les entités (shops) rédigés par les membres se présentaient sous forme de commentaires (textes) et d'évaluations quantitatives de 1 à 5 étoiles pour un ensemble d'aspects tels que la facilité d'utilisation, l'autonomie de la batterie, etc. pour les produits, et la facilité de commande, le service client, les délais de livraison, etc. pour les magasins. Les autres membres pouvaient donc évaluer les avis selon une échelle inutiles, plutôt utiles, utiles et très utiles, et ainsi contribuer à déterminer le degré de visibilité de l'avis, et accroître le statut de l'évaluateur. Un membre pouvait donc obtenir le statut *Advisor*, *Top Reviewer ou Category Lead* (le plus élevé) en fonction des notes cumulées sur l'ensemble de ses avis sur une période donnée. Ces membres sont automatiquement sélectionnés chaque mois en fonction de la qualité de leur

[267] qui à l'origine était un site général d'avis des consommateurs créé en 1999, acquis en 2003 par DealTime, plus tard Shopping.com, et par eBay en 2005

évaluation, et référencés sur *Epinions Web of Trust*[268], où un membre peut décider, selon les données dont il dispose, de faire confiance ou bloquer un autre membre.

Plateforme de e-commerce, Amazon[269] à l'origine était une librairie en ligne qui permettait aux membres d'écrire des critiques de livres sous forme d'avis textuels ou de notes (1 à 5 étoiles). La note d'un livre utilisait la moyenne de tous les avis attribués et les utilisateurs, y compris les non-membres, pouvaient attribuer une note sur ces avis (utiles ou inutiles) avec le nombre de votes utiles ainsi que le nombre total de votes affichés avec chaque avis. En fonction du nombre de votes utiles que chaque critique (personne ayant donné son avis) a reçu, ainsi que d'autres paramètres non révélés publiquement, Amazon détermine le rang de chaque critique, et les évaluateurs qui font partie des 1000 meilleurs se voient attribuer le statut de Top 1000, Top 500, Top 100, Top 50, Top 10 ou # 1 *Reviewer*.

Amazon dispose également d'un système de personnes favorites (*favourite reviewers*), dans lequel chaque membre peut choisir d'autres membres comme critiques préférés, et le nombre d'autres membres dont un critique spécifique est répertorié comme personne préférée influence également le classement de ce critique. Le modèle de réputation d'Amazon a été longtemps critiqué. D'une part, le critique d'Amazon # 1 est généralement celui qui publie le plus de critiques possible sur un très grand nombre de livres, nécessitant un énorme effort de lecture ce qui n'est pas toujours le cas, donc présentant un biais du système de *ranking*. D'autre part, les critiques ayant atteint le classement des 100 meilleurs ont signalé une augmentation soudaine des votes négatifs visant à réduire leur score de réputation.
Les modèles avancés de cette catégorie calculent donc, comme on peut le constater, une moyenne pondérée de toutes les notes, où la valeur de la note peut être déterminée par des facteurs tels que la fiabilité / réputation de l'évaluateur, la récence ou non de l'évaluation, l'écart entre la note et le score actuel par exemple.

---

[268] Le Web of Trust Epinions est un schéma simple dans lequel les membres peuvent décider de faire confiance ou de bloquer un autre membre. La liste des membres de confiance d'un membre représente le Web de confiance personnel de ce membre. Comme déjà mentionné, le Web of Trust influence la sélection automatisée des Top Reviewers et des Conseillers. Le nombre de membres (et leur statut) qui font confiance à un membre donné contribuera à ce que ce membre obtienne un statut plus élevé. Le nombre de membres (et leur statut) qui bloquent un autre membre aura un impact négatif sur ce membre obtenant un statut plus élevé (Josang & al., 2007, survey, p.25). https://www.oasis-open.org/committees/download.php/28303/JIB2007-DSS-Survey.pdf

[269] Site web de la plateforme Amazon https://www.amazon.fr

- L’approche basée sur les flux (*flow based*)

Les modèles basés sur les flux fournissent une méthode automatisée pour agréger les informations de confiance/ réputation. Ils utilisent les chaînes de Markov comme base mathématique, et fournissent une méthode automatisée pour regrouper toutes les informations de réputation disponibles. Cependant, ils ne sont pas en mesure d'exprimer le niveau de confiance dans les valeurs de réputation obtenues. En règle générale, chaque fois que l'utilisateur *I* a une transaction avec un autre utilisateur *J*, il peut évaluer la transaction comme positive, neutre ou négative. Dans un modèle de réputation basé sur les flux, ces notations sont agrégées afin d'obtenir une chaîne.

Les systèmes qui calculent la réputation par itération transitive à travers des chaînes en boucle ou arbitrairement longues peuvent être donc appelés modèles de flux.

Certains modèles de flux supposent une valeur constante de confiance / réputation pour l'ensemble de la communauté, valeur pouvant être répartie entre les membres de la communauté. En effet, les participants ne peuvent accroître leur confiance / réputation qu'au détriment des autres. Le PAGERANK de Google, l'algorithme Appleseed[270] et le schéma de réputation d'Advogato[271] (communauté open source de programmers où les membres s’évaluent selon leurs compétences avec des positions : *Apprentice* (*lowest*), *Journeyer* (*medium*) or *Master* (*highest*) appartiennent à cette catégorie. En général, la réputation d’un participant augmente en fonction du flux entrant et diminue en fonction du flux sortant. Dans le cas de Google, de nombreux hyperliens vers une page web contribuent à une augmentation du PANGERANK, tandis que de nombreux hyperliens d'une page Web contribuent à une diminution du PageRank pour cette page Web.

Les modèles de flux n'exigent pas toujours que la somme des scores de réputation / confiance soit constante. On peut prendre comme exemple le modèle EIGENTRUST qui calcule les scores des agents dans les réseaux *P2P* par multiplication et agrégation répétées et itératives

---

[270] C.-N. Ziegler and G. Lausen. Spreading Activation Models for Trust Propagation. In *Proceedings of the IEEE International Conference on e-Technology, e-Commerce, and e-Service (EEE ’04)*, Taipei, March 2004

[271] T. Grandison and M. Sloman. A Survey of Trust in Internet Applications. *IEEE Communications Surveys and Tutorials*, 3, 2000 ; R. Levien. *Attack Resistant Trust Metrics*. PhD thesis, University of California at Berkeley, 2004

des scores de confiance le long des chaînes transitives jusqu'à ce que les scores de tous les agents membres de la communauté *P2P* convergent vers des valeurs stables.

- Les modèles bayésiens

Les systèmes bayésiens sont basés sur le calcul des scores de réputation par mise à jour statistique. Le score de réputation *a posteriori* (c'est-à-dire mis à jour) est calculé en combinant le score de réputation *a priori* (c'est-à-dire précédent) avec le nouveau score (Josang & Ismail, 2002[272]; Mui & al., 2001/2002[273] ; Josang & al., 2005[274]). Le score de réputation peut être représenté sous la forme d'un tuple bêta PDF[275] (α, β) (où α et β représentent respectivement la valeur des notes positives et négatives), ou sous la forme de valeurs probabilistes, et éventuellement accompagné de la variance ou d'un paramètre de confiance[276]. L'avantage des systèmes bayésiens est qu'ils fournissent une base théoriquement solide pour le calcul des scores de réputation.

- L'approche axée sur la logique floue (*fuzzy logic*)

La réputation peut être représentée comme concept flou sur le plan linguistique, où les fonctions d'appartenance décrivent dans quelle mesure un agent peut être perçu comme par exemple digne de confiance ou pas digne de confiance. La logique floue (FL) est une approche informatique basée sur des « degrés de vérité » plutôt que sur la logique booléenne habituelle « vrai ou faux » (1 ou 0) sur laquelle l'ordinateur moderne est construit.

---

[272] A. Jøsang and R. Ismail. The Beta Reputation System. In *Proceedings of the 15th Bled Electronic Commerce Conference*, June 2002.

[273] L. Mui, M. Mohtashemi, and C. Ang. A Probabilistic Rating Framework for Pervasive Computing Environments. In *Proceedings of the MIT Student Oxygen Workshop (SOW'2001)*, 2001
L. Mui, M. Mohtashemi, and A. Halberstadt. A Computational Model of Trust and Reputation. In *Proceedings of the 35th Hawaii International Conference on System Science (HICSS)*, 2002.

[274] A. Withby, A. Jøsang, and J. Indulska. Filtering Out Unfair Ratings in Bayesian Reputation Systems. *The Icfain Journal of Management Research*, 4(2):48–64, 2005.

[275] Abréviation du terme *probability density functions*

[276] A. Jøsang and R. Ismail. The Beta Reputation System. In *Proceedings of the 15th Bled Electronic Commerce Conference*, June 2002

L'objectif de ce procédé est de produire des *outputs* en réponse de manière générale à des *inputs* complexes, non clairement définis ou incomplets, et le plus proche possible du raisonnement humain. En effet, l'approche de FL imite le mode de prise de décision chez l'homme qui implique toutes les possibilités intermédiaires entre les valeurs numériques OUI et NON. Concept proposé par Lotfi Zadeh de l'Université de Californie à Berkeley dans les années 1960, il partait du postulat que la compréhension informatique du langage naturel ne pourrait se résumer en termes absolus de 0 et 1, et que tout ne peut finalement pas être descriptible en termes binaires. Si la logique conventionnelle qu'un ordinateur peut comprendre prend une entrée précise et produit une sortie définie comme VRAI ou FAUX (ce qui équivaut au OUI ou au NON de l'homme), l'inventeur de la logique floue souligne que contrairement aux ordinateurs, la prise de décision humaine comprend une gamme de possibilités entre OUI et NON : certain, probablement sûr, peut être, pas du tout certain, etc. Plus concrètement, il s'agit de l'ensemble des *outputs* permettant de répondre à un *input* spécifique (tel que perçu par le raisonnement humain). Le schéma ci-dessous permet de résumer la distinction entre la logique floue et le raisonnement booléen.

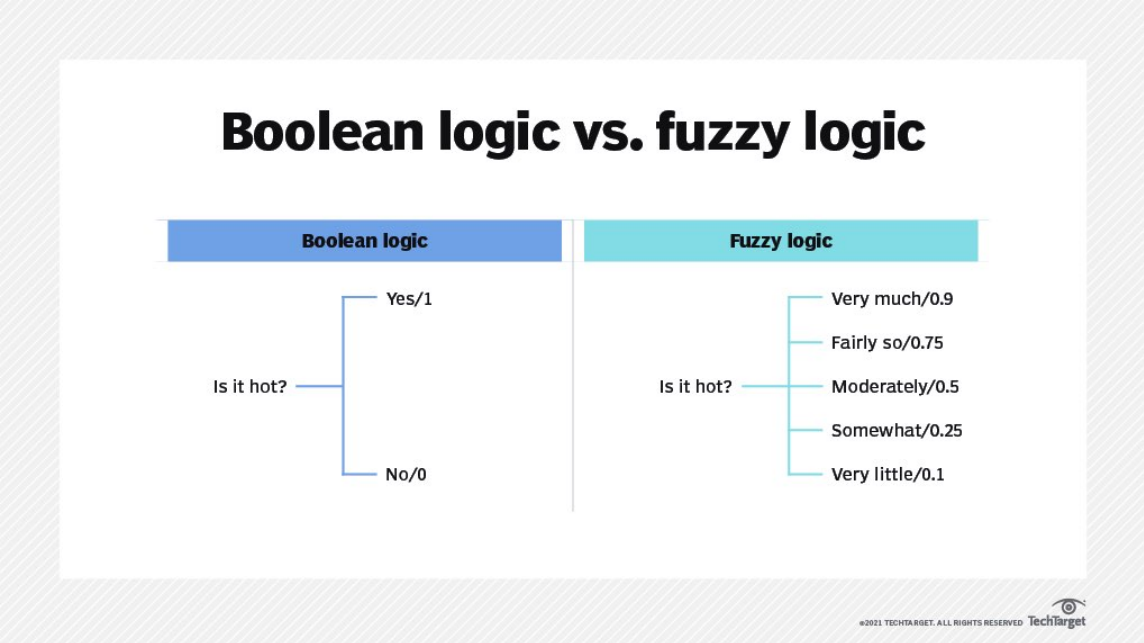

Figure n°8 : Distinction entre les logiques booléennes et floues, source TechTarget[277]

[277] La différence entre la logique floue et celle binaire de TechTarget https://searchenterpriseai.techtarget.com/definition/fuzzy-logic

Les modèles construits autour de la FL fonctionnent selon une architecture à quatre (4) niveaux : le module de *fuzzication*, le procédé IF - THEN, l'implémentation de la méthode et le résultat ou *output* (la *defuzzification*).

- Sur la mise en place du module de *fuzzication*, il s'agit de transformer les *inputs* du système (qui sont des nombres nets) en ensembles flous. Le signal d'entrée peut être de manière générale réparti sous cinq principaux critères : excellent, bien, plutôt positif, négatif, très négatif / médiocre (formulations qui diffèrent selon le contexte) ;
- Le choix du procédé IF - THEN[278] vise, à partir des inputs choisis, à déterminer l'ensemble des outputs qui permettent de résoudre la problématique ;
- L'implémentation de la méthode utilise des algorithmes d'IA et plus précisément de ML afin de simuler le processus de raisonnement humain en faisant une inférence floue sur les entrées et les règles IF -THEN (apprentissage machine et entrainement du *scorer*) ;
- La dernière étape consiste à transformer l'ensemble flou obtenu à l'issue du processus par le moteur d'inférence en une valeur nette comprise entre un intervalle donné.

Le système de réputation REGRET proposé par Sabater & Sierra (2001, 2002) est le premier a proposer une approche de mesure basée sur la logique floue. Dans le schéma de Sabater & Sierra, ce qu'ils appellent la réputation individuelle est dérivé d'informations privées sur un agent donné, ce qu'ils appellent la réputation sociale est dérivé d'informations publiques sur un agent, et ce qu'ils qualifient de réputation dépendante du contexte est dérivée d'informations contextuelles. Plusieurs autres modèles s'appuient également sur cette technique au compte desquels on peut par exemple retrouver les approches proposées par Carbo & al. (2002) , Buchegger & Le Boudec (2003), Song & al. (2005), Fullan & al. (2005), Sabater & al. (2006), Tian & al. (2009), Li & al. (2013).

[278] Qu'on peut traduire par Si - Alors (la / les données d'entrée et l'ensemble des résultats possibles)

- De l’apport de la logique subjective ou de la théorie des croyances (*Subjective logic ou belief theory*)

La *belief theory* ou théorie des croyances est une approche liée à la théorie des probabilités, mais où la somme des probabilités sur tous les résultats possibles ne correspond pas nécessairement à 1, et la probabilité restante est interprétée comme une incertitude. Elle étend les modèles probabilistes à une notion explicite d'incertitude, mais le calcul de la réputation dépend de la structure du réseau de confiance/ réputation et nécessite souvent de rejeter des informations. La *subjective logic* est enracinée dans la théorie de Dempster – Shafer (Josang & al., 2002/ 2005). Elle fournit une base mathématique pour traiter les opinions, et a la capacité naturelle d'exprimer l'incertitude de manière explicite. Intuitivement, l'incertitude intègre une marge d'erreur dans le calcul de la réputation en raison de la quantité (limitée) d'informations disponibles sur la confiance. La SL qui est le plus souvent utilisée dans les modèles hybrides (réputation & confiance) perçoit la réputation comme *input* de la confiance. Jøsang (1999,2001) fait partie des premiers à proposer une métrique de croyance / confiance appelée opinion dénotée (*opinion denoted*), qui exprime la croyance d’un agent *A* en la véracité de la déclaration d’un autre agent *X*[279].

Yu et Singh (2002) ont également utilisé la théorie des croyances pour représenter les scores de réputation. Dans leur schéma, deux types de scénarii (résultats) sont possibles à savoir qu'un agent *A* est digne de confiance ou non digne de confiance. De manière assez simple, les évaluations fournies par les agents individuels sont des mesures de croyance déterminées en fonction de l’historique de comportements de *A* avec des agents individuels comme dignes de confiance ou non, en utilisant des valeurs seuils prédéfinies pour ce qui constitue un comportement digne de confiance et non fiable. Ces mesures de croyance sont ensuite combinées en utilisant la règle de Dempster-schafer afin de calculer le score de réputation.

---

[279] À titre d'exemple, supposons qu'Alice ait besoin de faire réviser sa voiture et qu'elle demande à Bob de recommander un bon mécanicien automobile. Quand Bob recommande David, Alice aimerait avoir un deuxième avis, alors elle demande à Claire son opinion sur David. Lorsque les références sont exprimées sous forme d'opinions, chaque scénario (chemin) de confiance transitif (*trust transitivity)* Alice → Bob → David et Alice → Claire → David peut être calculé, avec l'idée que les références de Bob et Claire sont actualisées comme une fonction la confiance d'Alice en Bob et Claire respectivement. Enfin, les deux scénarii peuvent être combinés suivant une forme de logique subjective. Les opinions peuvent être mises à jour suivant l’approche bayésienne décrite ci-dessus

À l'issue de la phase d'analyse, il sera possible d'attribuer une réputation à un agent, qui dans une approche générale se présente sous forme de score ou indice avec pour certains, des étiquettes associées à :

- Une échelle sous forme textuelle avec des variantes selon les modèles : positive, négative ou neutre ; très mauvais (VB: *very bad*), mauvais (B : *bad*), neutre (N), bon (G : *good*) et très bon (VG : *very good*)[280] ; inutiles, plutôt utiles, utiles et très utiles sur la plateforme Epinions par exemple ;
- Pour d'autres, notamment les modèles construits sur la structure des places de marché ou du e-commerce (Amazon, Ebay), les avis sont présentés sous forme de note ou étoiles (1 à 5 étoiles) ;
- Pour certains encore des classements (top 10, n°1, etc.) : PAGERANK (1998), Advogado[281] par exemple ;
- Ou encore sous forme d'intervalle (*tuple*) de valeurs probabilistes {-1, 1} qui donne une valeur numérique[282] : utilisé par plusieurs des modèles Josang & Ismail, 2002[283]; Mui & al., 2001/2002[284] ; Josang & al., 2005[285] ; Yu & Singh (2002) ;

Mesurer la réputation dans la littérature scientifique, nous l'avons vu, suppose trois principales étapes : déterminer le type d'environnement au sein duquel seront collectées les données, identifier les métriques de réputation sur lesquelles s'appuyer dans l'évaluation, et

---

[280] pour Sabater-Mir et al.(2006)

[281] T. Grandison and M. Sloman. A Survey of Trust in Internet Applications. *IEEE Communications Surveys and Tutorials*, 3, 2000 ; R. Levien. *Attack Resistant Trust Metrics*. PhD thesis, University of California at Berkeley, 2004

[282] C'est notamment le cas des modèles qui utilisent une approche bayésienne / probabiliste ou construits autour de la logique subjective

[283] A. Jøsang and R. Ismail. The Beta Reputation System. In *Proceedings of the 15th Bled Electronic Commerce Conference*, June 2002.

[284] L. Mui, M. Mohtashemi, and C. Ang. A Probabilistic Rating Framework for Pervasive Computing Environments. In *Proceedings of the MIT Student Oxygen Workshop (SOW'2001)*, 2001
L. Mui, M. Mohtashemi, and A. Halberstadt. A Computational Model of Trust and Reputation. In *Proceedings of the 35th Hawaii International Conference on System Science (HICSS)*, 2002.

[285] A. Withby, A. Jøsang, and J. Indulska. Filtering Out Unfair Ratings in Bayesian Reputation Systems. *The Icfain Journal of Management Research*, 4(2):48–64, 2005.

enfin choisir la méthode mathématique (qui dépend en partie des deux premières phases) de calcul. Plusieurs constats peuvent être faits.

Premièrement, il est important de retenir que du type d'environnement au sein duquel interagissent les agents, dépend la méthode de collecte des données à déployer. En effet, et comme nous l'avons précédemment vu, les modèles construits autour d'une structure centralisée, utilisent dans la collecte des données et dans leur analyse, une autorité centrale. *A priori*, ce type de modèles sont construits autour de deux principales métriques de la réputation : l'identité et la confiance. En effet, le concept d'identité et notamment les attributs identitaires renseignés lors de l'entrée au sein du système servent à garantir que l'agent qui évalue a, en amont, déjà interagi avec l'entité évaluée, ce qui garantit une forme de fiabilité de son jugement. L'objectif principal de ces systèmes de réputation est donc, hormis la mise à disposition d'éléments réputationnels, de garantir en certains points (ce qui par ailleurs n'est pas toujours avéré comme nous l'avons déjà mentionné) la pertinence, l'authenticité de l'ensemble des informations collectées. Ce qui les distingue des systèmes de réputation décentralisés qui, de par leur structure (ouverts), offrent la possibilité à n'importe quel agent de proposer un jugement de valeur sur n'importe quelle entité.

Selon la structure du modèle, la phase de collecte des données vise à identifier d'une part les diverses sources de collecte, et d'autre part, à catégoriser les informations. Le processus de catégorisation utilise plusieurs métriques de la réputation : l'identité souvent substituée à l'image dans certains modèles ; la notion d'influence, de notoriété ou de popularité ; et la confiance. À partir de ce corpus initial, le calcul de la réputation s'articule autour de plusieurs méthodes : l'approche par somme ou moyenne, les approches bayésiennes, les modèles construits autour des flux (*flow based*) ou la logique floue (*Fuzzy logic models),* et enfin les approches liées à la théorie des croyances (*belief theory*).

Si les deux premières sections de ce chapitre nous ont permis d'identifier les différents critères de classification des modèles conceptuels, il nous paraît intéressant de présenter dans la section suivante les grandes catégories de classification qui nous serviront à déterminer le rapport entre les pratiques professionnelles dans la mesure de la réputation et les approches théoriques.

3. Proposition d'une nouvelle taxonomie : nos dimensions de classification

L'objectif premier de cette partie était d'identifier des catégories de classification des modèles théoriques issus de la littérature scientifique et de les comparer avec approches méthodologiques proposées par les professionnels de la réputation. Pour ce faire, nous nous sommes dans un premier temps appuyée sur les tentatives de classification des auteurs qui avant nous se sont intéressés à la question. À partir de leurs travaux, nous avons identifié huit (8) principaux critères et attributs, qui nous permettent de distinguer quatre (4) principales catégories de modèles : les modèles sommaires, les modèles ad hoc (construits à partir d'une des métriques de l'e-réputation et qui sont adaptés à une utilisation singulière), les modèles sophistiqués et les modèles modulaires (flexibles ou génériques).

3.1. Les modèles primaires ou sommaires

Par modèles sommaires, nous entendons ceux qui fonctionnent selon une architecture et une méthode de mesure de la réputation très simpliste.

Au compte de ces derniers, on retrouve les premiers modèles construits selon l'architecture des sites d'enchères et/ou de e-commerce, et qui fonctionnent sous le principe du *C2C* que sont Ebay (2002) et Amazon (2002). Ces plateformes qui font partie des précurseurs de la mesure de la réputation et de la confiance des agents au sein d'environnements ouverts et incertains offrent la possibilité à des acheteurs d'attribuer une évaluation sur la qualité de service des produits d'un vendeur. Cette évaluation, qui peut soit être positive, négative ou neutre, utilise une forme simple S = P – N avec P comme nombre d'avis positifs et N relatifs aux avis négatifs attribués par les clients, score à partir duquel la plateforme attribue un *ranking* global du vendeur (le taux d'avis positifs) sous forme de pourcentage à partir de la formule $R = (P - N)/(P + N)$ ($1 \geq R \geq 0$) qui permet de déduire de la fiabilité de l'agent et donc de le récompenser si R>98%.

À la suite de ces modèles et toujours dans la même optique, Zacharia (1999) a développé sur la base de la théorie des jeux, deux autres mécanismes de réputation « SPORAS et HISTOS » appliqués aux plateformes de e-commerce. La particularité du modèle SPORAS est sa

robustesse par rapport aux changements de comportement d'un utilisateur et la mesure de la fiabilité de la réputation de ce dernier. En effet, une des caractéristiques de ce modèle est que désormais, seule la note la plus récente entre deux utilisateurs est prise en compte dans la mesure de la réputation. Intervient donc la notion d'actualisation, qui permet une mesure en temps réel, mais aussi et surtout une mise à jour perpétuelle des notes des utilisateurs. Un autre élément important de ce modèle qu'on ne retrouvait pas dans le premier, est la prise en compte du caractère fiable de la réputation de l'utilisateur sur la base des travaux de Glicko (Glickman, 1999), de la mesure de l'écart-type des valeurs de réputation qu'ils ont laissées. Dans le modèle SPORAS, plus la réputation d'un utilisateur est élevée, moins son évaluation subit de modifications après les mises à jour, et les utilisateurs ayant une faible réputation subissent de fait plus de modifications de leur évaluation.

Le modèle HISTOS (Zacharia, 1999) part du postulat d'un manque de personnalisation de la valeur de la réputation du modèle SPORAS. En effet, dans ce modèle et contrairement au précédent, la valeur de la réputation est ici subjective et les informations collectées sont issues à la fois d'interactions directes et indirectes entre les utilisateurs. Les évaluations ici s'effectuent à l'aide de graphes avec les agents représentés par des nœuds et les extrémités comme source d'information la plus récente d'un agent vers un autre. On note dans ce modèle l'existence d'un noyau central avec une structure pyramidale, et où la réputation d'un agent au niveau X du graph (avec $X > 0$) est calculée en tant que moyenne pondérée des valeurs de classement attribuées par les agents du niveau X-1 à cet agent.

3.2. Les modèles ad hoc

Afin de simplifier la classification des modèles et leur analyse, nous avons identifié une seconde catégorie permettant de les distinguer selon l'usage spécifique auquel ils sont destinés. Par ad hoc, nous entendons les modèles de mesure qui sont adaptés à une utilisation contextuelle spécifique. Il s'agit entre autres de ceux utilisés ou déployés uniquement pour l'un de ces usages : une typologie de domaine ou un secteur d'activité donné, une problématique singulière, ou un objet de mesure spécifique.

Dans la littérature scientifique, les modèles ad hoc qui sont les plus représentatifs sont regroupés selon deux dimensions : les modèles ad hoc centralisés et ceux décentralisés. Au compte des modèles ad hoc construits autour d'une structure décentralisée, on retrouve ceux appliqués aux plateformes de partage de documents (Lu & al., 2016 ; Skoric & al., 2016 ; Fan & al., 2011 ; Tian & yang, 2011 ; Zhou & al., 2007; Chen & al., 2007), au bookmarking social et plateformes de micro-blogging (Portmann & al., 2015), aux places de marché (Gutowska & Buckley, 2008 ; Song & al., 2005 ; Regan & Cohen, 2005 ; Tran & Cohen, 2004) ; aux espaces ou forums de discussion (Chen & al., 2006), aux réseaux sans fil ad hoc ou *manet*[286] (Almulla & al., 2011 ; Liu & Issarny, 2009 ; Buchegger & Le Boudec, 2003 ; Michiardi & Molva, 2002), aux moteurs de recherche (Page & al., 1998).

Les approches ad hoc centralisées elles, représentent une minorité dans la littérature scientifique et font notamment référence à certaines plateformes de e-commerce et aux sites d'avis (Abdel Hafez & al., 2016 ; Ransi & Kobti, 2014 ; Fang & al., 2012 ; Vogiatzis & al., 2010) et à certains systèmes multi-agents (Urzica & al., 2011 ; Bertocco & Ferrari, 2008).

3.3. Les modèles sophistiqués

L'une des problématiques de la mesure de la réputation en ligne d'agents est d'arriver à collecter des données assez exhaustives sur un objet à travers diverses sources disponibles, d'implémenter des mécanismes permettant d'identifier les potentiels agents pouvant présenter des risques ou biais dans la mesure de la réputation, de détecter les biais dans les données, etc. Afin de pallier ces problématiques, certains auteurs se sont proposé d'implémenter des mécanismes permettant tant de garantir la sécurité des agents, de détecter les biais dans les données collectées et les agents frauduleux (dans les espaces d'échange), d'implémenter une diversité de sources de données, d'utiliser plusieurs des métriques de l'e-réputation, etc.

L'un des éléments pris en compte dans la classification des modèles sophistiqués est la question de la granularité qui fait référence aux contextes d'utilisation des modèles théoriques. Ces modèles utilisent dans la mesure de la réputation plusieurs typologies de sources (à minima trois afin de figurer dans cette catégorie), avec l'implémentation de mécanismes de détection des biais dans les données collectées. L'un des premiers modèles

[286] Mobile ad hoc network ou wireless ad hoc network

qu'on retrouve dans cette catégorie est celui REGRET de Sabater et Sierra (2002). C'est un modèle hybride c'est-à-dire qu'il permet à la fois d'évaluer la réputation et la confiance d'un agent. Le modèle REGRET est adapté pour les systèmes multi-agents génériques, avec la prise en compte d'un mécanisme permettant de détecter les données frauduleuses ou biaisées. Construit à partir des métriques de confiance et de réputation, il utilise autant les données explicites (avis, commentaires, etc.) que celles de nature implicite (appartenance à un groupe par exemple). Sur le même principe, on retrouve le modèle TRAVOS (Teacy & al., 2006) et celui de Jurca & Faltings (2003) qui utilisent une approche probabiliste (et plus particulière la théorie des jeux : le dilemme du prisonnier). Ils sont construits à partir de la métrique "identité" et incluent notamment un principe de *log* (connection via un couple identifiant + mot clé) avec la particularité d'avoir implémenté un mécanisme sécuritaire visant à punir ou exclure les agents frauduleux. Cette propension à mettre en place des mécanismes sécuritaires se retrouve également dans le modèle FIRE de Huynh & al. (2004, 2006). À l'instar des concepts de confiance et d'identité, d'autres métriques de la réputation sont utilisées dans cette catégorie. On retrouve notamment le concept d'image dans les modèles REPAGE (Sabater & al., 2006) et BDI+REPAGE (Pinyol & al., 2012), concept associé à celui de risque dans l'approche proposée par Tian & al. (GROUPREP, 2006), et celui d'influence et de popularité qu'on retrouve dans le modèle PATROL de Tajeddine & al. (2007) par exemple.

3.4. Les approches modulaires / génériques (flexibles)

Dans la littérature sur la mesure de la réputation en ligne, certains auteurs ont proposé des modèles sous forme modulables et/ou flexibles pouvant être adaptés à une pluralité d'utilisation ou sous forme de bêta test en open source afin de faire l'objet d'améliorations. Cette catégorie de modélisation permet donc de regrouper diverses catégories de modèles conçus pour être adaptés à une multitude de contextes. Peut être qualifié de modèle multi contextes, un modèle qui, dans sa conception, est donc modulaire, adaptable à divers contextes, environnements ou applications.

À la différence des modèles sophistiqués, les approches modulaires même si elles peuvent être déployées dans une pluralité d'environnement, ne prennent pas en compte la totalité des attributs qu'on retrouve dans la structure des modèles sophistiqués : à titre d'exemple,

certains n'implémentent pas de mécanismes de détection des biais dans les données, d'autres ne sont construits qu'autour d'une métrique de réputation ou encore offrent un nombre limité de sources de collecte des données avec parfois l'utilisation de méthodes de mesure les plus simplistes. On retrouve au sein de cette catégorie, les modèles qui font fi d'un environnement spécifique et peuvent être adaptés (d'où leur structure modulaire) à des contextes d'utilisation multiples. À titre d'exemple, on peut citer ARM de Yu & al. (2014), le modèle FR TRUST de Javanmardi & al. (2014), ceux adaptables à minima à deux typologies d'environnements comme c'est le cas pour le modèle de Yan & al. (2015) déployé autant sur les plateformes de e-commerce que sur les plateformes socionumériques, et une autre catégorie d'approches sous forme de bêta test, adaptée à l'ensemble des places de marché et au sein de laquelle on retrouve le modèle ART de Fullan & al. (2005) ou le modèle de Kerr & Cohen (TREET, 2009). Le tableau ci-dessous reprend de manière synthétique, les quatre principales catégories de classification des modèles utilisées dans nos travaux (avec pour chacune, les attributs servant à les étayer), et le tableau de classification de l'ensemble des modèles théoriques se trouve en annexe 1.

CATÉGORIES	ATTRIBUTS
SOMMAIRE	- Architecture simple - Méthode de calcul (somme ou moyenne) - Sites d'avis et premiers systèmes d'e-commerce
AD HOC	- Centralisé ou décentralisé - Déployé au sein d'un seul environnement - Un "objet" d'évaluation spécifique
GÉNÉRIQUE	- Environnements multiples - Plusieurs métriques de la réputation - Méthode d'évaluation générique (modulable)
SOPHISTIQUÉ	- Environnements multiples - Catégories des sources de données exhaustives - Plusieurs métriques de la réputation évaluées - Détection des données frauduleuses ou biaisées - Mécanisme sécuritaire - Hybride (réputation + confiance)

Tableau n°4 : Nos dimensions de classification des modèles théoriques de la réputation

Conclusion du chapitre 4

La mesure de la réputation, nous le voyons, a fait l'objet de nombreuses études dans la littérature scientifique surtout anglo-saxonne. Les modèles de la réputation décrits ont été conçus pour de nombreux usages allant des applications les plus simples de type plateformes de recommandation ou d'avis clients à celles les plus complexes liées à l'intelligence artificielle et aux objets connectés en passant par celles plus génériques liées aux réseaux sociaux et à leur analyse ou encore aux moteurs de recherche.

Afin donc d'en proposer une classification, plusieurs auteurs se sont attelés à décrire des catégories et/ou dimensions visant à en faciliter la lecture. C'est notamment le cas de Sabater et Sierra (2005), Koutrouli et Tsalgatidou (2006), Marti and Garcia-Molina, (2006), Wang and Vassileva, (2007), Hoffman et al., (2009), Yuan et al., (2012), Vavilis et al., (2014), Hendrikx et al., (2015), Granatyr et al. (2015). Les principales thématiques autour desquelles s'articulent les dimensions de classification sont notamment relatives à la collecte des données (le type de données et les sources), à la structure même des modèles (approche centralisée, décentralisée, granularité, l'existence de mécanismes de gestion des biais) et à la méthode d'attribution de la valeur finale de la réputation (type de visibilité, stockage, méthode mathématique de calcul et format).

À partir de ces grandes dimensions, et afin de déceler un éventuel lien entre les modèles et les approches méthodologiques professionnelles mobilisées dans la mesure de la réputation, nous avons défini quatre (4) grandes catégories de classification (avec pour chacune plusieurs critères pris en compte) : les approches primaires, ad hoc, génériques et sophistiquées, dimensions étayées par plusieurs indicateurs ou attributs. Ces catégories d'analyse que nous avons identifiées nous serviront par la suite (dans le chapitre 7 notamment) de base pour notre travail d'analyse des pratiques professionnelles.
Avant d'y revenir, il nous semble intéressant de questionner le marché de l'e-réputation, ce qui suppose un retour sur son évolution, les principaux acteurs qui le composent, et plus particulièrement les méthodes de mesure utilisées.

CHAPITRE 5 : LE MARCHÉ DE LA MESURE DE LA RÉPUTATION EN LIGNE

Introduction du chapitre 5

L'e-réputation fait l'objet d'un marché foisonnant et a connu un essor à partir des années 2009 - 2010. On assiste à partir de ces années, à la naissance d'une multitude de prestataires de services proposant une offre multiple et variée : " *e-réputation, veille marketing, veille digitale, référencement, community management, communication interactive, veille en intelligence économique, buzz tracking, brand monitoring, marketing communautaire, webmarketing, marketing digital, sentiment analysis, social analytics, relation blogueurs, e-influence, sentiment monitoring, veille image, engagement dans la conversation web, réputation digitale, online sentiment, mesure de la visibilité, RP 2.0, suivi de notoriété, buzz monitoring, identité numérique, etc.*", Boullier et Lohard[287].

S'intéresser au marché de la réputation suppose de proposer une cartographie des acteurs qui évoluent en son sein d'une part, mais aussi, et surtout d'identifier l'ensemble des services et/ ou prestations proposés. Pour ce faire, la première section de ce chapitre vise à proposer une réflexion sur les mutations qu'a connu le secteur, et notamment via le passage du concept de réputation (au sens traditionnel) à celui d'e-réputation tel que perçu aujourd'hui par les praticiens.

La seconde partie de ce chapitre vise à présenter un état des lieux du marché de l'e-réputation. Le marché de la réputation, nous le verrons, est construit autour d'une offre et des prestataires divers et variés. Nous proposerons à l'issue de l'analyse exploratoire du marché, une cartographie des acteurs, et une liste de l'ensemble des derniers que vous retrouverez en annexe 2. À partir de cette cartographie de l'ensemble des acteurs qui évoluent dans le marché et d'une identification des services proposés par ces derniers, nous tenterons de répondre à notre problématique principale (nous y reviendrons dans le chapitre 7) qui vise à savoir si les professionnels se sont appuyés dans les méthodes de mesure, sur des modèles théoriques.

---

[287] Boullier, Dominique, et Audrey Lohard. *Opinion mining et Sentiment analysis: Méthodes et outils*. OpenEdition Press, 2012. https://doi.org/10.4000/books.oep.198. En ligne https://books.openedition.org/oep/204

1. L'e-réputation : un marché en pleine mutation

Si on revient à notre postulat selon lequel les questions de réputation et d'e-réputation font référence au même objet, il n'en demeure pas moins qu'il existe un écart en termes méthodologiques entre les pratiques actuelles et celles utilisées auparavant, écart qui recouvre trois principaux aspects : les techniques et méthodes de collecte des données, les techniques et méthodes de mesure, et les divers objets ou attributs sur lesquels porte la mesure.

1.1. Les techniques et méthodes de collecte des données (les données utilisées) : échantillon versus quantité exponentielle de données

L'un des concepts clés qui a mué et sur lequel il convient de revenir dans une première approche est notamment lié à la collecte des données utilisées dans la mesure de la réputation. À l'origine, les questions de réputation s'appuyaient sur les baromètres d'image et de notoriété. Réalisé de manière périodique (mensuelle, trimestrielle ou annuelle), le baromètre visait à évaluer la perception en termes d'image (image perçue) d'une entité (organisation, entreprise, marque ou produit/ service), de sa notoriété et l'évolution de celle-ci (l'image) auprès de ses publics cibles.

Pour ce faire, et selon l'objectif visé, la mise en place d'un baromètre répondait à un certain nombre de problématiques : déterminer la perception d'une entité auprès de son public cible ou de manière plus générale au sein de son marché ou environnement concurrentiel, vérifier l'adéquation entre l'image voulue (la manière dont l'entité souhaite être perçue) et celle perçue (l'image réelle véhiculée par ses actions), évaluer la notoriété d'une entité, vérifier l'efficacité des actions déployées (marketing, publicitaire et de communication), mesurer la satisfaction des clients et consommateurs (mais aussi leur fidélité et leur aptitude à recommander l'organisation ou le produit/ service), identifier les potentiels risques auxquels s'expose l'entité, etc. Les objectifs d'un baromètre d'image et de notoriété sont vastes et peuvent se résumer en deux principales catégories avec une double visée, soit dans une première approche de mesurer l'évolution de l'image et/ou de la notoriété d'une entité comparativement à celle perçue lors des précédents baromètres à partir des diverses actions

(de communication, publicitaire ou marketing) déployées par celle-ci et de la satisfaction client, ou de réaliser une analyse comparative (sous forme de classement par exemple) des organisations évoluant dans un secteur d'activité donné (on parlera alors de baromètres d'image sectoriel).

Lorsqu'on s'intéresse à la mesure de l'image d'une entité et donc à la manière dont celle-ci est perçue par ses publics cibles, il s'agit entre autres de déterminer les diverses "*associations d'idées*" (*Corporate association* au sens de Brown, 1998, p.217) qui lui sont accolées, nécessitant une approche qualitative pluridimensionnelle et axée sur : les attributs cognitifs, affectifs et conatifs, les associations d'idées ou de marque relevant autant d'éléments cognitifs (c'est-à-dire d'expérience personnelle avec l'entité), mais aussi et surtout affectifs (via la publicité ou le bouche-à-oreille par exemple). Pendant des décennies la recherche sur la prise de décision des consommateurs était de nature principalement cognitive via notamment les attributs et éléments de marque tangibles (Malhotra, 2005) et la mesure de l'image de marque utilisait plutôt l'évaluation cognitive comme approche la plus courante (Keller, 2003, p.596). Ces deux dernières décennies ont vu naître une reconnaissance des aspects affectifs et émotionnels dans le comportement des consommateurs (Burk et Edell, 1989 ; Holbrook et Westwood, 1989, Aaker et al., 2001, Chun et Davis, 2002) avec une prise en compte des deux aspects (cognitifs et affectifs) comme importants et indispensables dans l'évaluation d'une organisation (Brown, 19989, Keller, 2003, Argawal & Malhotra, 2005), dimensions complétées par une approche conative (soit l'aptitude via notamment la satisfaction, à recommander ou non une entité).

En terme méthodologique, la collecte des données utilisait avant le développement des communications numériques, des approches par enquête auprès d'un échantillon représentatif du public cible, enquête qui selon l'objectif visé pouvait prendre deux principales formes : là où la détection d'attributs identitaires relatifs à l'image perçue utilise plutôt une approche qualitative (entretiens téléphoniques, interviews individuels ou de groupe par exemple), la mesure de la réputation globale d'une entité par rapport à son environnement concurrentiel et la mesure de sa notoriété (spontanée, assistée et *top of mind*) utilisent des approches quantitatives (questionnaires administrés en face à face, auto administrés en ligne[288], par voie postale ou téléphonique par exemple).

---

[288] via notamment le système CAWI (computer assistance for web interview)

Aujourd'hui, la mesure de l'e-réputation construite à partir des mêmes marqueurs utilise plutôt des méthodes dématérialisées, ce qui signifie que la collecte des données s'effectue la plupart du temps (même si certaines enseignes réalisent encore des bilans en face à face[289]) via deux principales méthodes : l'e-mailing et le *crawl*. Si la méthode du *crawl* est utilisée par tous les acteurs du marché dans la collecte des données (nous y reviendrons par la suite), d'autres approches viennent pour certains, en complément de celle-ci. C'est notamment le cas de la stratégie d'emailing. En effet, à l'issue d'une interaction avec une entité ou afin de l'évaluer sans interaction préalable, les données peuvent être collectées via un emailing à partir de deux principaux acteurs : soit en interne par l'entité par l'intermédiaire notamment de son service client (dans le cadre d'un achat) ou, dans une seconde mesure et de plus en plus aujourd'hui, via des acteurs externes, les sites d'avis ou de recommandation. Sur la mesure en interne, le processus qui est simple peut prendre plusieurs formes (enquêtes sommaires via notamment l'utilisation d'une échelle multiple ou sondages plus complets) selon l'objectif visé par la collecte des données.

Dans la mesure où l'entité ne serait pas en mesure de mettre en place sa propre stratégie de collecte des données, elle peut le déléguer à des organismes spécialisés dans le domaine. S'il est un acteur principal dans la littérature scientifique qui est à l'origine des questions de mesure en e-réputation, c'est bien les sites de recommandation (recommander) utilisés notamment par les acteurs du e-commerce, qui font partie des premiers à développer des modèles de mesure aux côtés des moteurs de recherche[290]. À l'image du processus de collecte utilisé par les services internes, les sites de recommandation fonctionnent en certains points sur le même principe avec quelques éléments de distinction selon leur structure.

---

[289] Cette approche est aujourd'hui très rare et ne se présente pas réellement sous forme d'enquête telle qu'utilisée dans la mesure de la réputation mais plutôt d'avis reprenant 2-3 questions et en complément des enquêtes en ligne, et s'effectue de manière générale dans les enseignes physiques ou dans la rue et parfois de manière implicite (venez-vous souvent dans le magasin ? Quels types de produits vous intéressent ?, etc.).

[290] Se référer à la section suivante sur les agences et le marché (la cartographie des acteurs du marché de la réputation)

En effet, si sur les plateformes décentralisées la collecte des avis est relativement simple (c'est notamment le cas de Google My Business[291] ou encore Trustpilot par exemple), certaines approches hybrides elles, à l'image de celles centralisées (comme la plateforme Net reviews ou avis vérifiés par exemple) proposent une autre approche. Les plateformes centralisées disposent d'un CRM qui permet, à l'issue d'une interaction avec une entité et via notamment l'envoi d'un mail d'invitation, de collecter les informations qui serviront à déterminer la réputation d'une entité.

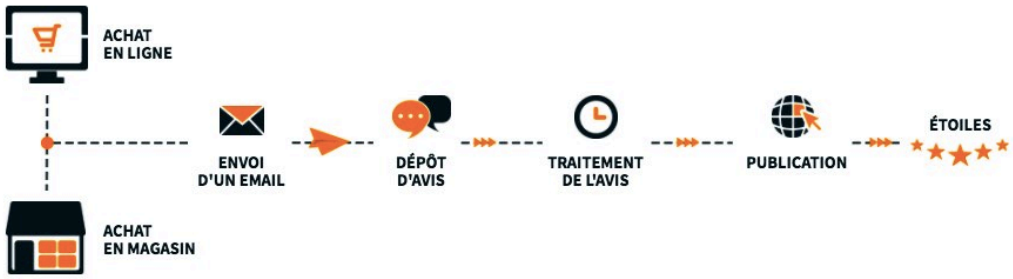

Figure n°9 : Processus de collecte des avis vérifiés[292], source Avis vérifiés

Les avis, une fois collectés, peuvent être mis en avant sur les plateformes corporate (sites web par exemple) ou encore, dans une approche classique, apparaître dans les résultats de recherche (organiques ou sponsorisés)[293]. Il est important de souligner que la collecte des avis et de manière générale la réputation d'une entité ne repose pas uniquement sur ses produits ou services, mais aussi l'ensemble des individus qui la compose, la réputation de ces derniers ayant forcément un impact sur celle de l'entité. En sus donc des données de réputation collectées sur l'objet de mesure, les entités s'intéressent de plus en plus à la réputation de ses "*salariés*" ou de son personnel, et des interactions qu'ils entretiennent avec les clients ou consommateurs.

---

[291] Où il suffit dès lors que l'entité dispose d'une page GMB d'y accéder via les moteurs de recherche et déposer un avis

[292] Processus sur lequel s'appuient les plateformes centralisées. Celles décentralisées en outre, le processus commence à la phase 3

[293] La mise en place des avis Google sur un site web par exemple, il suffit de créer un compte Google merchant center actif, cliquer sur Programmes merchant center, activer l'onglet avis clients puis suivre les principales phases afin d'intégrer le script. https://www.joptimisemonsite.fr/comment-recuperer-google-avis-client/

Si les stratégies d'e mailing propulsées notamment par les tiers de confiance sont une alternative à la collecte des données de mesure de la réputation, la principale méthode utilisée aujourd'hui s'appuie essentiellement sur des *crawlers* du web.

Qu'est-ce qu'un *crawler* ? « *Machine à échantillonner sur une méthode de proche en proche* »[294], le *crawler* ou robot d'exploration est programmé via notamment des scripts, afin d'exécuter un certain nombre de tâches de manière automatique et autonome. Le principe est simple : à partir des instructions ou fonctions préalablement assignées, le crawler va parcourir le web à la recherche d'informations qu'il va analyser puis enregistrer dans des bases de données, données qui par la suite feront l'objet d'indexation. Plus concrètement, en amont du crawl des données, des fonctions sont assignées au crawler avec notamment une liste initiale d'*URLs* appelée "*seed*" à explorer. À partir de cet ensemble de *seed* construit via notamment les indexations précédentes, les bases de données, les plans des sites, etc., et par l'utilisation des liens hypertextes, les robots recherchent de nouveaux documents à indexer, processus itératif qui se poursuit et se renouvelle au fil des navigations et au fur et à mesure de l'apparition de nouveaux documents et liens jusqu'à ce que les robots rencontrent des boucles, pages d'erreurs, liens morts, etc.

Si l'indexation de documents web à direction des moteurs de recherche est l'application la plus répandue des robots d'exploration (avec notamment comme *search bot* les plus connus : Googlebot, Bingbot, Alexa crawler), il n'en demeure pas moins qu'ils peuvent également être utilisés dans divers autres domaines comme par exemple le *SEO ou* l'analyse concurrentielle. À l'image des *crawlers* des moteurs de recherche, les solutions de *listening* et de veille en e-réputation fonctionnent sur le même principe. Connectés à plusieurs sources différentes (moteurs de recherche, presse en ligne, blogs, forums, sites d'avis, bases de données, *API* des réseaux sociaux numériques, etc.), les solutions, à l'aide des moteurs de recherche intégrés à leur plateforme et par l'utilisation notamment de requêtes (comme précédemment mentionné), vont sillonner le web à la recherche de métadonnées qui se réfèrent aux clés

---

294 Lakel, A. (2021). *My web intelligence* : un outil pour l'analyse du web et des réseaux. *I2D - Information, données & documents*, 1, 96-103. https://doi-org.ezproxy.u-bordeaux-montaigne.fr/10.3917/i2d.211.0096

d'entrée[295]. Dès lors que des contenus pertinents sont identifiés, « *les liens hypertextes sont explorés pour récupérer les documents cités* »[296], puis les données sont nettoyées, qualifiées, classées et catégorisées afin d'être analysées.

Si les premières approches (celles d'e mailing) sont une méthode d'enquête numérisée et donc la numérisation de la 1ère ère, la méthode du crawl, et il est important de le souligner, marque une rupture entre les deux ères. En effet, cette dernière approche met l'accent sur les phénomènes d'urgence, de surinformation, de surcharge informationnelle ou *infobésité,* et annihile les questions de segmentation en amont des besoins et des informations. On *crawl* désormais l'ensemble des contenus disponibles sur un objet donné sans questionnements préalables[297] et on collecte *a posteriori* les avis, les commentaires, et parfois des contenus qui ne sont pas des avis à proprement parler. C'est notamment le cas des crawls réseaux sociaux (*social media monitoring*) ou des crawls du web 1.0 proposés par des outils comme My Web Intelligence[298] ou Hyphe[299].

---

[295] En effet, dès lors qu'une requête est émise, les robots reçoivent un ensemble d'*URLs* qui correspondent à la fonction d'entrée demandée. Le contenu de ces *URLs* est par la suite analysé avec pour finalité de déterminer leur pertinence : si le contenu est pertinent, les métadonnées et *URL*s sont sauvegardées puis indexées dans des bases de données, et, dans la mesure où elles ne le seraient pas, elles sont marquées de la mention "vue" afin de prévenir les robots qu'ils ont déjà fait l'objet d'une analyse et éviter de les *crawler* de nouveau. Le contenu considéré comme pertinent est enfin envoyé vers un extracteur de liens en charge de collecter l'ensemble des hypertextes (issus du *SEO on page* : maillage interne et *off page* via le *netlinking* par exemple), liens qui seront de nouveau analysés en suivant le même procédé qui se renouvellera dès l'apparition de nouveaux documents.

[296] Lakel, A. (2021). *My web intelligence* : un outil pour l'analyse du web et des réseaux. *I2D - Information, données & documents*, 1, 96-103. https://doi-org.ezproxy.u-bordeaux-montaigne.fr/10.3917/i2d.211.0096

[297] Même si la mise en place d'une stratégie de *social listening* nécessite néanmoins en amont un cadrage général avec des requêtes spécifiques relatives à l'objet de mesure, il est important de souligner que ces dernières (les requêtes doivent être assez larges comme celles de longue traîne ou substituts par exemple) afin d'obtenir une certaine exhaustivité. Les données pourront par la suite être filtrées et catégorisées selon les besoins et finalités de l'étude.

[298] Qui est un outil d'extraction (*crawl*), d'archivage, de qualification et de visualisation du Web au service des *digital methods*. L'objectif visé par l'outil est « *de fournir, à tous les experts et chercheurs qui souhaitent développer des études dans le domaine de l'intelligence numérique et des humanités digitales, un dispositif basé sur l'analyse des prises de parole en ligne* », Lakel (2021). Lien de téléchargement https://github.com/MyWebIntelligence

[299] Qui est un outil de curation de corpus web doté d'un robot d'exploration Web axé sur la recherche. Comme ils le soulignent « *Dans Hyphe, un corpus Web est construit comme un ensemble de pages Web organisées et organisées par l'utilisateur. L'interface utilisateur de Hyphe vous guide à chaque étape de la méthode de conservation du corpus Web et vous aide à surveiller votre collecte* ». Accessible en ligne sur https://hyphe.medialab.sciences-po.fr

1.2. Les techniques et méthodes de mesure : approche traditionnelle versus algorithmique

Afin d'appréhender la mutation des techniques de mesure, il est important de remonter à l'essence du concept de réputation tel qu'envisagé par les chercheurs. Si on reprend les définitions théoriques de la réputation dans la littérature, partant d'un déficit ou manque d'informations sur une entité, elle est conçue comme construit social, comme la perception d'un acteur sur une entité quelle qu'elle soit (au sens d'individualité, organisation, produit ou service), perception fondée et construite à partir des divers attributs (au sens de Fombrun, 1996, 1997) de l'entité en question, mais aussi et surtout des actions ou interactions passées entre les deux entités et prenant la forme d'une évaluation des traces (Origgi, 2013), de l'opinion (D'Almeida, 2017) et des signaux dits informationnels (Alloing, 2015) et visant entre autres à réduire les incertitudes, à établir ou maintenir une relation de confiance, somme toute à garantir une certaine forme de contrat social implicite (Cornell et Shapiro, 1982).

La notion de "garantie de contrat" suffit à elle seule à expliciter et étayer cette mutation dans la technicité. Pour les théoriciens de la réputation et plus particulièrement à travers la théorie des jeux, la réputation d'une entité (au sens d'organisation) qui peut être bonne ou mauvaise, se fonde sur les actions passées de celle-ci lorsqu'elle interagit avec ses diverses parties prenantes, actions antérieures qui permettent de prédire, d'anticiper celles à venir, au même titre que celles de ses divers publics. Dans l'optique où la perception des parties prenantes sur l'entité serait positive, cette dernière (l'entité) se voit dans l'obligation implicite de maintenir une forme de constance ou d'adapter ses futures interactions afin de conserver sa réputation et éviter de rompre le contrat implicite « *interaction réussie = bonne réputation* ». *A contrario*, et dès lors qu'il existerait un risque dans l'établissement de la relation, l'entité se verrait attribuer une mauvaise réputation. Sur la détection du risque, tout se résume ici au concept de prédiction. À partir d'un nombre minimum d'interactions et de connaissance des différents acteurs (entité qui évalue et celle évaluée), il est possible de prédire les différentes possibilités de jeu (au sens d'actions futures) de la part des diverses parties, prédiction qui est le fondement même et à l'origine des travaux en intelligence artificielle (IA) et notamment en apprentissage automatique ou *Machine learning* (ML).

Sous discipline de l'IA, le *ML* est un domaine scientifique qui vise, à partir d'algorithmes, à identifier des "récurrences ou répétitions" (patterns) dans un flux de données, en vue de mettre en place des analyses prédictives. Ces algorithmes sont entraînés via un processus itératif, à effectuer (de manière autonome) à partir d'un jeu de données préalablement organisé et nettoyé, des tâches, ce qui permet au fil des apprentissages, d'améliorer leurs performances jusqu'à atteindre une autonomie complète dans le but d'être déployés ou appliqués à un nouveau jeu de données. Là où les approches traditionnelles présentaient des limites eu égard au flux constant de données disponibles, le *ML* offre, la possibilité de fournir des *insights* à partir de vastes ensembles de données, d'où la multiplication de cas d'applications : moteurs de recherche généralistes (Pagerank ou complément de requête Google par exemple), fil d'actualité des plateformes sociales numériques, assistants numériques, objets connectés, moteurs de recommandation. Le champ des possibles de l'application des techniques de *ML* est illimité et ne cesse de croître avec le développement des techniques d'IA. Le maître mot dans l'utilisation d'approches de *ML* est la prédiction à partir de l'analyse de données. Plus concrètement, comment cela fonctionne-t-il ? Il est tout d'abord important de souligner qu'il existe plusieurs typologies d'apprentissage machine selon l'objectif de l'analyse, et qu'on peut regrouper selon trois (3) approches (avec pour chacune une diversité d'algorithmes à déployer) : l'apprentissage supervisé[300], l'apprentissage non supervisé et par renforcement.[301] Dès lors qu'on définit la finalité souhaitée (modélisation prédictive dans le premier cas, apprentissage des caractéristiques dans le second ou en utilisant l'approche par renforcement), il s'agit de développer le modèle de *ML* qui sera utilisé dans la résolution du problème.

Adaptés à la mesure de la réputation, les algorithmes de *ML* développés via les suites logicielles permettent, à partir des métadonnées *crawlées* sur le web social et relatives à un objet d'étude et se référant à un ensemble de requêtes prédéfinies[302], de procéder à la classification du corpus dans les divers onglets ou *widgets* des tableaux de bord, afin d'extraire les informations essentielles qui seront par la suite interprétées.

---

[300] Technologie la plus utilisée

[301] Accessible via les sites Datascientest et Talend https://datascientest.com/machine-learning-tout-savoir ; https://www.talend.com/fr/resources/what-is-machine-learning/

[302] Se référer à la section 3 sur les services proposés (le *social listening*)

1.3. Des problématiques liées à la réputation : l'objet sur lequel porte la mesure (image versus totalité des attributs)

Un des concepts clés qui a également mué dans la mesure de la réputation aujourd'hui est notamment lié aux problématiques sur lesquelles porte la mesure. En effet, à l'origine, les questions de réputation s'articulaient autour d'un concept : l'image perçue par les divers clients ou consommateurs d'une entité notamment le concept de *corporate reputation* qui en a découlé, et la représentation de cette image au sein du secteur d'activité dans lequel l'organisation évolue (on parlera alors de notoriété : spontanée, assistée ou t*op of mind*).

Pour ce faire, et à partir de baromètres d'images comme nous l'avons vu à travers les techniques et méthodes de collecte et de mesure, les entreprises de sondages (qui étaient alors en charge de la gestion des questions d'image et de réputation : Ifop, Ipsos, Kantar, etc.) effectuaient des enquêtes auprès des divers publics, afin de déterminer leur perception de l'image véhiculée par l'entité (comparativement à celle voulue). On mesurait la connaissance générale de l'entité et sa position au sein de son environnement concurrentiel (aspect cognitif), les aspects culturels et liés aux valeurs (dimension affective visant à créer du lien et développer l'affect envers l'entité) et l'aspect conatif qui en découle soit la fidélité et l'aptitude à recommander (ou interagir de nouveau avec celle-ci) lié notamment à la satisfaction client.

Si la réputation "traditionnelle" s'appuyait de fait sur le concept d'image corporate de l'entité et ses attributs, la réputation telle qu'imaginée et perçue aujourd'hui revêt un aspect holistique, avec une approche de mesure portant sur une pluralité d'aspects et qui répond à plusieurs nouvelles problématiques apparues notamment avec l'essor du web social. Au compte de ces dernières, on retrouve : les questions d'indexation et de *ranking* (notamment de positionnement dans les résultats de recherche), d'influence (avec entre autres la cartographie d'acteurs), de *personal branding* (ou marque personnelle avec notamment les principaux acteurs qui représentent une entité), et les questions légales (notamment en cas de diffamation, d'atteinte à la vie privée, de droit à l'oubli)[303].

Ce premier cadre relatif à la mutation de l'e-réputation, posé, il convient dès lors de s'intéresser à l'ensemble des acteurs qui composent le marché de la réputation en ligne.

---

[303] Se référer à la section suivante sur les offres proposées dans le domaine légal

2. Les agences et le marché : un marché français en pleine expansion

Plusieurs auteurs se sont attelés à proposer une cartographie des différents acteurs évoluant dans le marché de l'e-réputation. Déjà en 2009, le site "demain la veille" et notamment Aref Jdey proposait une cartographie regroupant plus d'une centaine de prestataires selon quatre (4) catégories : les éditeurs d'applications, les *pure players*, les prestataires de service et les applications gratuites.



Figure n°10 : The e reputation market map Aref JDEY, source : site demain la veille[304]

[304] Cartographie créée avec les infostratèges https://www.les-infostrateges.com/actu/0911842/une-cartographie-du-marche-de-l-e-reputation

Peggy Cadel (2010)[305] regroupe les acteurs de l'e-réputation selon d'une part leur domaine de compétences (3 typologies) : les *pure players*, les acteurs historiques de la veille, les agences de communication/ marketing/ médias ; et d'autre part selon les caractéristiques de leurs offres (3 typologies également) : les éditeurs d'outils, les prestataires de service et les solutions mixtes (qui englobent à la fois les outils et prestataires de service). Camille Alloing lui, proposait en 2015 sur son blog, une liste de 140 acteurs[306] évoluant dans le domaine de la veille et de la gestion en e-réputation. À partir de ses travaux (et notamment sa première version de 2011), l'agence Blueboat proposait en mai 2011 sur son blog[307], une cartographie des agences d'e-réputation en France.

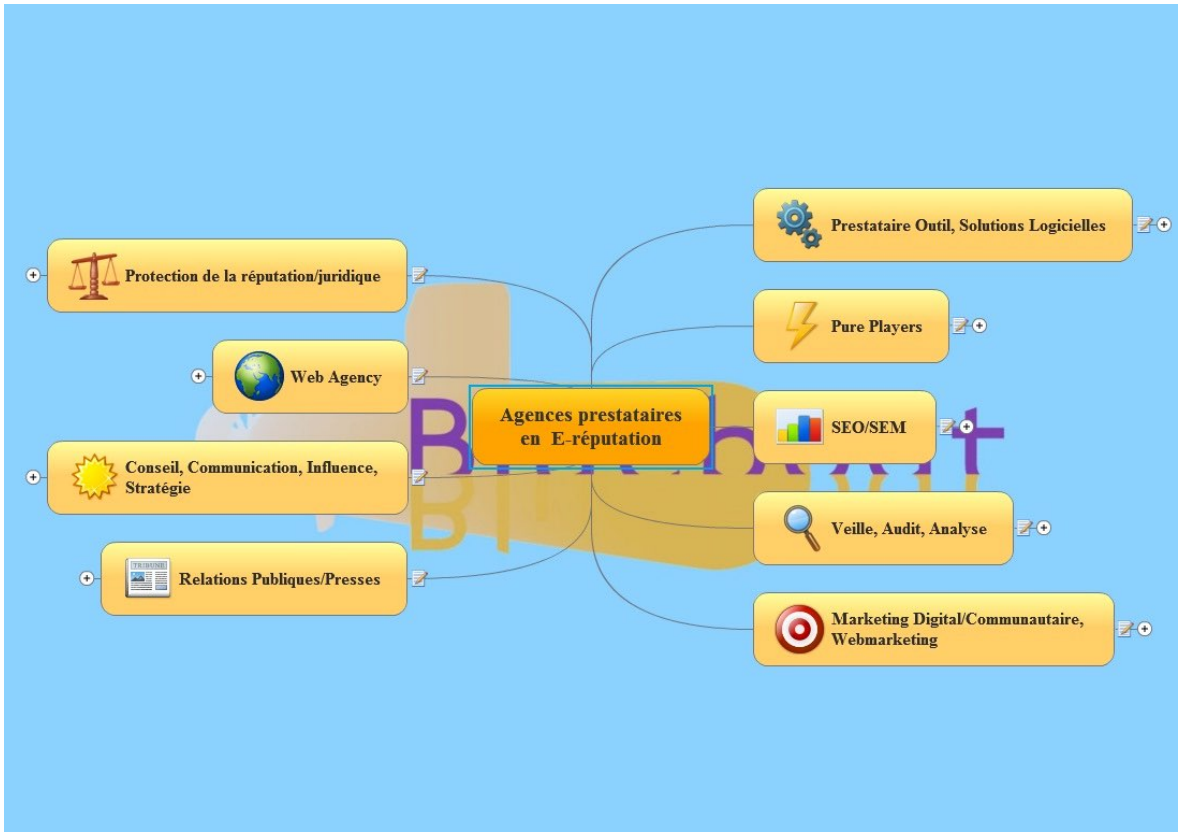

Figure n°11 : Cartographie des acteurs du marché de la réputation, Source : Blueboat

[305] Cadel, Peggy. « Le marché de l'e-réputation ». *Les Cahiers du numérique* Vol. 6, nº 4 (1 juin 2011): 111-21. En ligne http://www.cairn.info/resume.php?ID_ARTICLE=LCN_064_0111.

[306] Liste actualisée et issue de sa première version proposée en 2009 https://cadderep.hypotheses.org/agences-en-e-reputation

[307] Agence spécialisée dans l'e-réputation. Cartographie qu'on peut retrouver en ligne sur leur blog http://www.blueboat.fr/agences-e-reputation-france/

L'agence, à travers sa carte heuristique, entend proposer une classification par domaine d'activité des différents acteurs qui composent l'écosystème. On distingue donc à partir de leur carte, 123 prestataires répartis en 9 catégories[308] d'agences que sont les prestataires d'outil/ solutions logicielles, les *pure players*, les agences SEO/SEM[309], les agences spécialisées dans la veille/ l'audit / l'analyse, les agences web marketing, les agences de relations publiques/ presse, les agences conseil en communication / influence / stratégie, les *web agencies* et les agences spécialisées dans la protection de la réputation juridique.

Afin de proposer une taxonomie des acteurs évoluant dans le marché de la mesure de la réputation en ligne, nous nous sommes appuyée sur l'ensemble de ces travaux précédemment réalisés, en discutant pour chacune des taxonomies, les concepts qui nous paraissent essentiels. Il est essentiel de noter que nous ne discutons pas les choix méthodologiques de ces derniers, mais plutôt des dimensions ou attributs constitutifs de leur taxonomie. Dans la proposition de classification d'Aref (2009) qui selon nous est incomplète, nous discuterons trois (3) des quatre (4) catégories utilisées. En effet, dans sa proposition, les éditeurs logiciels, les *pure players* et applications gratuites nous semblent essentiels comme attributs. Cependant, l'attribut prestataires nous semble trop vaste, ces derniers pouvant être recoupés en sous-ensembles.

C'est notamment ce qu'a tenté de faire Peggy Cadel (2010) en proposant une distinction entre ce qu'elle qualifie *d'acteurs historiques de la veille* et les *agences* (de communication, marketing et médias) dans sa première typologie et dans la seconde une distinction entre les éditeurs d'outils ; les prestataires de service et les solutions mixtes (qui englobent à la fois les outils et prestataires de service). Cette approche soulève également certains questionnements dans la mesure où certaines catégories se voient imbriquées : par exemple, la veille ou le *social listening* étant la phase amont indispensable dans toute stratégie de mesure de la réputation, les *pure players* au même titre que les agences sont en partie des acteurs de la

[308]Les catégories d'acteurs dans le marché de la réputation selon l'agence Blueboat. Accessible en ligne sur https://books.openedition.org/oep/docannexe/image/204/img-2.jpg

[309] Spécialisées dans l'optimisation de la visibilité des entités notamment dans les résultats des recherches

veille; c'est également le cas des solutions mixtes dans lesquels on retrouve les éditeurs d'outils et les prestataires de services.

Pour pallier cette problématique, la cartographie présentée par l'agence Blueboat propose de regrouper les acteurs selon uniquement leur métier. Cette approche peut, de notre point de vue, et ce dû à son exhaustivité (9 attributs), être plus synthétique : en effet, certaines agences SEO/SEM/SEA se revendiquent être des agences de web marketing ; les prestataires d'outils / solutions proposent (via leurs plateformes) de la veille / audit/ analyse et comme déjà mentionné précédemment, la veille/ l'analyse s'avèrent indispensables à toute stratégie de mesure et sont transversales aux *pure players* et agences. Un élément essentiel apparu dans ces travaux est la prise en compte des agences spécialisées dans la protection juridique.

Toutes ces approches nous le voyons, révèlent des attributs intéressants dans la classification des acteurs, mais souffrent, de notre point de vue, d'être trop réductrices pour certaines, ou *a contrario,* trop exhaustives pour d'autres, mais aussi et surtout incomplètes, le marché de l'e-réputation étant encore nouveau et en perpétuelle construction.

Notre proposition s'articule autour de six (6) principales catégories de classification des approches proposées par les acteurs dans la mesure de la réputation en ligne : les *pure players*, les éditeurs de solutions logicielles, les agences ad hoc, les agences généralistes, les applications gratuites (*free tools*) et les tiers de confiance (les *recommander sites* ou sites d'avis).



Figure n°12 : Proposition d'une cartographie des acteurs du marché de la réputation

L'une des catégories, qui par ailleurs est transversale à la quasi-totalité des propositions de classification (2009, 2010, 2011) nous l'avons notée et, qui nous paraît essentielle dans la cartographie, est celle relative aux ***pure players***. Par *pure players*, et de notre point de vue[310], nous entendons les acteurs dont le cœur de métier est la mesure de la réputation en ligne, qui se présentent dans leurs discours et supports corporate comme agence spécialisée en e-réputation, et proposent de fait des services axés uniquement autour de cette question. Il nous apparaît essentiel de distinguer ces acteurs eu égard au caractère intangible du concept d'e-réputation qui, nous l'avons rappelé, est un concept polymorphe faisant appel, nous avons pu le constater, à une approche pluridisciplinaire avec une pluralité de services renvoyant au même concept. L'objectif de cette catégorie est donc d'identifier les acteurs dont l'offre

[310] Indépendant, nous tenons à le rappeler, des perceptions des autres acteurs ayant proposé cette catégorie, mais aussi et surtout de sa signification dans les lexiques

s'articule intrinsèquement autour de cette problématique de mesure, d'évaluation, et plus largement de gestion de l'e-réputation. Au compte des *pure players*[311], on peut citer entre autres, les agences Alterbuzz, Blueboat, E réputation paris, Netino, Opinion Act, QSN Digital, Reputation net, Sos e réputation.



Figure n°13 : Exemple de pure players en e-réputation

La seconde catégorie est relative à celle des **éditeurs de solutions logicielles**. La mesure de la réputation en ligne de par son approche pluridisciplinaire fait appel autant à une analyse d'une entité et de son environnement concurrentiel (qualifiée selon les usages de veille, *social listening*, audit), à la détection de potentiels relais ou profils d'influence (*hub, writers, spreaders*) en passant par des stratégies de mise en avant de pages sur les diverses plateformes socionumériques (gestion, modération de contenus), au suivi des mentions sur une entité, etc.

311 Vous retrouverez la liste complète au même titre que la catégorisation des agences généralistes, ad hoc, éditeurs, applications gratuites et tiers de confiance en Annexe n°2

Afin d'y parvenir, plusieurs solutions existent sur le marché selon l'usage, la problématique, et l'objectif visé dans la démarche. On retrouve donc dans cette catégorie, les acteurs qui ont développé une solution visant à appréhender une ou plusieurs des métriques de l'e-réputation (veille en e-réputation ou *social listening*, *social media monitoring*, cartographie d'acteurs). Il est intéressant de noter qu'au sein de ces acteurs, plusieurs proposent une prestation d'accompagnement et/ ou de conseil (comme il peut être le cas pour les *pure players*). C'est notamment le cas de Digimind, Linkfluence, Talkwalker, Sindup, KBCrawl par exemple. Le choix de les identifier dans cette catégorie et non comme *pure players* est tout simplement dû à l'objet premier de leur renommée[312] et à la perception des acteurs[313] qui utilisent les solutions ou font appel à leurs services.



Figure n°14 : Exemple d'éditeurs de solutions en e-réputation

[312] Lors de nos entretiens par exemple, nous avons demandé aux éditeurs dans quelle catégorie ils se situent au sein de notre cartographie

[313] Il s'agit des agences qui utilisent les solutions logicielles dans leurs stratégies de gestion de l'e-réputation

À côté de ces derniers, certains chercheurs ont développé des solutions en *open source* et donc gratuites, qu'on retrouve dans notre catégorie des **applications gratuites** (*free tools*). On serait tenté, *a priori*, d'intégrer ces derniers dans la classification précédente. Cependant, et à la différence des solutions logicielles en *Saas* des éditeurs qui proposent un accompagnement à la prise en main et pour certains une prestation de service (Linkfluence, Talkwalker, Sindup, Synthesio, Mention, Digimind par exemple), aucune prestation d'accompagnement à la prise en main ou de conseil sur les outils et leur utilisation n'est proposée. De plus, ces outils permettent de gérer, et sont construits autour d'une des problématiques de l'e-réputation, et non une offre générale allant de la veille sur les diverses plateformes sociales à la mise en place de stratégies de gestion, d'où notre choix de les identifier comme catégorie singulière. Au compte des applications gratuites, on retrouve à titre d'exemples, Crowdtangle qui permet de t*racker ou suivre* le partage des liens sur les réseaux sociaux, Netvibes[314], Social Searcher[315], des outils adaptés au micro blogging Twitter comme Social bearing ou Twazzup ou encore Hyphe[316] et My Web Intelligence[317].

Une importante partie des acteurs évoluant dans le marché de l'e-réputation propose une offre généraliste au sein de laquelle on retrouve une prestation de mesure de l'e-réputation. Par **agences généralistes**, nous faisons donc référence aux agences *webmarketing*, de communication digitale, d'influence, de relations publiques/presse et médias dont la mesure de la réputation n'est pas le coeur de métier, mais qui, de par leur offre générique et holistique, disposent d'un module de gestion de la réputation en ligne. Cette catégorie, il est important de le noter, recouvre la majorité des prestataires qui composent le marché de la réputation.

---

[314] Qui permet de faire de la veille en e-réputation. Accessible sur https://www.netvibes.com/fr

[315] Qui permet également de faire de la veille en e-réputation. Accessible sur https://www.social-searcher.com

[316] Accessible en ligne sur http://hyphe.medialab.sciences-po.fr/

[317] Lien de téléchargement https://github.com/MyWebIntelligence



Figure n°15 : Exemple d'agences généralistes en e-réputation

À côté de ce vaste univers d'acteurs, certains (une minorité) proposent, en lieu et place d'une mesure globale de la réputation, une offre spécifique contribuant à la gestion de la réputation, mais portant sur un attribut singulier. C'est notamment le cas des **agences ad hoc** (*ad hoc agencies*) qui proposent soit un recours légal en cas de diffamation, de contribuer à la construction de l'e-réputation notamment dans les résultats des moteurs de recherche, ou encore de nettoyer / supprimer / noyer les contenus négatifs. Ces agences, grâce notamment à l'essor des questions de crises numériques, de bad buzz, de diffamation, de droit à l'oubli, offrent une prestation ciblée en réponse à des besoins spécifiques. On retrouve notamment dans cette catégorie des cabinets d'avocats qui proposent une approche légale en cas d'atteinte à la réputation (Haas Avocats, Porlon Avocats, Alain Bensoussan, Picovschi par exemple), des agences spécialisées dans le domaine de la santé (Healthcie ou Ma réputation en ligne par exemple), les agences spécialisées sur les questions de déréférencement ou suppression de contenus négatifs (Netwash, Riposte, Netoffensive, etc.).



Figure n°16 : Exemple d'agences ad hoc en e-réputation

Un des acteurs qui est le parent pauvre (inexistant dans les autres taxonomies) et non réellement référencé comme tel dans les résultats de recherche avec les requêtes utilisées pour notre taxonomie, et qui par ailleurs fait paradoxalement partie des acteurs majeurs et connaît de plus de plus d'engouement, est l'ensemble de ceux qu'on retrouve dans les sites de recommandation ou *recommander sites* (sites d'avis) considérés comme **tiers de confiance**. Dans une approche générale, on distingue trois (3) catégories de tiers de confiance selon leur structure : les plateformes décentralisées ou ouvertes qui offrent la possibilité à tout internaute de laisser un avis sur n'importe quelle entité (même sans interaction avec celle-ci et au sein lesquelles on retrouve entre autres *Google my business* par exemple), les sites centralisés construits autour de la satisfaction client qui collectent des données réelles issues d'interactions (Avis vérifiés[318], Custplace[319], Ekomi[320], Trusted shops[321], par exemple) et, à

[318] Exemple de *recommander site* avec une structure centralisée https://www.netreviews.com/fr/

[319] Exemple de *recommander site* avec une structure centralisée https://fr.custplace.com

[320] Exemple de *recommander site* avec une structure centralisée https://www.ekomi.fr/fr/

[321] Exemple de *recommander site* avec une structure centralisée https://www.trustedshops.fr

l'intersection de ces derniers les sites hybrides, qui utilisent une approche décentralisée avec la possibilité de collecter des données vérifiées. Il est important de souligner que dans la littérature scientifique sur la mesure de la réputation, les premiers modèles étaient construits autour des approches e-commerce (et d'influence via notamment les moteurs de recherche) et celles des sites de recommandation ou *recommender sites*. Il nous semble donc évident et essentiel d'intégrer les acteurs en charge de la collecte et de la mise à disposition des avis, importants selon nous[322], dans la mesure où certains proposent en sus, des services de modération de contenu ou de gestion de la réputation[323]. Comme mentionné, à partir des mots clés de recherche que nous avons choisis, ces derniers ne sont pas référencés comme acteurs à l'image des éditeurs, agences ou *pure players*. Cependant, au vu de notre approche théorique précédemment évoquée dans le chapitre 4 et de notre approche terrain (les entrevues réalisées), il ressort qu'ils deviennent aujourd'hui incontournables dans le marché de la réputation, avec une approche de collecte des données certes "simpliste", mais visant toutefois, à proposer les *inputs* nécessaires à l'évaluation d'un objet, d'où leur intégration comme acteur dans notre proposition de cartographie[324].

À partir de ce premier cadrage, notamment les six (6) principales catégories supra-citées, et afin de définir une cartographie de l'ensemble des acteurs qui composent le marché de la réputation, il nous semble intéressant de nous appuyer sur l'ensemble des prestations proposées dans l'offre de mesure[325].

---

[322] Comme souligné dans le chapitre précédent sur la modélisation théorique de la réputation

[323] C'est notamment le cas d'Avis vérifiés par exemple via sa solution Review management+ https://www.netreviews.com/fr/solutions/reputation-builder/

[324] Même si nous avons tenu à les intégrer à la cartographie des acteurs, il est important de souligner que eu égard au fait qu'ils ne sont pas associés au même titre que les autres à nos requêtes, ils n'ont pas fait l'objet d'un travail de recherche exhaustif comme il peut être le cas pour les autres acteurs et feront de fait l'objet de perspectives dans nos futurs travaux

[325] Et dans une vision plus large, dans l'offre de gestion et de construction de l'e-réputation dans les environnements numériques

3. Les services proposés : une offre foisonnante et sur mesure

L'état de l'art présenté dans la section précédente nous a permis d'avoir une vision générale du marché de la réputation en ligne qui, comme nous avons pu le constater, est foisonnant et en perpétuelle croissance et construction. Que la mesure de la réputation soit le coeur de métier (comme cela peut être le cas pour les *pure players*) ou une des activités dérivées (pour les agences généralistes entre autres), les prestataires de service rivalisent d'ingéniosité et proposent des prestations aussi diverses que variées, mais complémentaires, tant sur du *consulting* que de la formation.

Ces offres peuvent être regroupées selon deux principales catégories : l'offre de mesure qui relève de l'intelligence économique, de la veille, de l'audit ou du *social listening* et sur laquelle nous nous attarderons plus particulièrement, et dans une approche plus générale l'offre allant de la construction à la gestion de la réputation qui est indissociable de la première. Même si cette dernière offre n'est pas l'objet central de nos travaux, nous proposerons un bref retour sur quelques unes des techniques de construction et de gestion de la réputation en ligne (en particulier sur les approches émergentes et les moins connues du grand public), ces techniques étant *in fine* la résultante d'une stratégie de mesure.

Il est important de souligner que nos travaux de recherche et notamment les méthodes ci-après présentées ne s'appuient pas sur un seul terrain d'observation. Les approches méthodologies présentées sont à la fois issues de nos entrevues réalisées auprès d'éditeurs de solutions logicielles et d'agences, mais aussi et surtout de manière plus pratique de mises en situation. En effet, certains éditeurs nous ont proposé des démonstrations du processus de mesure à partir de cas clients concrets. Nous ne présenterons de fait pas un terrain d'observation spécifique dans cette partie, mais nous appuierons plutôt sur les exemples concrets auxquels nous nous référerons au fur et à mesure de la déclinaison du processus de mesure.

3.1. L’analyse de la présence en ligne : entre intelligence économique, veille, audit et social listening

Toute démarche de mesure de la réputation en ligne suppose en amont un état des lieux de la présence numérique de l’entité dont on mesure l’e-réputation. Qu'elle soit perçue comme relevant de la veille, de l’intelligence économique, de l’audit ou du *social listening*, cette phase préliminaire vise à partir d’un crawler du web (le plus souvent automatique même s’il existe des crawler manuels), à collecter les *inputs* qui serviront de base, qui seront utilisés dans la mesure de la réputation, et répondant à une ou plusieurs de ces problématiques : comment est-elle perçue sur les diverses plateformes ? Quel sentiment ressort des interactions et expériences avec les clients / consommateurs ? Quels contenus apparaissent dans les résultats de recherche (sur les moteurs de recherche, médias sociaux, sites de *ratings*) lorsqu’on recherche l’entité ? Quelle est la teneur ou polarité du contenu (positif, négatif, neutre) mis en avant ? Quels types de contenus et quelles plateformes sont les plus adéquats afin d’accroître la visibilité ou notoriété ? Comment est perçue l’entité par rapport à son environnement concurrentiel ? Quels secteurs, domaines ou activités fonctionnent le mieux ? Et lesquels présentent d’éventuels risques pour l’entité ? Comment gérer (effacer ou noyer) les avis négatifs ? Quelle est la résultante des actions menées par l’organisation et comment les optimiser ? Quelles sont les sources influentes (de manière positive ou négative) ? etc.
La réussite de la stratégie repose essentiellement sur le cadrage préalablement effectué en amont de la collecte des données : l’identification des diverses sources à surveiller, le choix des requêtes, et plus généralement le paramétrage de la solution, ce qui permettra de garantir une recherche efficace.

La question du choix des requêtes est essentielle, car de leur précision dépendra la qualité et la pertinence des données collectées. Il existe de fait une nomenclature spécifique selon l’objectif de l’étude avec notamment l’utilisation d’opérateurs booléens (de type *or, and, ni*, etc.), d’opérateurs de proximité, de troncature (asso pour association par exemple), d’*alias* (ou systèmes de sous requêtes visant par exemple à regrouper divers champs sémantiques) permettant d’affiner le filtrage des données comme nous montre l’exemple ci-dessous (à partir d’une requête sur le café) :



Figure n°17 : Exemples d'opérateurs booléens utilisés pour le requêtage, source site web corporate Sproutsocial[326]

L'étape suivante consiste au paramétrage de la solution qui prend en compte plusieurs phases : les sources de collecte des données[327], les langues[328], le pays, la temporalité. Sur le *sourcing*, il est important de noter que les éditeurs de solutions logicielles de l*istening* sur le marché disposent de millions de sources regroupées selon plusieurs typologies au compte desquelles on retrouve entre autres les sites de presse nationaux / internationaux, les blogs, les forums, les sites d'entreprise, les réseaux sociaux (qui présentent des disparités selon les plateformes et concernent autant ceux nationaux qu'internationaux), les bases de données externes, les sites d'avis clients (qui se développent de plus en plus), etc. Les sources sont référencées en multi langue et selon le pays.

---

[326] Site web corporate de la plateforme Sproutsocial https://sproutsocial.com/features/social-media-listening/

[327] Plus de 50 millions (d'après un de nos enquêtés) pour la plateforme Sindup par exemple

[328] 82 langues pour la plateforme Digimind par exemple y compris l'argot, 187 langues pour Talkwalker

Dès lors que le paramétrage sourcing, langue, pays est effectué, il convient de définir la temporalité ou périodicité de collecte des données souhaitée (hebdomadaire, mensuelle, trimestrielle, etc.), puis de procéder à un affinage de la requête au besoin. L'affinage qui utilise le principe de l'entonnoir, vise à créer au départ des requêtes assez larges puis au fur et à mesure de la collecte et afin d'éviter les bruits et/ ou le cas échéant obtenir plus de données dans la mesure où les premières recherches seraient trop sélectives, de procéder selon le cas à certaines modifications : exclusion de termes, ajout d'opérateurs, exclusion de sources, ajout de nouveaux mots clés, *clipping* (ajout manuel de nouvelles sources ne figurant pas dans le bouquet de sources initiales, mais pertinentes pour l'étude).

À partir de là, et afin de catégoriser les données collectées et selon les plateformes, il est possible de créer des dossiers de veille reprenant l'arborescence de la veille selon les objectifs de l'analyse : veille concurrentielle (relative notamment au secteur d'activités au suivi des actions et de la réputation des principaux concurrents), veille commerciale (retombées des actions commerciales mises en place), juridique, e-réputation (perception des parties prenantes sur l'entité avec notamment les publications sur les sites de réseaux sociaux numériques), actualités (suivi en temps réel des publications sur les plateformes de médias en ligne), etc. Le paramétrage des solutions de *listening* offre d'énormes possibilités selon les débouchés de l'analyse, et dès lors que les données sont collectées, la prochaine étape consiste au traitement algorithmique via notamment les approches en intelligence artificielle.

Sur la collecte des données, plusieurs approches d'intelligence artificielle peuvent être utilisées dont les techniques de fouille de textes ou *text mining*. Ensemble de techniques liées à l'IA, la fouille de textes qui fait à la fois appel aux techniques langagières, linguistiques, statistiques, informatiques et sémantiques vise à partir de référentiels, à automatiser la structure et classification des documents d'un corpus et d'en identifier / extraire les *patterns*, unités langagières (mots ou *token*), co-occurrences de termes, formes grammaticales et lexicales, leur signification et les relations qui les lient et procéder à une *lemmatisation*[329].

[329] C'est-à-dire à identifier les différentes déclinaisons ou flexions d'un terme

De manière plus concrète, lorsque les requêtes sont exécutées via la solution de *listening*, l'algorithme pré testé sur un corpus va d'abord fouiller le web social selon la profondeur de la solution utilisée et le paramétrage initial (sources de collecte, langues, pays, etc.) et utiliser un référentiel composé d'un ensemble de variables sémantiques (mots clés, ontologies, listes à plat, termes normalisés, par exemple) et statistiques (réseaux bayésiens par exemple) afin d'identifier et procéder à la classification ou catégorisation de l'ensemble des textes dans le corpus. Les catégories de classification des données qui sont essentielles et modulables selon les solutions rappelons-le, se déclinent sous forme de tableaux de bord avec une myriade de widgets et filtres paramétrables selon les solutions et au sein desquels on peut retrouver : les données d'analyse concurrentielle ou sectorielle, des données de requêtage, l'analyse de sentiment, les publications sociales, l'analyse d'influence, par exemple. Nous nous appuierons ici sur une étude de cas réalisée sur la compagnie aérienne Air France[330].

- La catégorie relative à la perception de l'entité et notamment au requêtage et qui reprend plusieurs indicateurs notamment les recherches les plus récurrentes sur un "objet", les mots clés les plus utilisés, la progression des recherches (en hausse si le sujet est d'actualité ce qui peut dans certains cas être les prémisses d'une éventuelle crise, statique ou encore en baisse), les meilleurs résultats, etc. Si on se réfère à l'exemple du tableau de bord d'une des solutions du marché (*Digimind*), avec la requête "Air France" appliquée à l'Allemagne et avec des résultats uniquement en langue allemande, on constate que le top 4 des recherches est affilié " au check-in, au crash du vol 447 Rio-Paris de 2009, au statut des vols et à la valeur boursière du groupe Air France-KLM (*aktie*)". Il est intéressant de noter que la requête a été effectuée courant novembre dernier pendant la seconde vague de la crise de Covid d'où les requêtes associées "corona, hotline, etc.". L'un des éléments clés qu'offre l'interface est le *ranking* des recherches, ce qui permet à l'entité concernée (Air France ici) de mettre en place des actions marketing et de communication par exemple afin de rassurer les clients de la compagnie sur les points évoqués. Une autre des possibilités et non la moins importante qu'offre l'interface est, comme nous l'avons noté, la possibilité de détecter la progression des diverses requêtes. Si on

[330] Lors d'une démonstration sur une des plateformes de *social listening* par un de nos interviewés

revient sur notre exemple, on constate que, toujours sur la même requête (Air France), les mentions relatives au check-in des vols, à leur statut, et aux vols en "*premium economy*" ont progressé ce qui est la résultante des annulations et reports de vols liés à la crise sanitaire (la requête ayant été effectuée rappelons le en novembre dernier), celles relatives au crash du vol 447 et aux remboursements (*rückerstattung*) restent stables et toujours dans le top 10 et enfin, les recherches portant sur le groupe, le cours de l'action du groupe, le numéro de la hotline, les vols ont par contre baissé.



Figure n°18 : Classification des requêtes, interface de démonstration du tableau de bord de la plateforme Digimind

Toutes ces données qui apparaissent via l'interface de manière brute seront par la suite analysées et interprétées par des analystes de données (*data analyst*), interprétation qui est l'une des phases les plus importantes, car c'est à partir d'elles que seront décidées les actions à mettre en œuvre. Si on prend notre exemple des recherches et notamment la quatrième relative aux actions boursières, on peut postuler qu'elle figure parmi le top 4 dû à l'inquiétude des personnes ayant investi sur les actions du groupe et à l'affût d'éventuels mouvements sur le cours de l'action (surtout si elle est en baisse).

Il en est de même pour les recherches relatives aux remboursements, à la hotline, aux vols et à leur statut, au crash du vol 447 qui, même si on constate une baisse dans leur position restent néanmoins dans « le top 10 », ce qui se justifie par le contexte dans lequel la recherche a été effectuée pour les trois premiers et nécessitant une communication de la part du groupe sur ses sujets. Les résultats de ces requêtes sont classés avec une mise en avant des "meilleurs résultats des recherches" renvoyant aux requêtes.

- La catégorie relative au sentiment, à la tonalité ou polarité du corpus reprend le questionnement Quintilien (qui, quoi, quand, où, comment). Il est important de souligner que l'analyse de sentiment est très controversée, les algorithmes n'arrivant pas à déceler certains traits tels l'humour, le sarcasme, etc. Cette tonalité (il est ici question de la réputation et de sa valeur) qui pouvait auparavant être perçue comme bonne ou mauvaise (Boistel, 2014) est aujourd'hui déclinée qualitativement en plusieurs sentiments et utilise les techniques de *Natural Language Processing* (NLP). Plus concrètement, à partir d'un référentiel, l'algorithme de *listening* attribue une polarité à un corpus, et ce à partir de la teneur des mentions faites sur une entité, et portant sur plusieurs critères : le contenu du message qui est diffusé (*what* : il s'agit ici d'un extrait du contenu relatif à la requête avec les concepts clés se rapportant à la tonalité), la date à laquelle le contenu a été mis en ligne (*when),* la ou les diverses plateformes sur lesquelles il a été publié, repris (*where*), la personne à l'origine de la publication (*who*) et enfin (l'un des critères les plus importants dans cet onglet) la teneur du contenu (*how)* qui reprend autant la tonalité du message véhiculé comme on peut le constater sur l'exemple ci-dessous, mais aussi les éléments connexes servant à la mesure tels les émoticônes utilisées et les émotions véhiculées de manière implicite.



Figure n°19 : Analyse de sentiment, source interface de démonstration du tableau de bord de la plateforme Digimind

Dans l'exemple ci-dessus, la requête initiale porte sur la polarité attribuée aux compagnies aériennes sur une période de 7 jours (du 12 au 18 novembre 2020) : la recherche a permis comme indiqué dans l'interface de faire émerger plus de 58 000 (59 192) mentions sur 6 compagnies aériennes réparties comme suit : Lufthansa (18112), British Airways (17615), Easyjet (12156), Air France (10730), Alitalia (4782). Sur la totalité des mentions, 14 000 sont classées comme positives, 7 000 négatives et 5 000 neutres. Il est important de souligner dans certains cas, que l'algorithme de *sentiment analysis* comme nous l'avons déjà noté se trouve dans l'incapacité d'attribuer une tonalité à certains contenus. C'est notamment le cas de cet exemple dans lequel les mentions liées au sentiment ne font pas apparaître la totalité des recherches et 33115 mentions supposées ne pas contenir de "sentiment" ne figurent de fait pas dans l'analyse.

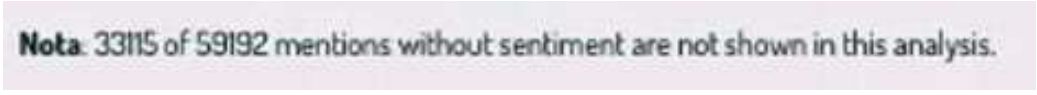

Figure n°20 : Notification de contenus non pris en compte dans l'analyse de sentiment, source interface de démonstration du tableau de bord de la plateforme Digimind

Dans une approche générale, et comme le montre cet exemple, les mentions se présentent comme suit : la plateforme sur laquelle le contenu a été publié (il s'agit ici de la plateforme de micro blogging twitter), la personne à l'origine de la publication (avec la possibilité de retrouver le message d'origine et l'activité de celle-ci en cliquant sur "*view activity*"), la date de publication (ici le 14 novembre 2020), l'audience totale qui a été touchée par le message, le *ranking* du message selon l'audience qui a été exposée, les mentions similaires.



Figure n°21 : Exemple de mention sur une entité sur Twitter, source : interface de démonstration du tableau de bord de la plateforme Digimind

Il est intéressant de noter le concept mis en gras "***peine***" et qui par ailleurs est en grande partie (mais pas uniquement) à l'origine du sentiment attribué au contenu et qui, même si dans ce contexte a été employé à bon escient (la teneur du message étant effectivement négative) pourrait dans un autre contexte porter à confusion. À partir donc du référentiel de sentiment, l'algorithme permet de classer ce contenu comme négatif avec les concepts clés servant à catégoriser le contenu en vue de l'analyse et de l'interprétation des résultats (catégorisation et unités de sens qui dépendent de la solution choisie : dans cet exemple, il est indiqué que le contenu provient d'une plateforme *earned média* et qu'il concerne la relation, la satisfaction client, ce qui permettra à la compagnie de mettre en place des actions correctives sur la gestion de la relation client, ou contacter la personne mise en cause directement, etc.).

Sur l'analyse des émotions, les plateformes de *listening* utilisent comme vu précédemment un référentiel avec une pluralité de typologies de sentiments. À titre d'exemple, on peut retrouver dans le référentiel des sentiments relatifs à la surprise, à la joie, à la tristesse, au dégoût, à l'amour.

L'analyse des émotions et du sentiment se voit complétée par celle des emojis, via les approches en IA qui facilitent leur reconnaissance et leur classification :

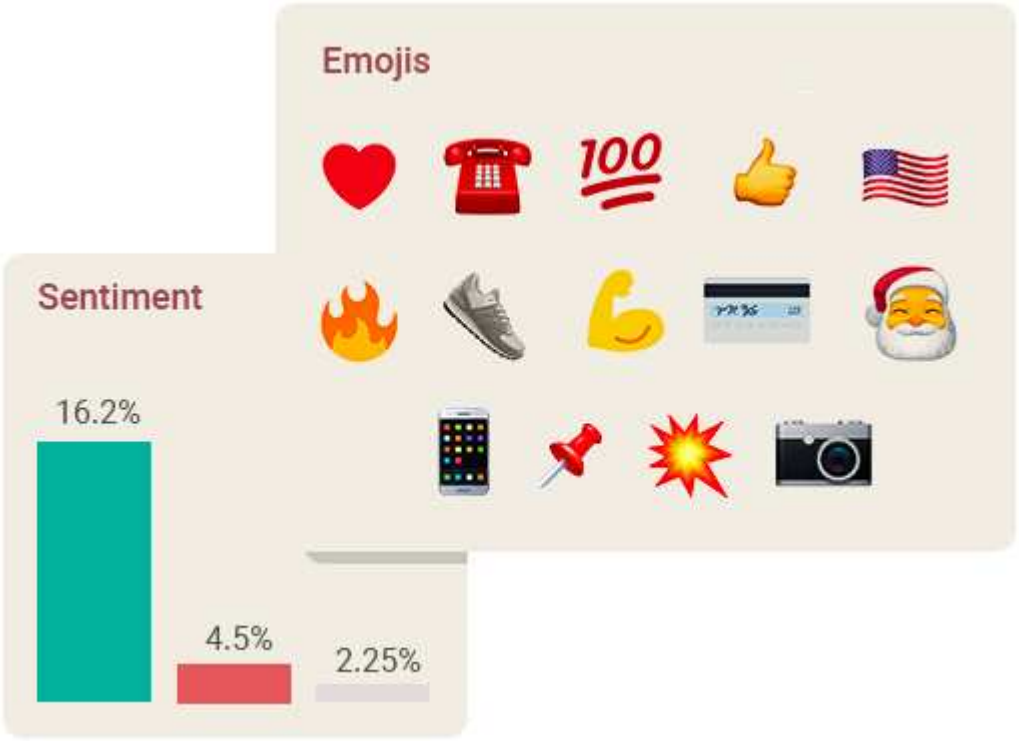

Figure n°22 : Typologie d'emojis pris en compte dans l'analyse du sentiment, Source site web corporate de Linkfluence

- L'onglet relatif à la réputation quant à lui, reprend la vision dans son ensemble de la réputation globale de l'entité sous forme de tableau de bord composé de plusieurs indicateurs (certains étant spécifiques à certaines plateformes[331]) selon la plateforme utilisée[332] :



Figure n°23 : Critères de réputation du tableau de bord, source : Interface de démonstration (Source Digimind[333])

Si on reprend notre requête sur la compagnie Air France KLM, le tableau de bord ci-dessus, qui représente l'image de l'entité telle que perçue par ses divers publics est construit à partir de plusieurs critères : les tendances, top 10 des hashtags les plus utilisés, le top 10 des autosuggestions Google, le classement par type de média des mentions, la réputation par

---

[331] Il est important de souligner comme mentionné que les variables utilisées dépendent et sont variables selon les solutions du marché et nos analyses sont basées sur les données issues de nos entretiens et l'analyse des sites corporate des solutions

[332] Notre exemple utilise le tableau de bord de Digimind

[333] Un des éditeurs de solution de veille en e-réputation https://www.digimind.com/fr/

thématique, les tendances par thématique, la répartition des mentions selon qu'elles appartiennent au *owned ou earned* média, le taux de satisfaction, les émotions relatives à la marque.

Sur les tendances, et à partir d'une recherche effectuée sur la marque entre octobre et novembre 2020, elles permettent de déterminer le nombre de personnes parlant du groupe (27500 pour cette période), le nombre de mentions total (qui est le nombre de fois que la marque a été citée 62200 fois ici), le taux de *reach* ou de couverture de la totalité des publications sur la marque (18,2 millions) et les interactions sociales portant sur l'entité (au nombre de 666 600).

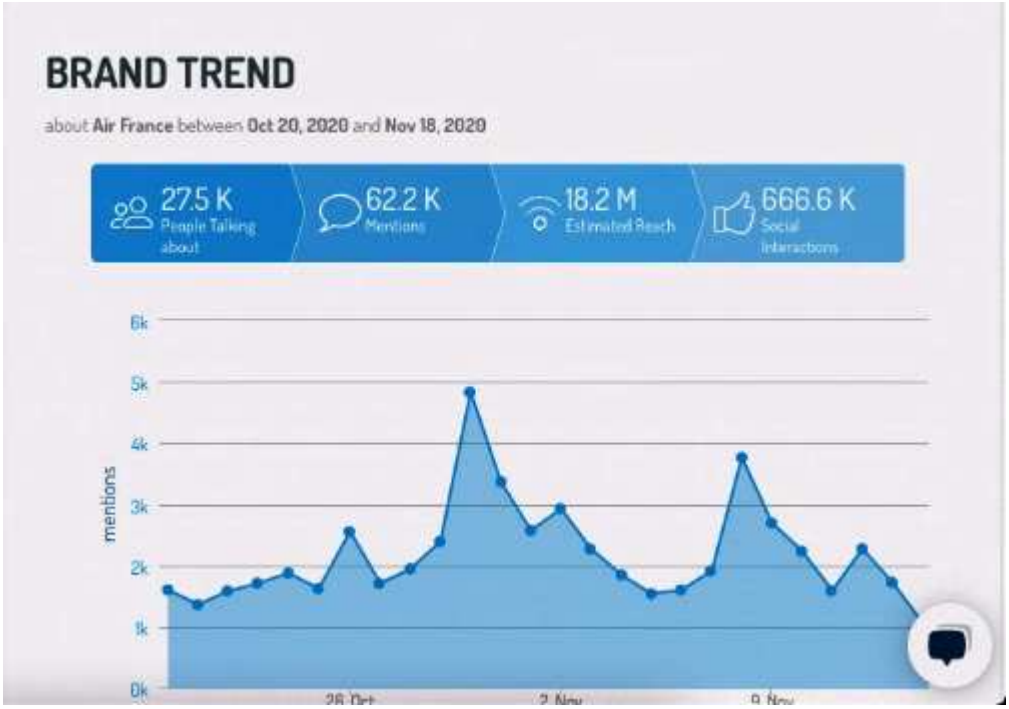

Figure n°24 : Interface de démonstration du tableau de bord relative aux tendances sur une requête (Source : Digimind)

Les tableaux de bord proposent des données démographiques assez complètes sur les clients, consommateurs et plus généralement les internautes qui parlent d'une entité, données se référant à titre d'exemple à leur position géographique, leur occupation, la tranche d'âge concernée, les plateformes les plus utilisées.

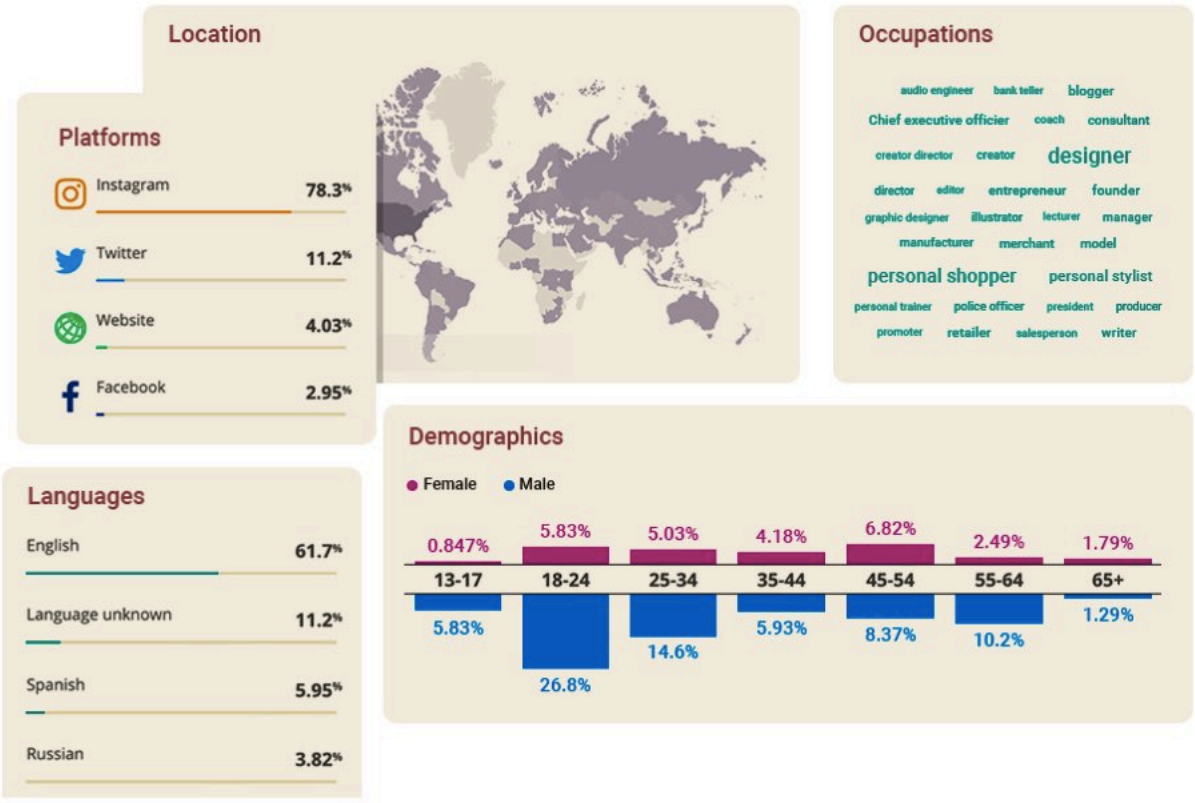

Figure n°25 : Exemple de (re) présentation de données démographiques, source : site web corporate de Linkfluence

Afin de proposer une analyse complète avec des *insights* clés, le tableau de bord permet également d'identifier et classifier à partir de la totalité des mentions et des personnes qui produisent le contenu, les divers supports médias utilisés avec pour chacun le nombre de mentions (comme le montre l'exemple ci-dessous).



Figure n°26 : Exemple de présentation des mentions par type de média, source : Digimind

Au compte des indicateurs liés à la réputation, on retrouve également les diverses thématiques abordées par les internautes et pour chacune, le sentiment qui y est associé. Les thématiques peuvent par exemple être liées aux différents produits ou services proposés par une entité (comme c'est le cas pour l'exemple ci-dessous avec la recherche sur la compagnie de voyage Air France) :

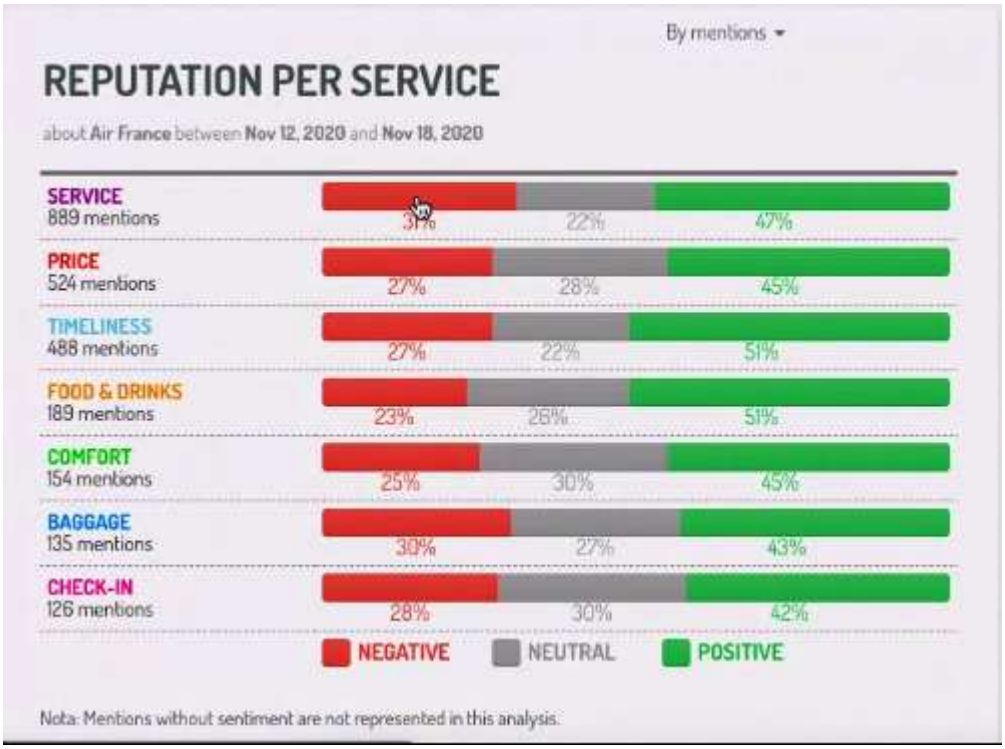

Figure n°27 : Exemple de classification des services d'une entité par tonalité, source Digimind

- L'onglet relatif à la veille concurrentielle : l'une des problématiques clés dans la mise en place d'une stratégie de *listening* ou de veille est de collecter des informations sur le marché au sein duquel évolue l'entité, ce qui suppose définir les principaux concurrents qui lui sont associés et la perception de ces derniers vis-à-vis de l'entité dont il est question. À l'image d'une stratégie de s*ocial listening* classique pour une entité donnée, la veille concurrentielle ou stratégique utilise les mêmes indicateurs que ceux précédemment énoncés avec pour finalité d'identifier dans une première approche les principaux concurrents associés à une entité (marque ou produit) ce qui inclut la détection des nouveaux entrants au sein du marché, cartographie des acteurs à partir de laquelle il sera par la suite possible pour chacun des concurrents, de suivre et analyser les publications, mentions, retombées des actions et leur incidence sur le marché dans une approche plus générale (le graphique ci-dessous offre une représentation de l'évolution des tendances du marché automobile à partir des recherches et conversations des utilisateurs et des actions mises en oeuvre).



Figure n°28 : Analyse concurrentielle de la réputation d'une entité et de son marché , source : site web corporate Digimind

Le champ des possibles en termes de *widgets* ou d'indicateurs compris dans les tableaux de bord est, nous le voyons, quasi infini (ou du moins très large), et diffère selon la plateforme utilisée (Digimind[334], Linkfluence[335], Sindup[336], talkwalker[337], Synthesio, Mention, Meltwater[338], etc.), les objectifs visés par l'analyse et les besoins clients, mais dans une approche plus générale les indicateurs clés renvoient à la portée des publications, aux données analytiques démographiques, à l'environnement concurrentiel, aux *hashtags* les plus utilisés si on s'intéresse particulièrement au microblogging Twitter, et plus généralement à la réputation d'une entité.

---

[334] Exemples de plateformes spécialisée dans la veille en e-réputation https://www.digimind.com/fr/

[335] Exemples de plateformes spécialisée dans la veille en e-réputation https://www.linkfluence.com/fr/

[336] Exemples de plateformes spécialisée dans la veille en e-réputation https://fr.sindup.com

[337] Exemples de plateformes spécialisée dans la veille en e-réputation https://www.talkwalker.com/fr

[338] Exemples de plateformes spécialisée dans la veille en e-réputation https://www.meltwater.com/fr/solutions/social-listening

Il est important de souligner que les plateformes social média ne se résument pas aux réseaux sociaux (comme on serait tenté de le penser *a priori*), mais regroupent en plus de ces derniers qui en sont une typologie, les plateformes de publication (blogs, forums, etc.) ; de discussion, conversationnelles, de gestion de commentaires (Ask, quora, reddit, etc.) ; de partage (audio, vidéo, etc.) ; collaboratives, de networking, etc[339]. La phase d'audit, de veille en e-réputation ou de s*ocial listening* s'appuie essentiellement sur l'ensemble de ces catégories de plateformes selon comme nous l'avons vous, l'aptitude et la granularité des solutions du marché à accéder aux données. De manière générale, cette phase de veille de la présence en ligne vise à s'appuyer sur les données *analytics* des diverses plateformes sur lesquelles l'organisation est présente et dans les environnements au sein desquels interagissent ses divers publics, afin d'identifier les métriques clés sur lesquelles portera la stratégie de gestion. Les outils de reporting que proposent les solutions dans l'analyse des *analytics* visent à centraliser sur un tableau de bord notamment corrélé à *Google Analytics*, les métriques avec les indicateurs clés à surveiller. Il convient de souligner que cette phase d'analyse des pages, publications et l'ensemble des données disponibles est réalisée à la fois pour l'entité dont on mesure la réputation, mais aussi et surtout les acteurs de son environnement concurrentiel. À partir de là, il s'agit de déterminer les meilleures pratiques et les plateformes les plus adéquates selon les objectifs stratégiques de l'organisation, et de mettre en place la stratégie d'optimisation de la visibilité.

Si les approches de veille ou *social listening* sont essentielles et visent à évaluer dans une certaine mesure la réputation de manière générale d'une entité, elles ne sauraient être utilisées telles quelles et indépendamment de l'apport d'autres approches qui lui sont connexes (la construction et la mise en place des actions issues de l'analyse dans une approche générale). Ces dernières, même si elles relèvent en grande majorité de la gestion de la réputation et non de la mesure à proprement parler, viennent en amont de la stratégie d'écoute du web, mais aussi et surtout comme support au déploiement des actions à mettre en œuvre à l'issue de la phase d'écoute ou de veille. Nous présenterons dans la section suivante, et notamment à partir de nos données de terrain, quelques-unes des approches proposées par les professionnels qui composent le marché.

---

[339] Se référer au chapitre 3 de la première partie sur les réseaux socionumériques et leur analyse

3.2. La gestion de la réputation : de la gestion de crise au recours légal

S'il est aujourd'hui une question essentielle dans la gestion de la réputation en ligne, c'est bien celle liée à la crise et aux diverses approches mises en œuvre afin d'anticiper, de prévenir ses effets. En effet, l'un des objectifs dans la veille ou l'audit en e-réputation est de déceler en avant-première, ce qu'on qualifie de signaux faibles qui sont l'ensemble des éléments de langage à teneur négative ou péjorative sur une entité et susceptibles, une fois cumulés, repris, remixés sur l'ensemble des espaces numériques, de porter atteinte à l'image de celle-ci. Certaines agences proposent des services allant de la gestion des contenus sur les plateformes numériques au recours légal en cas notamment de diffamation ou d'atteinte à la vie privée.

À mi-chemin entre la communication (notamment de crise), le marketing (*black hat seo*) et le droit, l'approche de gestion des contenus sur l'ensemble des plateformes sociales et les moteurs de recherche utilise plusieurs techniques selon l'origine des publications et leurs effets qui peuvent être plus ou moins importants sur l'image et la réputation de l'entité, et la stratégie à déployer peut aller d'un simple droit de réponse sur les plateformes corporate de l'entité avec une justification de l'action ou la personne mise en cause, à la suppression des contenus incriminants et pour les cas extrêmes à la mise en place d'une action judiciaire.

3.2.a. La modération de contenus sur les media sociaux

La première méthode utilisée dans la gestion des contenus sur les plateformes *social media* est relative à la modération de contenus. Méthode visant à s'assurer que l'ensemble des avis, contributions et commentaires laissés par les internautes respectent les chartes de publication et ne sont ni diffamatoires, ni insultants, ni haineux ni à caractère raciste, etc., la modération peut s'effectuer en amont ou en aval de la publication. La phase amont suppose la mise en place d'une charte (dans les CGU par exemple) permettant de filtrer automatiquement les publications jugées inappropriées avant leur mise en ligne. Technique de modération automatique, elle utilise les méthodes de traitement automatique qui comme nous l'avons déjà évoqué (notamment via l'analyse de sentiment) peut présenter certaines failles : certains

contenus négatifs peuvent échapper au filtrage mis en place ou *a contrario*, d'autres positifs peuvent être perçus comme négatifs et supprimés (les tournures ironiques par exemple). À l'inverse, la seconde typologie qui peut utiliser une approche manuelle ou automatique, consiste à autoriser de manière "provisoire" les publications des avis des internautes tels qu'exprimés, puis par la suite, au cas par cas et selon la teneur, de procéder à leur analyse, ce qui suppose par ailleurs une réactivité de la part des modérateurs, et une veille continuelle.

Il est important de souligner, comme le notent les agences spécialisées dans le domaine, que les méthodes automatiques qui permettent de filtrer les contenus sont utilisées de pair avec celles manuelles, c'est-à-dire que les contenus validés automatiquement font l'objet d'une vérification manuelle et nécessitent peu importe la teneur (positifs ou négatifs) une réponse. À l'instar de la modération effectuée sur les plateformes de publication (notamment la partie relative aux commentaires sur les sites corporate et blogs) qui est relativement simple, la modération sur les réseaux sociaux revêt une autre dimension eu égard à la quantité exponentielle de données diffusées à la fois sur les pages corporate des entités en question, mais aussi et surtout celles des concurrents.

Dès lors, et afin de collecter l'ensemble de ce corpus, les agences utilisent pour certaines, les solutions de *listening* proposées par les éditeurs sur le marché et adaptées à l'écoute des audiences et mentions sur les médias sociaux. Ce sont entre autres : Awario[340] solution permettant à la fois via son interface d'écouter les mentions sur Twitter, Facebook, Reddit, Instagram, les blogs, forums et la presse en ligne, d'y répondre et publier du contenu ; Hubspot[341] qui propose entre autres un système d'*alerting* par mail par exemple lorsqu'une marque, entreprise, produit ou mot clé surveillé est mentionné et d'identifier les internautes à l'origine des mentions ; Mention[342] ; Hootsuite[343] qui permet d'écouter les mentions, mais

---

[340] Solution de *listening*, accessible en ligne sur https://awario.com

[341] Outil de gestion des plateformes socionumériques. Accessible en ligne sur https://www.hubspot.fr/products/marketing/social-inbox?hubs_post=blog.hubspot.fr%2Fmarketing%2Fsocial-listening&hubs_post-cta=HubSpot

[342] Solution d'écoute et de monitoring du web social. Accessible en ligne sur https://mention.com/fr/

[343] Solution complète de gestion des plateformes socionumériques. Accessible en ligne sur https://www.hootsuite.com

aussi de planifier la publication de ses contenus ; Agora pulse[344]; Meltwater[345]; Brandwatch[346]; Tweetreach qui permet de suivre les mentions sur Twitter[347] au même titre que Twazzup[348]; Keyhole[349] qui utilise Instagram et Twitter, etc. Dès lors que les mentions sont collectées, elles sont catégorisées dans les tableaux de bord, et il conviendra par la suite pour chacune de mettre en place la stratégie adéquate de réponse ou suppression.

Certaines agences se sont spécialisées dans le domaine (Rankwell, Webhelp, Atchik, Netino, Solike[350], 50A[351]) et à l'instar des développeurs de solutions logicielles, d'autres ont également développé des solutions propriétaires de modération de contenu interne : c'est notamment le cas de Netino[352] avec son outil moderatus, Atchik[353], Webhelp[354] par exemple. Si la modération des contenus sur les médias sociaux semble relativement simple et qu'il est possible de supprimer les avis négatifs, les moteurs de recherche (Google en l'occurrence) et les sites d'avis, qui rassemblent à eux seuls les avis les plus vus et influents sur la réputation d'une entreprise ou organisation, n'offrent malheureusement pas cette possibilité. Ce qui pousse certains acteurs de la réputation à recourir à d'autres pratiques.

---

344 Plateforme de gestion des réseaux sociaux. Accessible en ligne sur https://www.agorapulse.com/fr/

345 Solution de veille média et social listening. Accessible en ligne sur https://www.meltwater.com/fr

346 Solution de social listening. En ligne sur https://www.brandwatch.com

347 Outil de suivi des mentions sur la plateforme de microblogging Twitter. https://tweetreach.com

348 Outil de suivi des mentions sur la plateforme de microblogging Twitter. http://new.twazzup.com

349 Agence spécialisée dans le social listening https://keyhole.co

350 Offre de modération de contenus proposée par l'agence Solike. Accessible en ligne sur https://www.solike.eu/definition-moderation-de-contenu/

351 Agence social media et e-réputation. Accessible en ligne sur https://www.50a.fr/agence-social-media

352 Solution de modération de contenu sur les plateformes socionumériques https://netino.fr/moderation/

353 Solution de modération de contenu https://atchik.com/activites/moderation/

354 Solution de modération de contenu https://webhelp.com/fr/nos-solutions/services-digitaux/

3.2.b. La suppression ou le *flooding* (noyage) des contenus négatifs dans les résultats de recherche

Lorsqu'on effectue une recherche sur une entité donnée sur le moteur de recherche Google par exemple, les premiers résultats auxquels on a accès sont notamment relatifs à sa page Google my business (GMB) qui reprend l'ensemble des avis des internautes et les liens vers les documents (articles par exemple) la mentionnant. Les avis, qui ne sont pas vérifiés, au même titre que les autres documents, présentent un risque en termes de réputation dans la mesure où il est impossible pour l'organisation de supprimer ceux qui sont diffamatoires ou qui portent atteinte à leur image. L'un des challenges auquel font face les agences spécialisées dans la gestion de la réputation est de fait de mettre en place des stratégies visant à supprimer ou à minima à noyer les avis négatifs publiés. Pour ce faire, et selon le type de contenu et l'agence, trois principales actions peuvent être mises en place : les actions à visée communicationnelle, l'approche marketing numérique et le recours légal.

La première approche, qui est relativement la plus simple à mettre en place (selon l'origine et la teneur), consiste à contacter directement les internautes à l'origine des contenus dénigrants, diffamatoires, etc., afin de mettre en place la stratégie adéquate. À titre d'exemple et si on se réfère à l'offre proposée par l'agence Net offensive[355] (qui est un des acteurs évoluant spécifiquement dans le domaine au même titre que plusieurs autres), plusieurs scénarii peuvent être envisagés : lorsque des commentaires à caractère diffamatoires sont publiés sur un ou plusieurs forums relatifs à un produit, service ou un individu (il s'agit là de *personal branding*), la stratégie proposée dans l'offre de communication de l'agence vise à prendre contact avec les éditeurs des forums concernés après les avoir identifiés, afin d'obtenir la suppression des avis abusifs. Autre cas de figure, dans la mesure où un individu (on parlera plutôt ici de personnalité publique) se verrait associé à des faits plutôt graves (détournement, harcèlement, etc.) dans un article sur un site de presse par exemple, l'agence se propose de négocier avec la rédaction du journal afin de limiter l'accès à l'information incriminante.

---

[355] Plateforme de gestion des contenus négatifs. Accessible sur https://www.netoffensive.com/contenus-negatifs.html

L'approche en communication consiste également lorsqu'il s'agit d'avis publiés sur *Google my business (GMB)* par exemple, soit de le signaler ou de contacter la personne à l'origine de l'avis en privé afin de mettre en place une stratégie de relation client visant à le "*reconquérir*" et obtenir la modification des avis, où, le cas échéant de contacter les clients "satisfaits" afin d'encourager la publication d'avis favorables visant à déréférencer ou "*noyer*" les avis négatifs et faire remonter le score de réputation. Afin de signaler un avis sur une fiche GMB, il suffit de se rendre sur son profil, "gérer l'établissement", "avis", choisir les contenus en question puis les signaler (via l'icône sous forme de drapeau) en renseignant le type d'infraction[356]. Il est important de souligner que signaler un contenu ne garantit pas sa suppression, d'où la nécessité de mettre en place une stratégie de relation ou de satisfaction client et de contacter ces derniers pour qu'ils modifient leurs avis, encourager la production d'avis positifs ou en dernier recours supprimer et recréer un nouveau compte *GMB* (solution qui n'est pas recommandée dans la mesure où les anciens clients ou consommateurs insatisfaits pouvant de nouveau laisser des avis négatifs) et qu'il est par ailleurs important de s'assurer que ces derniers soient réellement diffamatoires et inappropriés avant d'entamer une telle procédure.

Cette première approche, même si elle semble *a priori* relativement facile à déployer, nécessite à la fois comme nous l'ont indiqué les agences qui proposent le service une grande expérience dans le domaine, un important portefeuille relationnel, mais aussi et surtout des compétences en relations publiques / presse. Ces dernières (les actions en communication) n'étant pas toujours efficaces (certains contenus étant par exemple liés à des faits relativement graves de harcèlement, détournement, trafic, etc.), l'approche en marketing peut s'avérer une solution. L'approche marketing digital dans la gestion des contenus négatifs sur une entité utilise les techniques de *black hat SEO* : il s'agit non pas réellement de supprimer les contenus négatifs, mais plutôt de les noyer dans les résultats *SERPs*, de limiter leur accès. Le principe est simple : partant du postulat qu'au-delà de la première page des résultats des moteurs de recherche, le taux d'audience sur les suivantes est fortement réduit, l'objectif est de fait de produire et mettre en avant les contenus à forte valeur ajoutée sur l'entité afin de

---

356 Google distingue quatre types de contenus ne respectant pas leur charte de publication : contenu inapproprié ou incitant à la haine ou à la violence, annonces publicitaires ou spams, hors sujet ou associé à des conflits d'intérêts.

faire reculer la position des résultats négatifs. Pour ce faire, cette approche utilise par exemple certaines techniques de *black hat* évoquées précédemment telles que la création de pages satellites, de *PBN, le cloaking*, les redirections trompeuses. Le principe des pages satellites est de créer de fausses pages qui par ailleurs sont invisibles des internautes, mais visibles par les bots qui crawlent le web afin de les indexer, et d'y publier du contenu très optimisé voire à la limite du spam avec en utilisant des requêtes spécifiques. Le *cloaking* consiste lui, à programmer un serveur afin qu'il diffuse sur une même page web, différents contenus ou *URL* selon que le visiteur est un moteur de recherche (avec une suroptimisation des termes clés) ou un internaute (via une version plus lisible). À l'image du *cloaking*, les redirections trompeuses visent à duper les moteurs de recherche afin qu'ils indexent les pages d'origine, puis par la suite de rediriger les internautes vers une nouvelle *URL* de destination. Les techniques de *black hat* comme nous le voyons sont nombreuses et les agences spécialisées dans le traitement des avis négatifs rivalisent d'ingéniosité afin de les appréhender. Dans la mesure où ces diverses approches de suppression des contenus négatifs proposées par plusieurs agences (reputation.net[357], Alterbuzz, Netwash, Veille e réputation[358], etc.) se révèleraient inefficaces, l'approche ultime proposée est le recours légal.

3.2.c. Le recours légal en cas de diffamation

Le dernier recours dans la mesure où l'ensemble des techniques et méthodes s'avèreraient inefficaces est celui juridique. Depuis quelques années, on constate la floraison d'offres légales sur les questions de réputation et leur gestion qui jadis, étaient l'apanage des agences de communication, relations publiques et marketing. Les actions juridiques déployées dans le cadre de la gestion de la réputation s'appuient sur certains articles relatifs à trois principales lois : la loi informatique et libertés, la loi sur la liberté de la presse et la dernière le droit à l'oubli.

---

[357] Agence spécialisée dans la veille en e-réputation. https://réputation.net

[358] Agence spécialisée dans la veille en e-réputation. Accessible en ligne https://www.veille-reputation.com

Dans le titre II portant sur les traitements relevant du régime de protection des données à caractère personnel prévu par le règlement (UE) 2016/679 du 27 avril 2016 (Articles 42 à 86) de la loi informatique et liberté du 6 janvier 1978[359] et plus précisément le chapitre II sur les droits des personnes concernées (Articles 48 à 56), il est stipulé que (nous ne détaillerons pas tous les titres, chapitres et articles, mais présenterons uniquement quelques-uns de ceux sur lesquels s'appuient l'approche légale de la gestion de la réputation) la personne concernée a le droit d'accès[360], de rectification[361] et d'effacement des données[362] lui concernant. En ce qui concerne la loi du 29 juillet 1981 portant sur la liberté de la presse[363] et même si celle-ci n'a pas spécifiquement été pensée pour les applications numériques, elle délimite, notamment le paragraphe 3 du chapitre IV sur les délits contre les personnes (les articles 29 à 35), le périmètre des infractions surtout ceux de diffamation, de dénigrement et d'atteinte à la dignité ou à l'honneur et les divers recours légaux qui peuvent être déployés. Autre principe régissant le recours légal est notamment lié au droit à l'oubli qui s'applique depuis le 13 mai 2014 aux questions de réputation et qui vise à protéger la vie privée de tout individu résidant dans l'Union Européenne, notamment les articles 15 (droits d'accès de la personne concernée), 16 (droit de rectification) et 17 (droit à l'effacement) du chapitre III relatif aux droits de la personne concernée[364].

---

[359] Extrait de l'article de loi portant sur la protection des données à caractère personnel. Accessible en ligne sur https://www.legifrance.gouv.fr/loda/id/JORFTEXT000000886460/

[360] Article 49 qui s'exerce dans les conditions prévues à l'article 15 du règlement (UE) 2016/679 du 27 avril 2016.

[361] Article 50 qui s'exerce dans les conditions prévues à l'article 16 du règlement (UE) 2016/679 du 27 avril 2016

[362] Article 51 qui s'exerce dans les conditions prévues à l'article 17 du règlement (UE) 2016/679 du 27 avril 2016.

[363] Extrait de l'article sur la liberté de la presse. Accessible en ligne sur https://www.legifrance.gouv.fr/loda/id/LEGITEXT000006070722/

[364] Extrait de l'article relatif aux questions de réputation et aux droits de la personne dans la protection de la vie privée. Accessible en ligne sur https://www.cnil.fr/fr/reglement-europeen-protection-donnees/chapitre3#Article17

Ce dernier peut en effet, à compter de cette date et selon certaines conditions (critères édictés par CNIL[365]), demander le retrait (de manière partielle ou totale) de certaines données qui apparaissent dans les résultats des recherches et qui par ailleurs portent atteinte à son image.

En principe, le droit au déréférencement[366] qui ne s'applique pas pour les personnalités publiques et organisations s'exerce en utilisant simplement un formulaire (de droit à l'oubli ou de demande de suppression) qui, par la suite, fera l'objet d'un examen, examen qui ne garantit pas la désindexation. Plusieurs acteurs du marché se sont spécialisés dans les actions judiciaires à déployer (outre le procédé de déréférencement) en cas d'atteinte à la réputation en faisant notamment appel à des cabinets d'avocats pour les problématiques que les actions à l'amiable n'arriveraient pas à résoudre : rédaction de procès-verbaux, mise en place d'un droit de réponse, identification technique et juridiques des auteurs et engagement de leur responsabilité, fixation du litige, qualification des faits, constitution du dossier et recueil d'éléments de preuve, saisie des instances judiciaires compétentes afin d'obtenir réparation (indemnisation par exemple), etc. Au compte de ces derniers, on retrouve des cabinets d'avocats (Haas avocats, cabinet spécialisé sur l'e-réputation[367], Avocats Picovschi[368], Porlon avocats[369], Alain Bensoussan avocats[370] par exemple), mais aussi des agences spécialisées en e-réputation et qui travaillent avec ces derniers Net Offensive[371], Alterbuzz[372], Netwash[373],

---

365 Extrait de l'article sur le droit au retrait de certaines données numériques. Accessible en ligne sur https://www.cnil.fr/sites/default/files/typo/document/Droit_au_dereferencement-criteres.pdf

366 Extrait de l'article relatif au droit au déréférencement. Accessible en ligne sur https://www.cnil.fr/fr/le-droit-leffacement-supprimer-vos-donnees-en-ligne

367 Prestation d'e-réputation proposée par le cabinet Haas Avocats https://www.haas-avocats.com/nos-competences/e-reputation/

368 Prestation d'e-réputation proposée par le cabinet d'avocat Picovschi https://www.avocats-picovschi.com/comment-l-avocat-peut-defendre-votre-e-reputation_article_1256.html

369 Prestation d'e-réputation proposée par le cabinet Porlon Avocat https://www.porlon-avocats.com/e-reputation-diffamation-denigrement-et-injure-sur-internet

370 Prestation d'e-réputation proposée par le cabinet Bensoussan https://www.alain-bensoussan.com/avocat-e-reputation/

371 Prestation proposée par l'agence Netoffensive https://www.netoffensive.com/contenus-negatifs.html

372 Prestation sur le nettoyage des données de réputation sur les moteurs de recherche proposée par l'agence Alterbuzz https://www.alterbuzz.com/nettoyer-son-e-reputation/e-reputation-google/

373 Prestation proposée par l'agence Netwash https://www.net-wash.fr

SOS e-réputation[374], etc. Il est important de souligner que les recours légaux, à l'image de l'ensemble des approches sus-citées, peuvent dans certains cas s'avérer inefficaces comme nous l'indique l'exemple du jugement du 9 mars 2021 du tribunal judiciaire de Paris sur un professionnel libéral (dentiste) souhaitant supprimer sa fiche *GMB*, demande qui a été refusée[375].

3.3. La construction de la réputation : l'approche par popularité

À partir de notre cartographie des acteurs du marché de la réputation, on constate qu'une grande partie des prestataires s'est spécialisée dans les techniques et méthodes d'aide à la construction de la popularité sur notamment les résultats des moteurs de recherche, et que ces offres occupent aujourd'hui une importante place dans la gestion de la réputation dans son entièreté. Nous nous intéresserons ici particulièrement à une des techniques proposée par certaines agences généralistes et ad hoc : l'offre de "*grey* et de *black hat seo*". Nous avons choisi de présenter cette offre qui ne relève certes pas de la mesure de la réputation, mais plutôt aide à sa construction, pour la simple raison qu'elle est construite autour d'une des métriques clés à l'origine d'une des catégories des modèles de réputation : l'influence. Les prestataires qui proposent cette offre s'appuient de fait, entre autres, sur le modèle PAGERANK de Brin et Page (1998).

Avec les mises à jour récurrentes des moteurs de recherche (notamment les dernières en date de Google) et eu égard au fait que les techniques de référencement naturel ne soient efficaces que sur le long terme, certains acteurs proposent des services subsidiaires qui par ailleurs sont interdits et sanctionnés par les grands acteurs du marché.

En effet, les algorithmes des moteurs de recherche et celui de Google en particulier différencient les liens naturels "en *dofollow ou follow*" de ceux sponsorisés, et Google

---

[374] Prestation de gestion des contenus négatifs. https://www.sos-e-reputation.com/notre-demarche-enfouissement-suppression-mauvaise-reputation/

[375] on peut la retrouver en intégralité sur le site d'actualité du droit des nouvelles technologies (Legalis) https://www.legalis.net/jurisprudences/tribunal-judiciaire-de-paris-5eme-ch-1ere-sec-jugement-du-9-mars-2021/

interdit même l'utilisation de liens sponsorisés avec la mention "*dofollow*" et l'obligation de la mention "*nofollow*" sur ces derniers vise à indiquer qu'ils n'ont pas d'impact sur la popularité du site et que les algorithmes ne devraient pas à en tenir compte. Le *grey hat* (à mi-chemin entre le *black et le white hat*) et le *black hat* (techniques qui ne respectent pas la charte d'utilisation des moteurs de recherche) s'appuient de fait essentiellement sur la stratégie de liens non naturels pour l'un et frauduleux pour le second.

Sur l'achat de liens, certaines plateformes proposent aux entreprises qui le souhaitent de leur vendre (de manière ponctuelle ou sous forme d'abonnement) des rétroliens provenant de sites thématiques affiliés à leur domaine ou secteur d'activité. À la différence d'une stratégie de netlinking *white hat* ou "organique" qui suppose une acquisition légale et quelque peu naturelle de *backlinks* via une stratégie éditoriale et de contenu (référencement) efficiente et qui par ailleurs est recommandée, le principe de l'achat de liens est simple : après avoir réalisé un audit du site et à partir de la proximité sémantique, il s'agit de sélectionner les sites les plus pertinents au sein desquels publier les liens qui redirigeront vers le site en question. Le choix des sites s'appuie sur plusieurs métriques selon les solutions proposées par les grands acteurs du marché en termes de pratiques liées au référencement. Au compte des métriques essentielles, on retrouve par exemple le TF/CF (*trust flow*[376] */ citation flow*[377]) de Majestic SEO qui permet d'évaluer l'indice de confiance et de qualité d'une page web, le DA (*domain authority*), le TB (*total backlinks*) et le PA (*page authority*) de Moz[378], le Tr (*trustrank*), le DR (*domain rating*) de Ahref[379], le RD (*referring domains*) mesuré par la plupart des solutions (Semrush[380] ou Majestic seo par exemple).
Plusieurs acteurs du marché de la réputation proposent de fait une offre double de référencement naturel cumulée à certaines approches *grey hat*. C'est notamment le cas de

---

[376] Mesure la qualité des liens qui renvoient vers une page

[377] Relatif à la popularité, la quantité de liens qui pointent vers une page

[378] Acteur spécialisé sur les questions de référencement. Accessible en ligne sur https://moz.com/products/pro/seo-toolbar

[379] Acteur spécialisé sur les questions de référencement. Accessible en ligne sur https://ahrefs.com/fr/website-authority-checker

[380] Plateforme de gestion de la visibilité sur les plateformes sociales. Accessible en ligne sur https://fr.semrush.com

l'agence Smartkeywords,[381] Optimiz.me[382], Semjuice[383], Boosterlink[384], Netlinka[385], Develink[386], Netoffensive[387], etc. Cette approche qui ne respecte pas intégralement les conditions d'utilisation des moteurs de recherche, mais n'est pas non plus frauduleuse est tolérée dans la mesure où les liens achetés sont de qualité.

A contrario, les techniques de *black hat* emploient des méthodes frauduleuses et sanctionnées via l'acquisition de rétroliens et utilisent plusieurs techniques : achats de liens (*paid linking*) non qualifiés et en grande quantité, utilisation de ressources tierces (annuaires par exemple), *Google Bombing*, *spamming*, s*crapping*, le contenu caché, le *keyword stuffing* (*black hat world*[388], 1re position[389], *Content marketing*[390], Semrush[391], Smart keyword[392]). Une des stratégies utilisées dans le *black hat* est notamment relative à l'insertion de liens dans l'encyclopédie Wikipedia. De par son *DA*[393] élevé, Wikipedia est de plus en plus utilisé comme référent par les acteurs qui proposent une offre de *black hat.* La stratégie est simple : utiliser les pages les plus anciennes de l'encyclopédie open source afin d'y insérer les liens *dofollow* qui redirigeront vers le site souhaité via l'utilisation d'approches permettant de dissimuler, rendre invisible les liens en modifiant par exemple le *CSS*.

Cette technique se voit corrélée à l'utilisation de sources tierces telles que les annuaires, blogs, forums ou les espaces commentaires des pages web par exemple. S'il y a quelques

---

[381] Site web de l'agence https://smartkeyword.io/netlinking-seo/

[382] Site web de l'agence https://optimiz.me/netlinking/

[383]Site web de l'agence https://www.semjuice.com

[384] Site web de l'agence https://www.boosterlink.fr/netlinking-achat-liens

[385]Site web de l'agence https://netlinka.com

[386]Site web de l'agence https://www.develink.com/solutions-achat-de-liens/

[387]Site web de l'agence https://www.netoffensive.blog/blog-seo/

[388] Accessible en ligne sur https://www.blackhatworld.com

[389] Accessible en ligne sur https://www.1ere-position.fr/blog/black-hat-seo-la-fausse-bonne-idee/

[390]Accessible en ligne sur https://www.redacteur.com/blog/seo-techniques-black-hat/

[391] Accessible en ligne sur https://fr.semrush.com/blog/seo-black-hat/

[392]Accessible en ligne sur https://smartkeyword.io/referencement-naturel-black-hat-seo/

[393] Domain Autority

années, l'utilisation des annuaires était monnaie courante, elle l'est beaucoup moins aujourd'hui, mais reste malgré tout encore appliquée par certaines agences. L'outil de l'agence Optimiz.me[394] permet entre autres (l'agence propose surtout une offre plus axée sur le référencement naturel) d'identifier les blogs ou forums avec des *backlinks* vers les concurrents d'une organisation et d'utiliser la section commentaire de ces sites pour y insérer les liens. La technique de *Google bombing*, même si elle est peu utilisée aujourd'hui, vise à s'appuyer sur des requêtes spécifiques (les plus utilisées de manière générale via notamment les outils de tendance) afin de renseigner plusieurs *backlinks* vers la page souhaitée tout en esquivant les filtres des différents moteurs. L'une des techniques les plus largement répandues (qui relève plutôt du *grey hat*) qu'utilisent plusieurs des agences spécialisées en *netlinking* et plus particulièrement dans la vente de liens, consiste en la création de *PBN* ou réseaux de sites web gérés par une même entité, et au sein desquels sont insérés les *backlinks*. La technique de *keyword stuffing* ou bourrage de mots clés consiste, comme l'indique l'intitulé, à répéter de manière intempestive les termes clés se rapportant à l'activité. L'approche par contenu ou texte masqué elle, vise à dissimuler derrière une image ou une police de taille zéro par exemple des documents sous format textuel avec une couleur identique au fond, uniquement lisibles par les *crawlers*. D'autres approches similaires sont utilisées par les prestataires : les redirections trompeuses, le *cloaking*, la création de pages satellites, le texte caché[395], etc. Ces techniques, qui comme nous le voyons, sont nombreuses[396], présentent un risque pour les organisations les utilisant dans la mesure où elles ne respectent pas la charte d'utilisation des moteurs de recherche dont l'organique et la qualité des sites et *backlinks* sont le maître mot, risques de sanctions liées notamment à des pénalités dans l'affichage des *SERPs*.

Certaines de ces techniques, nous l'avons vu, pouvant s'avérer risquées et présenter des pénalités pour les entités qui les utilisent, les prestataires proposent une offre alternative de référencement naturel (*SEO*).

---

[394] Accessible en ligne sur https://optimiz.me/netlinking/

[395] Nous y reviendrons dans les pratiques utilisées dans la gestion des avis négatifs

[396] Nous n'avons identifié que celles les plus courantes et les moins risquées

Conclusion du chapitre 5

Ce chapitre s'articule autour du marché de l'e-réputation, qui, comme nous l'avons vu, n'a pas cessé de croître depuis 2009, avec une offre, des pratiques et problématiques évolutives et encore en construction.

La première section visait d'ailleurs à présenter en avant-propos, les mutations d'une part en termes de technicité sur les approches de collecte et d'analyse des informations, et d'autre part, en termes méthodologiques sur les nouvelles problématiques et notamment les objets sur lesquels portent la mesure. Si à l'origine les questions de réputation (au sens traditionnel) faisaient en grande partie référence aux questions d'image (et les attributs associés via notamment les baromètres) et utilisaient des méthodes traditionnelles[397], le concept de réputation tel que conçu aujourd'hui, répond à des problématiques nouvelles et plurielles portant sur entre autres, les questions d'indexation, d'influence, de marque personnelle, de modération de contenus, de droit (notamment en cas de diffamation, d'atteinte à la vie privée), etc.

Ces mutations ont de fait poussé les acteurs qui évoluent dans le marché de la réputation à adapter leurs offres aux besoins des clients. Nous nous sommes dès lors proposée de définir, à la suite de plusieurs auteurs qui se sont avant nous intéressés à la question (Aref en 2009, Peggy Cadel, 2010, Camille Alloing, 2011 et l'agence Blueboat, 2011[398]), une cartographie des acteurs. Au compte de ces derniers, on retrouve les *pure players*, les éditeurs de solutions logicielles, les agences ad hoc, les agences généralistes, les applications gratuites et les *recommander sites* (sites d'avis ou de recommandation).

Si notamment pour certains acteurs (les *pure players* ou les éditeurs par exemple) la mesure de la réputation est le cœur de métier, d'autres, la majorité, proposent plutôt une approche générique axée sur les techniques et méthodes de gestion (modération de contenus, suppression ou déréférencement des contenus négatifs entre autres) et de construction de l'e-réputation dans les environnements numériques.

---

[397] À titre d'exemple les sondages, enquêtes en face à face, focus groupes, questionnaires de satisfaction, etc.

[398] Sur la base des travaux de Camille Alloing et de la liste qu'il a proposé sur son Blog https://cadderep.hypotheses.org/agences-en-e-reputation

Conclusion de la partie II

La seconde partie de nos travaux de recherche a une double visée : proposer d'une part une taxonomie des modèles de mesure de la réputation, et d'autre part, une approche en termes méthodologiques des pratiques professionnelles.
Classer les approches conceptuelles de modélisation de la réputation telles que conçues par la littérature scientifique suppose en amont de définir un ensemble d'items sur lesquels s'appuyer. Plusieurs auteurs se sont intéressés, avant nous, aux questions de réputation et plus particulièrement à leur évaluation, et ont de fait développé un ensemble de dimensions de classification étayées par des critères ou attributs. C'est notamment le cas de Sabater et Sierra (2005) ; Koutrouli et Tsalgatidou (2006) ; Marti and Garcia-Molina, (2006) ; Wang and Vassileva, (2007) ; Hoffman et al., (2009) ; Yuan et al., 2012 ; Vavilis et al. (2014) ; Hendrikx et al., 2015 ; Granatyr et al. (2015). À partir d'une revue de ces diverses approches, de l'analyse de la structure des modèles conceptuels identifiés (à l'issue de notre analyse exploratoire) et de l'analyse des méthodes théoriques de mesure de la réputation, nous avons proposé quatre (4) grandes dimensions de classification : les modèles sommaires, génériques, ad hoc et sophistiqués.

À partir de ce premier cadrage théorique, nous nous sommes, par la suite, intéressée dans le chapitre 5, au marché de la réputation en ligne. L'objectif était de distinguer les approches en termes méthodologiques mobilisées dans l'évaluation de l'e-réputation. Pour y parvenir, il nous semblait intéressant de présenter les mutations du marché de l'e-réputation. En effet, même si les questions de réputation (traditionnelle) et d'e-réputation font, selon nous, référence au même objet, il nous est apparu important de souligner les attributs qui ont mué dans le passage de la réputation à l'e-réputation. Ces derniers font notamment référence aux objets sur lesquels porte la mesure, aux techniques de collecte et d'évaluation. Nous nous sommes par la suite proposée, en nous appuyant sur des travaux existants, de cartographier le marché actuel de la réputation : les acteurs et leurs offres. En termes d'offres, les prestations vont des approches d'aide à la construction de la visibilité et de la popularité sur les plateformes socionumériques via notamment le marketing numérique, en passant par la gestion de la présence en ligne, à la mesure de l'e-réputation. Même si ce dernier point est le

plus important, nous tenons à souligner que seule une minorité des acteurs du marché (notamment les *pure players* et les éditeurs) se sont spécifiés sur la question de "mesure", et que la plupart propose une approche générale / générique de gestion, modulable selon les besoins. Cela nous a donc semblé intéressant de présenter quelques-unes des approches de gestion et de construction de la visibilité en ligne, ces dernières étant les *output* et *input* de la phase de veille et d'analyse, et l'évaluation étant un processus itératif. Sur la cartographie des prestataires, nous avons identifié six (6) catégories d'acteurs.

Il est important de souligner que la visée de ce chapitre n'était pas de nous limiter à une simple taxonomie des modèles théoriques et des approches méthodologiques professionnelles, mais plutôt de répondre à notre problématique de départ qui vise à déterminer un rapport entre savoirs académiques et pratiques professionnelles. Nous nous sommes aidée pour ce faire, et à l'issue de notre analyse des approches théoriques et professionnelles, de deux principaux outils : le tableau de classification des modèles théoriques (en annexe 1), et le tableau de classification des acteurs du marché (en annexe 2). Nous proposerons à partir de ces derniers, des éléments de réponse dans le chapitre 7. Il nous semble intéressant de présenter en amont, notre approche méthodologique et notre terrain.

PARTIE III : UNE ENQUÊTE AUPRÈS DES PROFESSIONNELS DE LA RÉPUTATION EN LIGNE

Introduction de la partie III

La dernière partie de nos travaux s'articule autour de deux principales sections : d'une part, la présentation des approches en termes méthodologiques mobilisées, et d'autre part, notre terrain de recherche. L'objectif à l'issue de cette partie est d'apporter les premiers éléments de réponse à notre problématique principale et les hypothèses qui en découlent.

Le chapitre premier présente notre double approche méthodologique (qualitative et quantifiante) dans la constitution de notre corpus d'analyse. Si l'utilisation d'une approche par entretiens semi-directifs (via un guide d'entretien) semble de prime abord logique eu égard à la nature de nos travaux, nous nous sommes également appuyée, notamment dans le processus de comparaison entre les approches théoriques et professionnelles de mesure de la réputation, d'une grille de lecture.

Le second chapitre vise, à partir du corpus constitué, à présenter nos résultats d'analyse d'une part en restituant dans leur intégralité les éléments clés des entretiens dans les six principales catégories d'analyse prédéfinies, puis, dans un second temps, à proposer notre interprétation des résultats. À l'issue de cette étape et afin de formaliser les approches théoriques et professionnelles de mesure de l'e-réputation, nous présenterons un processus d'évaluation de la réputation.

CHAPITRE 6 : LA MÉTHODOLOGIE DE TRAVAIL

Introduction du chapitre 6

Notre travail porte sur les modèles théoriques sur lesquels sont bâties les méthodes de mesure en e-réputation. Afin de répondre à notre problématique centrale et aux hypothèses de travail qui en découlent, nous avons choisi de nous appuyer sur une approche méthodologique mixte visant deux principaux objectifs.

Le premier objectif est de réaliser une analyse complète de la structure et des méthodes de mesure des modèles conceptuels de réputation, ce qui nous permettra d'identifier des items sur lesquels nous nous appuierons. Le deuxième objectif est, à partir d'une analyse du marché de la réputation, de déterminer les pratiques des professionnels en termes de techniques et de méthodes d'évaluation. Ces deux éléments d'analyse nous serviront à déceler un éventuel rapport entre les approches théoriques et professionnelles dans la mesure de la réputation.
Pour y parvenir, nous avons donc eu recours à deux approches méthodologiques complémentaires : une méthode quantitative à l'aide d'une grille de lecture et une qualitative avec l'utilisation d'un guide d'entretien.

Sur l'approche quantitative, nous nous sommes appuyée sur une grille de lecture conçue à partir des items issus de l'analyse des modèles théoriques. Elle s'articule autour de quatre (4) principales catégories : la structure générale, le contexte d'utilisation, la méthode de mesure et les données de mesure. Sur la méthode qualitative, nous avons choisi de mener une série d'entretiens semi-directifs auprès d'un échantillon de professionnels de la réputation en ligne.

1. De la nécessité d'une approche quantifiante : utilisation d'une grille de lecture

L'objectif principal de nos travaux de recherche est, rappelons-le, de déterminer l'existence d'un éventuel rapport, implicite ou objectif, entre les modèles issus de la littérature scientifique et les pratiques des professionnels en termes de mesure de la réputation. Notre travail avait de fait une double visée analytique d'une part des modèles conceptuels, et d'autre part des méthodes professionnelles servant à évaluer l'e-réputation.

Afin de déterminer le rapport entre les deux approches, il nous a semblé intéressant de nous appuyer dans une première perspective sur des catégories d'analyse comparative. En effet, à l'issue de la première phase d'analyse relative à celle des modèles conceptuels, nous avons identifié plusieurs catégories autour desquelles sont construites les méthodes théoriques de mesure. Ces catégories font notamment référence à l'architecture ou structure générale du processus d'évaluation, aux problématiques auxquelles répond la mesure, aux environnements divers au sein desquels peuvent être déployées les approches, aux techniques de mesure utilisées, à la catégorisation des informations collectées et aux valeurs finales issues de l'analyse.

Il est important, avant de présenter la grille de lecture, de noter que nous nous sommes appuyée en grande partie sur les critères de classification des modèles théoriques que nous avons identifiés dans le chapitre 4. La grille de lecture que vous trouverez en annexe n°3, vise à vérifier si les modèles théoriques de réputation ont inspiré les méthodes actuelles de mesure proposées par les praticiens, et a été utilisée en complément des entretiens semi-directifs. Elle s'articule autour de quatre (4) principales catégories : le contexte d'utilisation, la structure générale, la méthode de mesure et les données de mesure. Pour chacune d'elles, nous avons défini les sous-catégories servant à les étayer.

1.1. Le contexte d'utilisation ou l'objectif visé : la phase de cadrage

La notion de contexte pour nous, renvoie à la granularité des systèmes de réputation, soit leur aptitude à être adaptables ou non selon les contextes d'évaluation. En effet, et comme nous l'avons déjà évoqué, la mesure de la réputation fait référence à une myriade d'objets liés aux questions d'identité, d'influence, de confiance, etc., et les acteurs du marché rivalisent d'ingéniosité afin de capter de manière partielle ou totale tous ces attributs. Deux principales composantes sont prises en compte dans le concept de contexte : le type de contexte et la cible.

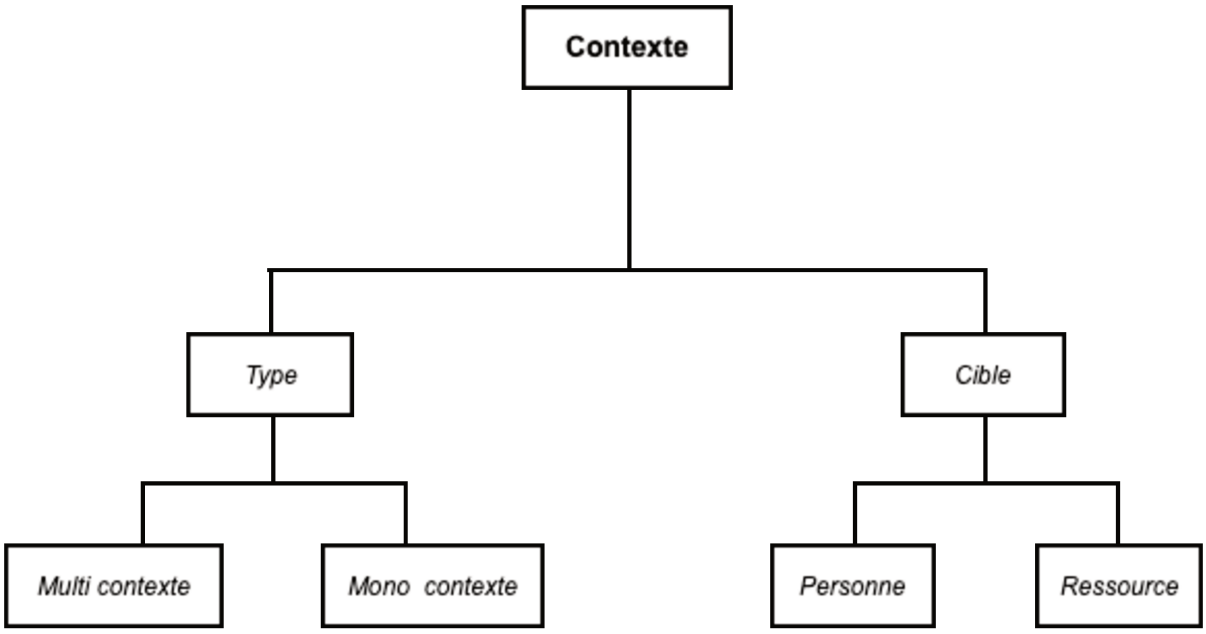

Figure n°29 : Contexte de mesure de la réputation

- Le **type de contexte** : la méthode de mesure de la réputation peut être comme nous l'avons vu holistique donc adaptable à plusieurs contextes de mesure (et faire référence aux modèles perçus dans notre classification comme génériques ou sophistiqués selon les attributs pris en compte) ou uniquement utilisée pour une analyse d'un objet spécifique de l'entité qui impacte sa réputation (comme cela peut par exemple être le cas pour les modèles ad hoc) ;

- La **cible** : sur l'entité sur laquelle porte la mesure de la réputation, il peut s'agir d'une personne (physique ou morale) ou d'une ressource (bien ou service). Si on s'intéresse aujourd'hui moins aux questions de mesure de la réputation des individus ou *personal branding*, dans certains cas, les opinions sur les parties prenantes (les dirigeants par exemple) sont souvent essentielles.

1.2. La structure (de la solution ou de la méthode)

Notre analyse exploratoire de la littérature scientifique (surtout anglo-saxonne) nous a permis d'identifier l'architecture ou la structure spécifique autour desquelles sont construits les modèles de mesure de la réputation, architecture qui s'articule autour d'une pluralité de critères. Afin d'analyser et classer les méthodes de mesure, nous proposons une structure générique comprenant 5 composantes clés : les données collectées (*inputs)*, le stockage, le traitement, le résultat (ou *output)* et le rapport d'audit (ces composantes servent de cadrage général afin d'identifier si de prime abord les méthodes de mesure proposées par les professionnels s'articulent autour de la structure que nous avons définie, et à partir de celle-ci, nous questionnerons plus en substance les diverses dimensions prises en compte).

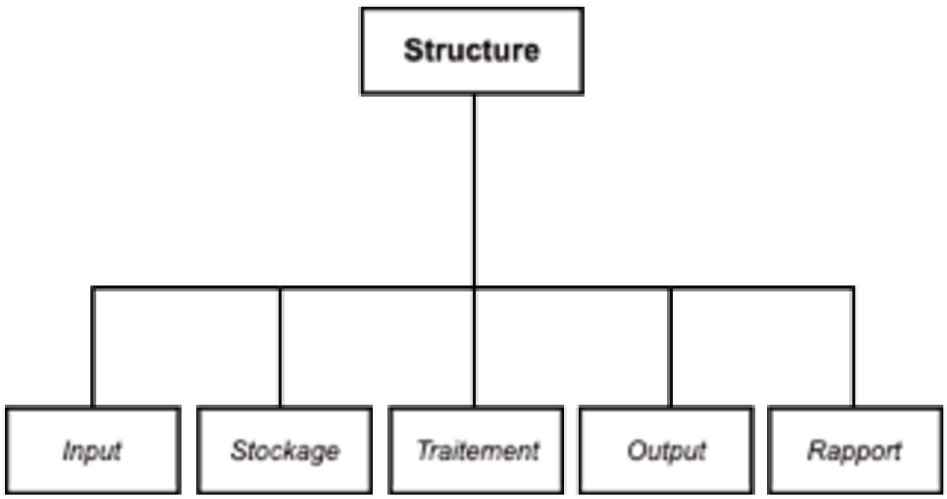

Figure n°30 : Structure générale des méthodes de mesure de la réputation

- L'***input*** : la notion d'*input* renvoie à l'ensemble des données collectées à partir des historiques de transaction (au sens large du terme et non uniquement à but marchand), d'avis ou de commentaires laissés à l'issue d'une interaction ou non, etc., et qui serviront à évaluer la visibilité d'une entité sur les diverses plateformes sociales. Il s'agit notamment d'identifier et définir les typologies de données extraites, le format sous lequel celles-ci sont extraites, les sources potentielles à exploiter (nous y reviendrons plus en substance dans la partie sur le set de données);
- Le **stockage** : dès lors que les données sont collectées, elles sont stockées dans des serveurs en vue de leur traitement. Si la question du stockage peut de prime abord être perçue comme relevant davantage de la mise en œuvre qu'un composant central (Hoffman et al., 2009), les stratégies de stockage influent néanmoins sur la structure du système de réputation. Certains modèles nécessitent une stratégie de stockage de manière aléatoire non contrôlable avec une volatilité des données privilégiant le local ; et d'autres considérés comme robustes, nécessitent un stockage permanent avec une conservation des données sur une plus longue périodicité (même pendant les périodes sans alimentation) permettant de fait de déceler les écarts ou changements de comportement lents sur de longues périodes et qui suppose de fait l'utilisation de serveurs dédiés auprès d'hébergeurs. Cet indicateur est notamment relatif aux éditeurs de solutions logicielles;
- Le **traitement** : il s'agit de l'ensemble des techniques et méthodes d'analyse des données qui répondent à une problématique de départ. Selon le but visé par la stratégie et l'objet spécifique sur lequel porte la mesure, les approches liées au traitement des données peuvent varier et s'appuient essentiellement sur trois principales caractéristiques (sur lesquels nous reviendrons plus en détail dans le prochain point : la méthode de calcul) : la technique utilisée (probabiliste, déterministe, etc.), la valeur issue de l'analyse (multiple, quantitative discrète, etc.);
- L'***output*** : se dit des résultats, des objets des tableaux de bord ou rapports, de la valeur finale attribuée à l'issue du traitement et de l'analyse des *inputs.* Comme nous l'avons déjà évoqué, le format de l'output diffère selon la quantité de données collectées, la temporalité souhaitée et l'urgence (par exemple : en cas de *bad buzz* ou crise, le

format privilégié sera une newsletter ou un système d'alertes (*alerting)* en lieu et place d'un rapport) avec une somme de variables (ou widgets) modulables et personnalisables selon les besoins des clients (nous y reviendrons dans le dernier point : le set de données collectées);

- Le **rapport d'audit** : par rendu, qui rappelons-le est distinct de l'output (le rendu est le format de présentation de l'ensemble des critères ou widgets de l'output), nous désignons la résultante de la stratégie de mesure dans sa globalité qui peut, selon les cas, être présentée sous une multitude de formats (le rapport d'audit par exemple).

1.3. La méthode de mesure utilisée

L'une des catégories que nous utiliserons également dans la classification des systèmes de réputation est la dimension de calcul. Composante essentielle dans la mesure de la réputation, c'est la représentation "*mathématique ou algorithmique*" de la réputation (pour des questions de faisabilité[399], nous tenons à souligner qu'il ne s'agit pas ici d'identifier stricto sensu les formules mathématiques utilisées par les acteurs du marché dans le calcul de la réputation, mais plutôt, à partir de concepts clés issus des entretiens et qui nous renvoient à des approches algorithmiques issues de la littérature scientifique, de proposer une classification selon l'interprétation que nous en faisons) et elle consiste en différentes stratégies et méthodes mises en œuvre afin d'obtenir les valeurs des différentes métriques de réputation. Pour ce faire, nous proposons trois principaux critères à prendre en compte dans le calcul : l'algorithme de mesure, la valeur de la réputation et la mesure de la fiabilité.

[399] Nous tenons à rappeler que nos entretiens ont été réalisés avec les personnes en charge de la gestion de la réputation des clients pour les agences, et les fondateurs/ dirigeants et products manager pour certaines solutions et non avec des développeurs de solution logicielle, d'où la difficulté voire l'impossibilité à obtenir des données sur les formules mathématiques ou algorithmiques utilisées. Nous avons de fait défini, à partir des modèles mathématiques, certains critères relatifs notamment à la technique de mesure et au format des données, afin de proposer une classification.

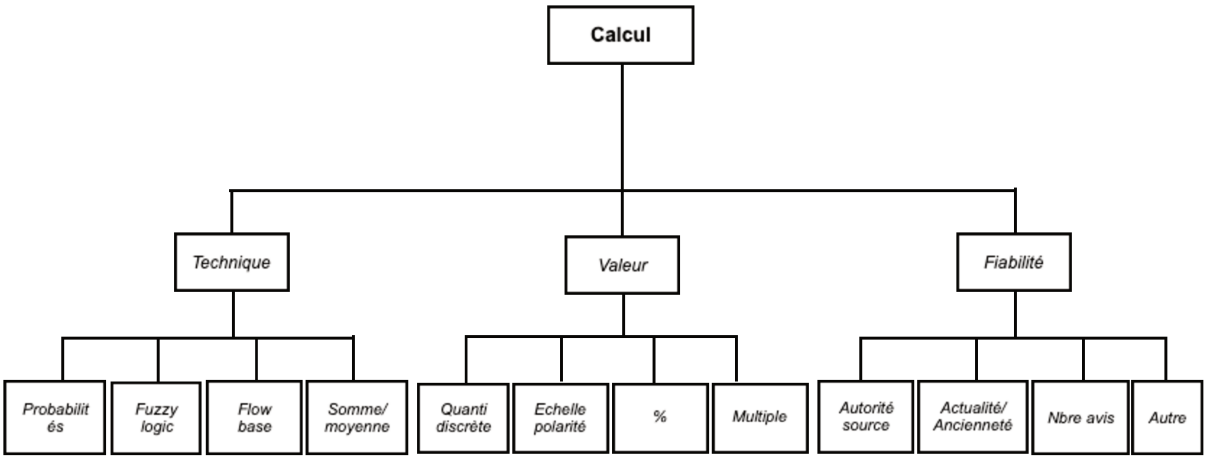

Figure n°31 : La dimension de calcul dans la mesure de la réputation

- La **technique de mesure** (l'algorithme) : notre analyse exploratoire de la littérature scientifique a révélé l'existence d'un certain nombre d'algorithmes utilisés dans les modèles, ce qui nous permet de regrouper les approches selon ces derniers (algorithmes). Au compte des méthodes utilisées, on distingue : celles probabilistes, les approches liées à la logique floue (*fuzzy logic* : intervalles bâtis sur des incertitudes), celles construites autour de la logique subjective (*subjective logic*), celles issues de la somme / moyenne de toutes les données et les *flow based* ;
- La **valeur de la réputation** : se dit du format sous lequel est présenté le résultat ou l'*output* et est intrinsèquement lié, dans une certaine mesure, à l'approche algorithmique utilisée dans le calcul. On distingue de fait selon la technique utilisée, mais aussi et surtout l'objet sur lequel porte la mesure, plusieurs typologies de variables : quantitatives discrètes, sous forme d'échelle de polarité, les variables multiples, les pourcentages et les variables de type "floue";
- La **mesure de la fiabilité** : une des valeurs essentielles dans la mesure de la réputation est le score ou indice de réputation selon les organismes de mesure, qui est attribué à partir de l'agrégation de diverses données (*inputs)*. Afin de garantir la fiabilité du score et de la valeur finale de la réputation, plusieurs critères peuvent être pris en compte dans la collecte des données : la notoriété ou l'autorité de la source, l'actualité ou l'ancienneté des avis, le nombre d'avis, caractéristiques que nous classifierons selon leur importance en 3 niveaux (3 étant le plus important).

1.4. Les données de mesure

Le *dataset* ou encore jeu de données, est l'ensemble des données utilisées dans la mesure de la réputation d'une entité. Notion essentielle dans la classification des méthodes de mesure, il fait référence aux types de données et/ ou formats de données collectées, aux sources potentielles de collecte du corpus ainsi qu'à la technique de stockage ;

- Le **type de données** : correspond au format sous lequel sont présentés les données soit textuel, visuel, au format vidéo ou encore sous forme d'emojis, etc. (que nous classifierons selon la récurrence et l'importance des cas d'usage);

- Les **sources** : nous avons défini un certain nombre de sources de collecte des *données (inputs)* qu'il convient de classer sur une échelle d'un à trois (3 étant la plus importante des sources) : les médias sociaux, les moteurs de recherche génériques et/ ou spécifiques, les portails, les blogs et forums de discussion, les plateformes d'avis, etc.;

- Le **stockage** : comme mentionné précédemment, la méthode de stockage du jeu de données peut être un élément important (certains modèles favorisant de par leur structure une forme de stockage plutôt qu'une autre). Nous distinguons à partir de là, le stockage en local ou en externe et celui aléatoire ou permanent (pas indéfini, mais sur une durée plus longue);

- l'***output*** ou le **résultat** : sur l'objet des tableaux de bord que nous avons déjà évoqué, on distingue afin de faciliter notre classification, un ensemble de métadonnées relatif notamment : à l'analyse de sentiment (*sentiment analysis)*, au score ou indice de réputation, aux données démographiques des diverses parties prenantes, à l'analyse de l'environnement concurrentiel ou encore à la détection de profils influents. Ces métadonnées rappelons-le, sont modulables et peuvent être agencées et paramétrées selon les besoins de l'analyse.

La grille de lecture, comme nous pouvons le constater, vise à répondre à la première question de notre problématique centrale : comment sont structurées, en termes méthodologiques, les approches théoriques de mesure de la réputation? La réponse à cette question nécessitait une analyse en amont des modèles conceptuels (se référer au chapitre 4), afin d'identifier les grandes catégories qui nous serviront dans le processus d'analyse comparative. La grille est construite autour de quatre (4) catégories de classification étayées par plusieurs critères. Si elle a été construite à l'issue de cette première phase d'analyse, il est important de souligner que la grille de lecture a été utilisée lors des entretiens, en complément du guide d'entretien.

2. Une approche inductive par entretiens semi directifs

De par sa nature, la recherche qualitative qui se veut exploratoire vise à appréhender un phénomène complexe tel que perçu par un interviewé de manière globale, proximale, directe et interprétative (Mucchielli, 2009) afin de co-construire le sens avec le chercheur. Paillé et Mucchielli (2012) définissent la recherche qualitative de terrain comme « *la recherche qui implique un contact personnel avec les sujets de la recherche, principalement par le biais d'entretiens et par l'observation des pratiques dans les milieux mêmes où évoluent les acteurs* »[400]. L'attribut *qualitatif* fait notamment référence autant à la typologie du matériau collecté lors de l'analyse (et qui se présente sous forme d'images, vidéos, témoignages, etc.), qu'à la méthode d'extraction des données utilisée qui se veut sémantique en respectant "*la norme de proximité*" (Paillé, 2007).

[400] Paillé Pierre, Mucchielli Alex, « Chapitre 1. Choisir une approche d'analyse qualitative », dans : , *L'analyse qualitative en sciences humaines et sociales.* sous la direction de Paillé Pierre, Mucchielli Alex. Paris, Armand Colin, « U », 2016, p. 13-33. DOI : 10.3917/arco.paill.2016.01.0013. URL : https://www-cairn-int-info.ezproxy.u-bordeaux-montaigne.fr/analyse-qualitative-en-sciences-humaines-et-social--9782200614706-page-13.htm , p.1

Il est important de noter ici que l'objectif visé par la recherche (exploratoire, descriptive, analytique, etc.) détermine la méthode de collecte à utiliser, méthodes d'entretiens qu'on peut regrouper selon trois (3) principales typologies (l'entretien dirigé, semi-directif et libre) et dont Geneviève Imbert (2010)[401] à partir des travaux de De Ketele et Roegiers (1996, p.172) nous en propose une synthèse.

Entretien dirigé (ou directif)	**Entretien semi-dirigé (ou semi-directif)**	**Entretien libre (ou non directif)**
Discours non continu qui suit l'ordre des questions posées	Discours par thèmes dont l'ordre peut être plus ou moins bien déterminé selon la réactivité de l'interviewé	Discours continu
Questions préparées à l'avance et posées dans un ordre bien précis	Quelques points de repère (passages obligés) pour l'interviewer	Aucune question préparée à l'avance
Information partielle et réduite	Information de bonne qualité, orientée vers le but poursuivi	Information de très bonne qualité, mais pas nécessairement pertinente
Information recueillie rapidement ou très rapidement	Information recueillie dans un laps de temps raisonnable	Durée de recueil d'informations non prévisible
Inférence assez faible	Inférence modérée	Inférence exclusivement fonction du mode de recueil

Tableau n°5 : Caractéristiques des typologies d'entretiens par Imbert (2010), inspiré des travaux de De Ketele et Roegiers (1996, p.172)

[401] Imbert Geneviève, « L'entretien semi-directif : à la frontière de la santé publique et de l'anthropologie », *Recherche en soins infirmiers*, 2010/3 (N° 102), p. 23-34. DOI : 10.3917/rsi.102.0023. URL : https://www.cairn.info/revue-recherche-en-soins-infirmiers-2010-3-page-23.htm

Afin de garantir cette forme de *proximité* dans la collecte des données tout en laissant le loisir à l'interviewé d'exprimer ses opinions (*avec cependant un cadre et des thématiques générales*), nous nous sommes appuyée sur une approche par entretiens semi-directifs.
L'entretien semi-directif qui suppose une démarche rigoureuse et éthique de la part du chercheur (Hopf, 2004), vise à clarifier ou étayer lors d'un dialogue ou échange avec les personnes interviewées, les objectifs et questions de départ qu'il se pose. À la différence d'un questionnaire qui cloisonne les réponses des interviewés, l'entretien semi-directif lui, se veut ouvert, nécessite une écoute attentive afin de collecter les éléments de réponses, mais aussi déceler les éléments de contexte (silences, absence de réponse, interjections, insistance sur certains éléments, etc.) et prend en compte plusieurs étapes : la phase d'échantillonnage (nous y reviendrons plus en substance dans le chapitre suivant sur le terrain de recherche), la construction du guide d'entretien et sa validation, la réalisation des entretiens et l'analyse des données. L'une des étapes les plus importantes dans la collecte des données est la construction du guide d'entretien. Nous avons, pour ce faire, identifié et défini les grandes thématiques du guide d'entretien (que vous retrouverez en annexe numéro 4) que nous souhaitions aborder avec des questions de relance en cas de silence :

- La première thématique est relative à la **perception dans une approche générale de la question de l'e-réputation et des notions connexes** : cette thématique est essentielle dans la mesure où, dans la littérature scientifique, elle est associée à une pluralité d'approches, de déclinaisons, de concepts connexes qu'il convient d'identifier;

- La seconde elle, fait référence aux **outils et/ou solutions utilisés dans la mesure de l'e-réputation** : il est intéressant de noter qu'à partir de notre analyse de la littérature scientifique sur la mesure de l'e-réputation, plusieurs tendances ont émergé sur les modèles théoriques utilisés. Afin de réaliser une analyse comparative entre ces modèles et les pratiques des professionnels de la réputation, il s'avère important de questionner les praticiens sur les composantes clés de leurs outils/ solutions : architecture, concepts clés, types/ sources de données;

- Et enfin, la troisième traite de manière plus complète **de l'offre proposée par les professionnels**, offre qui, nous le verrons, est aussi diverse que variée selon, à la fois les attentes et besoins, mais aussi et surtout le statut des entités : l'objet e-réputation nous l'avons vu, renvoie autant aux concepts d'influence, de confiance, de crise, de référencement, une pluralité de concepts que les professionnels tentent d'appréhender.

2.1. Déroulé des entretiens

Sur le déroulement des entretiens, il est important de noter qu'ils ont été réalisés sur une période de trois (3) mois soit d'octobre à décembre 2020. Après avoir défini notre échantillon, nous avons dans une première approche contacté les personnes via leurs pages professionnelles Linkedin (nous intéressant aux questions de réputation, il est nous est apparu évident d'utiliser ce canal, car la plupart des personnes rencontrées lors nos travaux sont actives sur les plateformes numériques dans l'objectif d'accroître ou maintenir leur réputation). L'objectif premier était de présenter en amont le projet de recherche, mais aussi et surtout d'établir une forme de relation de confiance entre les personnes interviewées et nous. En effet, et il est important de le souligner, la richesse du contenu collecté lors des entrevues est influencée en grande partie par la relation de confiance établie en amont, nos entretiens portant entre autres sur certains éléments quelque peu "*confidentiels*" (notamment sur la méthode de mesure utilisée, l'architecture des solutions pour ceux ayant développé leur propre outil et les exemples types de cas nécessitant une mesure de la réputation), certains interviewés nous ont d'ailleurs notifié avant les entretiens qu'il leur serait impossible de répondre à certaines de nos questions du fait de la confidentialité à laquelle ils sont tenus. Afin donc de réduire ce biais dans la collecte de nos données, nous avons envoyé à notre échantillon un mail reprenant l'objet de notre recherche, le caractère scientifique de celui-ci et non professionnel, l'objectif visé, le déroulé, mais aussi, et surtout notre engagement à anonymiser toutes les données collectées et à leur proposer une synthèse de nos travaux :

« ***Objet : Demande d'entretien téléphonique***

Bonjour M./ Mme XXXX

Dans le cadre de mes ***travaux de recherche*** *doctorale menés au sein du Laboratoire universitaire MICA (Université Bordeaux-Montaigne) sur le concept d'e-réputation et plus précisément sur les* ***modèles théoriques sur lesquels sont bâties les méthodes d'audit****, je me permets d'avoir recours à votre expertise, en sollicitant un entretien sur vos pratiques afin de recueillir votre avis de professionnel expérimenté.*

En effet, à partir d'une revue de la littérature scientifique (surtout anglo-saxonne) en e-réputation, j'ai répertorié plus de ***soixante-dix modèles théoriques*** *de mesure de l'e-réputation (Regret, Fire, Brs, Liar, Blade, Core, Fora, Travos, etc.) dont l'analyse m'a permis d'identifier les différents éléments et la structure sur lesquels ils sont bâtis. J'ai pu, à partir de cette analyse, élaborer une grille de lecture permettant de proposer une* ***taxonomie des modèles*** *et in fine d'en faire une* ***modélisation****.*

Je souhaiterais échanger avec vous sur vos pratiques en termes d'audit d'e-réputation et sur votre avis sur ces pratiques, plus largement, lors d'un ***entretien téléphonique*** *qui durera entre un quart d'heure et vingt minutes (****15 mn - 20 mn****) selon* ***vos disponibilités****. Mon objectif, de nature scientifique, est, à partir de ma grille de lecture, d'****identifier les éventuelles corrélations entre les pratiques effectivement mises en œuvre en termes d'audit et celles proposées dans la littérature scientifique dans la myriade de modèles théoriques****.*

Je tiens à vous rassurer sur le caractère des ***données*** *collectées lors de notre entretien qui seront* ***totalement anonymisées, traitées de manière confidentielle,*** *et ne seront utilisées qu'aux fins de ce travail de recherche.*

Votre participation à ce projet est tout à fait déterminante pour l'avancée de cette recherche et je vous serais extrêmement reconnaissante si vous acceptiez de partager votre expertise sur ce domaine particulier. Si tel est le cas, je serais ravie, de ***partager*** *avec vous, en avant-première, les* ***résultats de mes analyses*** *et la synthèse de ma thèse, sur la structure des modèles théoriques académiques existants, ainsi que sur les déclinaisons de ces modèles et pratiques effectives de mesure de l'e-réputation.*

Je reste dans l'attente de votre retour, que j'espère positif, et vous prie de croire à toute ma considération. Bien à vous, Liliane BARA »

Figure n°32 : Exemple de Mail / Inmail envoyé

À partir des premiers retours (dont le premier a été le jour même de l'envoi des mails), nous avons planifié les entrevues selon la disponibilité des personnes à interviewer, et nous avons par la suite envoyé des mails de relances à ceux n'ayant pas encore répondu. Avant le début de chacune des entrevues, nous avons rappelé le caractère anonymisé de l'ensemble du corpus collecté, et avons demandé l'autorisation des interviewés afin d'enregistrer les entretiens pour faciliter leur retranscription, demande qui a été acceptée par l'ensemble des personnes interviewées. Il est par ailleurs important de noter que les entretiens ne se sont pas déroulés de manière linéaire c'est-à-dire en suivant les thématiques du guide d'entretien, mais plutôt de manière spontanée, et cela en respectant les thématiques abordées de manière spontanée par les interviewés (par exemple, certains de nos interlocuteurs qui nous ont proposé une démonstration de leurs plateformes nous ont dans un premier temps présenté leur organisation puis leur solution, et au fur et à mesure de l'avancée de l'entrevue, nous utilisions le guide d'entretien et la grille de lecture comme support afin de rebondir ou spécifier certains points non abordés ou abordés de manière partielle), et les entrevues ont duré en moyenne une heure.

En amont du traitement de notre corpus, il convient de présenter la méthodologie menant à la réduction des données et aux catégories d'unités de sens.

2.2. Proposition de catégories d'analyse : du processus de réduction des données

À partir du corpus collecté et afin de proposer une lecture interprétative des données, l'utilisation d'une démarche d'analyse qualitative nécessite trois principales étapes : la réduction des données, la condensation et la présentation de ces données (Miles et Huberman, 2003) ou encore la phase de transcription-traduction, de transposition- réarrangement et de reconstitution- narration au sens de Paillé et Mucchielli (2003). Plusieurs approches peuvent être mobilisées, approches que Blais et Martineau (2006)[402] à partir des travaux de Thomas (2006, p.241) synthétisent de la manière suivante :

[402] Mireille Blais, Stéphane Martineau (2007). L'analyse inductive générale : description d'une démarche visant à donner un sens à des données brutes. Recherches qualitatives – Vol .26(2), 2006, pp. 1-18.ISSN 1715-8705 . Accessible en ligne http://www.recherche-qualitative.qc.ca/documents/files/revue/edition_reguliere/numero26(2)/blais_et_martineau_final2.pdf , p.3

Tableau 1
Comparaisons entre des approches d'analyse qualitative
(adapté de Thomas, 2006, p. 241).

	Approche générale d'analyse inductive	**Théorie ancrée**[3]	**Analyse du discours**	**Phénoménologie**[4]
Objectif général visé	Dégager les significations centrales et évidentes parmi les données brutes et relevant des objectifs de recherche.	Développer une théorie en s'enracinant dans l'analyse des données de terrain.	Fournir une description détaillée du langage et des textes en tant que pratiques sociales de même que de leur organisation rhétorique et argumentaire.	Explorer la manière dont une expérience de vie a été ressentie et comprise par la personne qui l'a vécue.
Résultat attendu au terme de l'analyse	Des catégories étant les plus révélatrices des objectifs de recherche identifiés au départ par le chercheur	Une théorie qui inclut des thèmes ou des catégories mis en relation	Un compte rendu descriptif des diverses significations dans le texte analysé	Une description cohérente et explicite du sens de l'expérience et du vécu des acteurs

Tableau n°6 : Comparaison des approches d'analyse qualitative de Blais et Martineau (2006) à partir des travaux de Thomas (2006, p.241)

À partir de ce tableau comparatif, deux principales approches semblent dans l'absolu inadaptées à nos travaux : l'analyse de discours et les approches phénoménologiques. En ce qui concerne l'analyse de discours, dans la mesure où nous ne souhaitons pas analyser les divers éléments langagiers et textuels des individus et en restituer les significations, elle s'est tout de suite avérée inappropriée. L'approche phénoménologique, même si elle a soulevé certains questionnements, n'a également pas été utilisée dans nos travaux.

La théorisation ancrée (Glaser et Strauss, 1967) elle, était, au premier abord, celle que nous souhaitions utiliser. Approche inductive favorisant l'innovation, la théorisation ancrée (*grounded theory*) vise, à partir des données empiriques, à développer une théorie sur un phénomène, et nécessite un retour et un ajustement entre les résultats de l'analyse et les données empiriques (Guillemette, 2006). Le recours à la théorie ancrée suppose une méthodologie quelque peu spécifique : l'un des concepts clés inhérent à son utilisation est la suspension temporaire du recours à des cadres théoriques existants (Opt. Cit)[403]. En effet, l'utilisation de cette approche suppose d'une part pour le chercheur de faire abstraction d'un cadre explicatif ou théorique en amont de son travail de recherche[404], c'est à dire et comme le soulignent Glaser et Strauss (1967, p.37) de faire fi des résultats des recherches déjà réalisées sur l'objet d'étude, dans le but de favoriser l'émergence de nouveaux phénomènes, de théoriser à partir du terrain de recherche et d'éviter de cloisonner, d'influencer les résultats. En un mot, tandis que dans une approche traditionnelle la revue de littérature s'effectue en amont, la théorisation ancrée elle préconise de la faire dans la phase d'analyse des données.

Ce premier postulat influe sur la manière d'énoncer l'objet de recherche qui se doit d'être dénué de toute problématique et questions de recherche avec pour cadre l'identification de paramètres du phénomène à étudier (Strauss et Corbin, 1998). L'objectif premier de nos travaux rappelons-le étant d'identifier un rapport entre les savoirs académiques et les pratiques des professionnels dans la mesure de la réputation, le choix de cette approche pourrait sembler évident. Cependant, et il est important de le noter, le recours à la théorisation ancrée suppose le développement d'une théorie (Thomas, 2006, p.241) visant à modéliser un phénomène et non simplement à identifier et à décrire les unités de sens les plus importantes qui serviront à étayer les hypothèses de travail comme c'est le cas dans nos travaux.

---

[403] François Guillemette (2006), L'approche de la Grounded Theory; pour innover? Recherches qualitatives – Vol .26(1), 2006, pp. 32-50. ISSN 1715-8705. Accessible en ligne sur http://www.recherche-qualitative.qc.ca/documents/files/revue/edition_reguliere/numero26(1)/fguillemette_ch.pdf

[404] Il est important ici de noter que cela ne signifie cependant pas faire totalement fi de toute théorie mais aborder le terrain avec le moins possible de suppositions préalables, de pré compréhensions et d'hypothèses à vérifier (Starrin, Dahlgren, Larsson & Styrborn, 1997, p.31)

Au vu de ces éléments et de nos choix épistémologiques, nous avons finalement choisi de nous appuyer dans nos travaux sur l'approche inductive générale qui, même si elle ressemble à la théorisation ancrée, ne sépare cependant pas les processus de codage ouvert (identification de sous thèmes) et axial (regroupement de ces derniers en catégories plus importantes , Thomas, 2006) et limite la modélisation à la description des catégories les plus importantes (Opt. Cit.). Comme nous l'indiquent Blais et Martineau (*2006*) « *L'analyse inductive générale est définie comme un ensemble de procédures systématiques permettant de traiter des données qualitatives, ces procédures étant essentiellement guidées par les objectifs de recherche. Elle s'appuie sur différentes stratégies utilisant prioritairement la lecture détaillée des données brutes pour faire émerger des catégories à partir des interprétations du chercheur qui s'appuie sur ces données brutes* »[405].

En effet, même si les objectifs et questions de recherche prédéfinis en amont influent en partie sur l'analyse des données de terrain, il n'en demeure pas moins que les résultats et l'interprétation qui seront faits doivent uniquement découler des données de terrain et non de postulats préétablis ou souhaités (Thomas, 2006). De par la nature exploratoire de notre travail de recherche donc, l'analyse inductive s'est imposée comme approche la plus appropriée dans la mesure où elle vise à partir des données brutes collectées lors de nos entretiens, à identifier des catégories corrélées à nos objectifs de recherche afin de les intégrer dans un cadre de référence, processus inductif qui nécessite quatre (4) principales étapes : la préparation des données brutes qui consiste à les aménager dans un format commun et à les sauvegarder ; la lecture attentive et approfondie qui vise pour le chercheur à prendre connaissance et à se familiariser avec le contenu des données (contenu qu'il peut par exemple résumer dans son intégralité ou en partie afin de faciliter la mémorisation sur un support) ; l'identification et la description des premières catégories dans laquelle il s'agit de déterminer les unités de sens les plus pertinentes dans les données afin de les étiqueter sous forme de catégories, et enfin la révision et le raffinement des catégories qui vise à attribuer au sein des catégories déjà constituées des sous-catégories (regroupant à la fois les avis convergents et/ ou divergents).

---

[405] Ibid

La phase de codification des données brutes qui amène à leur réduction est l'une des plus importantes dans l'analyse des données, car c'est à partir d'elle que sont identifiées les catégories essentielles qui viendront étayer les objectifs et questions de recherche.

Thomas (2006) à la suite de Creswell (2002) préconise l'utilisation d'un nombre restreint de catégories (entre trois et huit) afin de faciliter la lecture, avec un processus de codification en cinq étapes comme l'illustre Blais et Martineau (2006) qui reprennent leurs travaux :

Faire une lecture préliminaire des données brutes.	Identifier des segments de texte spécifiquement reliés aux objectifs de recherche.	Étiqueter les segments de texte pour créer des catégories.	Réduire les catégories redondantes ou similaires.	Créer un modèle qui intègre les plus importantes catégories.
→	→	→		
Plusieurs pages de texte	Plusieurs segments de texte	30 à 40 catégories	15 à 20 catégories	3 à 8 catégories
			→	→

Tableau n°7 : Processus de codage des données brutes de Blais et Martineau (2006), adapté de Thomas (2006, p.242), d'après Creswell (2002, p.266)

La visée finale de l'analyse inductive étant l'identification d'un ensemble d'unités de sens ou catégories à l'intérieur d'un modèle qui matérialise les données brutes collectées (Blais et Martineau, 2006), Thomas (2006) préconise cinq principales caractéristiques que doivent remplir ou posséder les catégories choisies : une forme d'«*étiquette*» attribuée à la catégorie sous forme par exemple de mot ou phrase courte ; des éléments descriptifs du contenu de la catégorie (signification, caractéristiques, etc.) ; des illustrations du contenu avec des données contextuelles (par exemple des verbatims) ; des liens hiérarchiques ou transversaux entre les diverses catégories afin de faciliter la compréhension et enfin, un cadre de référence ou modèle au sein duquel s'inscrit la catégorie.

Si nous reprenons le processus inductif d'analyse des données, dès lors que nous avons finalisé notre terrain, la première étape était la préparation des données brutes collectées.

Pour ce faire, après les avoir sauvegardées dans un dossier, nous avons renommé les entretiens puis, à l'aide d'un outil *open source* d'aide à la transcription oTranscribe[406] (qui ne transcrit pas les données, mais permet de faciliter la lecture de celles-ci), nous avons procédé à la retranscription manuelle des entretiens dans leur intégralité. Cette phase nous a permis, après avoir pris connaissance à plusieurs reprises des données retranscrites, d'avoir une vue d'ensemble des questions les plus mises en avant et récurrentes dans l'ensemble du corpus, et d'identifier, à partir des thématiques de notre guide d'entretien et celles de notre grille de lecture (*nous reviendrons sur les éléments constitutifs de la grille dans le point suivant*), les premiers extraits d'entretiens correspondants, que nous avons par la suite enrichis au fur et à mesure de l'analyse des données.

Sur la troisième étape, soit le processus de codification visant à identifier et décrire les premières catégories, nous avons d'abord défini pour chacune des catégories une forme d'étiquette servant à la nommer. Nous avons par la suite proposé pour chacune des catégories, une description synthétique de son contenu à partir des extraits servant à les étayer. À partir de ce premier codage, nous avons défini 20 catégories[407] (*que vous retrouverez en annexe n°5*). La seconde étape a consisté à réduire les données ou unités de sens que avons identifiées afin de n'en obtenir qu'un nombre plus restreint et faciliter l'analyse. Afin d'y parvenir et à partir des premières catégories, nous avons regroupé celles-ci selon leur similitude ou redondance à une des thématiques de notre objet d'étude, ce qui nous a permis d'en obtenir 10 (1.E-réputation : propositions de définitions ; 2.E-réputation et cadre conceptuel : réputation, identité, influence, confiance ; 3.E-réputation : de la construction à la gestion ; 4.E-réputation et gestion de crise ; 5.Les offres de mesure en l'e-réputation en pratique ; 6.Les livrables ; 7.E-réputation et perspectives futures ; 8.Les solutions et/ ou outils de gestion / mesure de la réputation ; 9.E-réputation et cartographie des acteurs du marché et 10.E-réputation : approche homme-machine).

---

[406] Plateforme de transcription utilisée. Accessible en ligne sur https://otranscribe.com

[407] Rappelons que dans le processus de codification extrait du tableau de Blais et Martineau (2006) inspiré des travaux de Thomas (2006) et Creswell (2002) sus-cité, cette première étape devrait permettre d'identifier entre 30 et 40 catégories (nous en avons plutôt définit 20) qui seront par la suite affinées

Enfin, et afin de faciliter l'analyse des données et produire un cadre de référence avec les catégories les plus importantes, nous en avons finalement choisi six (6) qui sont les plus représentatives des extraits d'entretien :

1. La catégorie « **L'e-réputation comme objet intangible nécessitant une approche pluridisciplinaire** » fait référence dans les extraits d'entretiens comme l'ont souligné les interviewés à la difficulté à proposer un cadrage, une définition une, unique et universelle du concept, celui-ci renvoyant à plusieurs objets (influence, confiance, identité, etc.) et déclinaisons tout en étant relatif aux besoins des clients ;

2. La catégorie « **L'e-réputation comme transposition de la réputation au numérique?** » fait notamment référence, et la plupart des interviewés s'accordent à le dire, au fait que le concept d'e-réputation n'est pas né *ex nihilo* et n'est pas un objet nouveau, mais une nouvelle forme, via notamment le numérique et ses caractéristiques, de la perception des organisations par leurs divers publics. Ce concept, se veut distinct de celui de réputation traditionnelle, propos nuancés par certains interviewés pour qui les deux concepts font référence au même objet ;

3. « **L'e-réputation, une offre sur mesure** » : lorsqu'on s'intéresse à la gestion de l'e-réputation dans sa globalité, le maître mot qui ressort des entretiens et tous s'accordent sur le sujet, c'est le "sur-mesure". Les demandes étant diverses et variées (allant des stratégies d'écoute sociale (*social listening*), à de la cartographie d'acteurs clés en passant par de l'optimisation, du référencement (*black hat seo* par exemple), les acteurs du marché proposent, selon leur statut, une offre au cas par cas ;

4. « **La mesure de l'e-réputation : une méthode en construction**», à l'exception d'une minorité d'interviewés qui disent ne disposer d'aucune méthode spécifique de mesure de la réputation (et de fait s'adapter aux besoins de leurs clients), la plupart s'accordent, sans pour autant parler de modèle (exception faite pour Linkfluence qui propose un *schéma*), à décrire plusieurs étapes quasi identiques dans le processus de mesure qu'il convient de schématiser ;

5. « **Des types de livrables : entre systèmes d'alerte (*alerting*), lettres d'information (*newsletters*), et comptes-rendus (*reporting*)** » **dans les tableaux de bord de monitoring (*dashboard*)** : il est intéressant de noter à partir des entretiens qu'il existe sur le marché de la mesure de la réputation en ligne trois principales typologies de livrables à l'issue de l'analyse, typologies qui dépendent de plusieurs facteurs : l'objectif et la nature de l'étude, la quantité de données collectées, la temporalité souhaitée, la / les solution(s) utilisée(s) ;

6. « **La mesure de l'e-réputation : un avenir mitigé ?** », cette catégorie renvoie dans les extraits d'entretiens, d'une part aux diverses restrictions en vigueur limitant l'accès aux données sur certaines plateformes (Facebook par exemple) et pouvant dans une certaine mesure constituer un frein à la mesure de la réputation, assertion cependant contrebalancée par d'autres interviewés pour qui les questions de réputation existant depuis la nuit des temps, elles connaîtront un essor indéniable avec peut-être l'apparition de nouvelles pratiques et méthodes. Comme le soulignent certains interviewés « *la crise fera toujours vivre la mesure et la gestion de l'e-réputation* ».

Conclusion du chapitre 6

Ce chapitre visait à présenter les approches méthodologiques mobilisées dans nos travaux de recherche, et qui se déclinent sous deux principales formes.

La première approche méthodologique consistait en l'utilisation d'une grille de lecture. Elle a été construite à partir de l'analyse préalable des modèles conceptuels de la réputation, et s'articule autour de quatre (4) catégories : la structure, le contexte d'utilisation, la méthode de mesure et les données utilisées.

Nous nous sommes dans un second temps appuyée sur une approche qualitative par entretiens semi-directifs. Pour ce faire, nous avons construit une grille autour de trois grandes thématiques : la perception de l'objet e-réputation et des concepts connexes qui lui sont associés, les solutions et/ ou outils utilisés dans les stratégies de mesure et notamment leur structure et les fonctionnalités essentielles, et enfin les offres de mesure proposées.

Ces deux grilles nous serviront par la suite, dans la collecte des données de terrain et leur analyse, processus que nous présenterons dans le prochain chapitre.

CHAPITRE 7 : LE TERRAIN ET LES RÉSULTATS

Introduction du chapitre 7

Il est important de souligner en amont que nos travaux de recherche se sont articulés autour de deux principaux terrains de recherche : l'analyse d'une part des approches de mesure comprises dans les modèles scientifiques et d'autre part, les méthodes proposées par les professionnels de la réputation. Le premier terrain ayant été spécifié dans le chapitre 4, nous nous intéresserons particulièrement ici au terrain professionnel.

Dans ce chapitre, il est question de présenter le processus de la collecte des données de terrain à leur analyse, et de proposer des éléments de réponse à notre problématique centrale qui vise à déterminer s'il existe un rapport entre les approches théoriques et professionnelles dans la mesure de la réputation.

Ce chapitre est composé de trois (3) sections. Dans la première, nous présenterons notre terrain de recherche, et notamment le processus d'échantillonnage (allant des éléments de cadrage au choix de notre échantillon en passant par le processus de mise en relation) avec l'identification des entités auprès desquelles nos entrevues ont été menées.

Ensuite, à partir notamment des deux grilles (d'entretien et de lecture) définies dans le chapitre précédent et des six catégories d'unités de sens choisies, nous présenterons les résultats issus de notre analyse.

La dernière section se résume à nos apports. Il s'agira de discuter les résultats issus de notre analyse en apportant les premiers éléments de réponse à notre problématique centrale.

1. Échantillonnage : d'un processus d'analyse exploratoire du web à la constitution du corpus

Nous nous proposons dans cette partie d'identifier les principaux acteurs évoluant dans le marché de la réputation en ligne et à partir d'un échantillon, d'étayer nos hypothèses de travail. Pour y parvenir, le choix de notre corpus s'est effectué en plusieurs temps relevant essentiellement de la recherche documentaire.

Notre stratégie d'échantillonnage s'est effectuée en trois (3) principales étapes : le cadrage, la constitution de l'échantillon et le traitement des données.

1.1. Le cadrage de la recherche

Dans cette première phase, il est essentiellement question de définir un périmètre général autour duquel s'articulera la collecte des informations : objet, dimension spatio-temporelle (zone géographique et durée de la collecte), les outils/ supports numériques utilisés, les requêtes choisies.

Comme nous l'avons déjà mentionné, notre étude vise à identifier d'une part les organisations de gestion de la réputation afin par la suite de déceler un éventuel lien entre les méthodes de mesure utilisées par celles-ci et celles proposées dans la littérature scientifique à travers les modèles théoriques. Pour des raisons de faisabilité et de praticité, et même si cela peut représenter, nous en sommes consciente, une limite de nos travaux, nous nous appuierons uniquement sur les entreprises qui disposent d'un siège en France (il est important de souligner que certaines sont des succursales de grandes firmes internationales notamment américaines). Le périmètre géographique est uniquement donc les entreprises françaises ou disposant d'une succursale en France.

Afin d'identifier ces dernières et pour plus d'exhaustivité dans les données, nous nous sommes principalement appuyée sur le moteur de recherche Google. Il est important de noter qu'en amont de ce travail de recherche documentaire, nous nous sommes appuyée (comme déjà mentionné dans le chapitre 5 sur le marché de la réputation en ligne) sur les travaux de plusieurs auteurs.

C'est notamment le cas de Camille Alloing qui proposait sur son blog (caddE-réputation)[408] une liste des agences en e-réputation. À partir de ce premier canevas et de l'analyse des offres de ces agences (via notamment leurs sites corporate, livres blancs et articles de blogs), nous avons identifié celles qui proposaient une offre de mesure ou plus généralement de gestion de la réputation et qui étaient encore actives (certaines agences ayant malheureusement cessé leur activité).

Nous avons par la suite, et afin d'actualiser la liste des agences et garantir une forme de représentativité de celles-ci, utilisé le moteur de recherche Google que nous avons questionné à l'aide de plusieurs mots clés sous forme de requêtes. La construction des requêtes est essentielle dans la collecte des données, car de leur pertinence et précision dépendra la qualité des données. Afin de garantir une ouverture dans la collecte des données et ce de la manière la plus exhaustive possible, mais aussi et surtout de réduire les "*bruits*" ou informations ne présentant pas un grand intérêt pour nos travaux, nous avons identifié les concepts clés inhérents à la mesure de la réputation (concepts issus de la littérature scientifique et de notre analyse exploratoire du web social sur la question) auxquels nous avons appliqué certains opérateurs booléens : « e réputation agences » or « agences e-réputation » , « mesure e-réputation » or « veille e-réputation » , « audit e-réputation agence » avec des variances (tiret après le e ou pas par exemple). Nous avons par ailleurs utilisé de manière concomitante un système « *d'alertes e-mail* »[409] afin d'agréger sur l'ensemble du web social (et non uniquement sur les moteurs de recherche) toutes les données relatives à notre objet d'étude.

---

[408] Blog de Camille Alloing qui répertorie les agences qui composent le marché de l'e-réputation. Accessible en ligne sur https://cadderep.hypotheses.org/agences-en-e-reputation

[409] Utilisés dans une démarche de veille, les systèmes d'alertes collectent et ce à partir de mots clés prédéfinis, les informations relatives à un "objet" spécifique et les envoie par mail à une fréquence choisie (journalière, hebdomadaire, etc. Nous avons utilisé ici le principal système d'alerte affilié au moteur de recherche google qui est Google alerte https://www.google.fr/alerts

1.2. La constitution du corpus

À l'issue de cette phase de cadrage et à partir de notre analyse exploratoire du web, nous avons identifié 207 organisations. Notre population est de fait composée de l'ensemble des acteurs ayant un siège en France et spécialisés dans la gestion de la réputation dans une approche très large (et dont nous avons proposé une taxonomie dans le chapitre 5). Ces acteurs, que nous avons regroupé sous six (6) principales catégories, sont composés des *pures players* ou organisations spécialisées et dont le coeur de métier est la mesure de la réputation (avec le développement en interne d'une solution propriétaire pour certains), mais aussi d'acteurs qui proposent une offre plus générale dans les domaines de la communication (digitale), de la publicité, des relations publiques ou du marketing (digital), et qui possèdent un service dédié à la mesure et à la gestion de la réputation. Plusieurs questions se sont posées à nous dans l'approche à choisir afin de constituer notre échantillon. Notre étude visant à la fois une approche qualitative, mais aussi et surtout quantifiante via notamment la grille de lecture, il nous est apparu plus opportun de cibler dans une première approche les acteurs dont la mesure de la réputation est le cœur de métier.

Nous avons dans un premier temps effectué une analyse de contenu des supports corporate des différentes organisations (avec le site web comme support central) et une vérification de leur activité (et des données financières notamment leur chiffre d'affaires sur société.com) afin d'identifier pour chacune, si elles sont encore en activité, mais aussi et surtout le positionnement corporate de l'organisation, les offres ou prestations effectivement proposées. Il nous apparaissait important de réaliser ce travail de vérification en amont, car certains écarts existaient entre les offres supposées de certains acteurs du marché et les prestations réelles proposées. Malgré ce travail de vérification en amont, nous avons tout de même rencontré certaines difficultés lors de nos entretiens, certaines organisations qui, sur leurs plateformes, proposaient pourtant une offre de mesure, nous ayant finalement déclaré de manière implicite ne pas proposer d'offres de mesure au sens où nous l'entendions, mais des services connexes (nous reviendrons sur ce point dans l'interprétation et la discussion des résultats).

À partir de ce travail, notre échantillonnage s'est effectué en deux temps : nous avons d'abord, à partir de l'analyse des offres et de notre cartographie des acteurs, appliqué un premier filtre qui nous a permis d'identifier une première catégorie d'acteurs composée essentiellement de *pure players* (acteurs dont la mesure de la réputation est le coeur de métier) et développeurs de solutions logicielles.

Nous avons par la suite et dans un souci de représentativité, identifié une seconde typologie d'acteurs au sein de notre population. Ce dernier échantillon regroupe les acteurs généraux (évoluant dans la totalité des sous-disciplines identifiées ci-dessus) qui proposent une offre générale de gestion de la réputation et dont la mesure est proposée de manière implicite en amont des stratégies de gestion.

1.3. La phase de traitement : de la prise de contact à la collecte

Après avoir défini notre échantillon, nous avons identifié pour chacun des acteurs, la personne ressource la plus à même de répondre à nos interrogations sur l'objet de recherche. Pour les agences généralistes ou spécifiques, il s'agissait entre autres des : directeur(trice) du pôle e-réputation, consultant(e) social média / e-réputation, chef de projet e-réputation, consultant e-réputation et influence, web intelligence and *social listening analyst*, chargé(e)s de comptes veille et réputation, social média manager, analyste e-réputation, chargé(e) d'études image et notoriété, consultants seo et e-réputation ; et pour les développeurs de solutions logicielles les *Ceo*, VP *solutions consulting*, *senior solution consultant, key account manager, strategic account executive*. Il est important de souligner que nous avons utilisé pour la collecte des données de contact et la prise de contact avec les personnes ressources les supports corporate des diverses organisations (sites web entre autres où sont présentés les divers profils, webinaires avec présentation des intervenants par exemple), mais aussi et surtout la plateforme professionnelle Linkedin.

Nous avons par la suite envoyé aux personnes qui ne figuraient pas dans notre réseau un premier mail de mise en relation avant de leur envoyer le mail de présentation de nos travaux et de demande de rendez-vous pour les entretiens.

Si certaines demandes de mise en relation n'ont d'emblée pas été acceptées, d'autres qui l'ont par ailleurs été n'ont par la suite pas donné lieu à des réponses favorables à nos demandes d'entretiens, et ce, même après avoir envoyé les mails de relance, et au final, nous avons reçu 19 réponses favorables qui ont donné lieu à des entrevues et réparties comme suit :

- Au compte des éditeurs de solutions logicielles, nous avons : Digimind, Kb crawl, Linkfluence, Sindup, Synthesio et Talkwalker ;
- Les pure players et agences ad hoc en e-réputation sont également représentés dans notre échantillon avec notamment : l'agence Atchik, Altyveille, blueboat, Human to human, Netino by webhelp, Net Offensive et Net réputation, Opinion act ;
- Et enfin, la dernière typologie d'acteurs comprise dans notre corpus est relative aux agences généralistes qui proposent une offre de réputation : Cision Insights, Firstlink, Otta, Rankwell et Vanksen.

2. Présentation des résultats et analyse des données

2.1. L'e-réputation, un objet intangible nécessitant une approche pluridisciplinaire

S'il est une chose sur laquelle tous s'accordent d'entrée de jeu et l'un des premiers éléments clés qui ressort de nos entretiens, c'est la difficulté des interviewés à proposer un cadrage "théorique", à définir de prime abord le concept d'e-réputation :

> *« C'est une très bonne question. Je ne pense pas qu'on ait de définition déjà réellement partagée [...]. À un moment, c'était posé quand je dis un moment c'est il y a 7-8 ans sur une définition de l'e-réputation à l'époque où les problématiques d'e-réputation étaient très liées aux réseaux sociaux. Mais du coup, après ce n'est pas vraiment une notion qu'on définit, sur laquelle on se penche concrètement d'un point de vue théorique »*[410]

[410] Extrait de l'entretien avec l'Interviewé 8

Cette difficulté à définir l'objet e-réputation est relatif au fait qu'il recouvre une myriade d'objets, de significations et de déclinaisons dû d'une part aux attentes et besoins des clients, mais aussi, et surtout d'autre part au statut des organisations qui en proposent la mesure comme l'a souligné l'un des acteurs majeurs du marché :

> *« Sur les définitions de la réputation, globalement je sais qu'il en existe une multitude - nous on est assez proches en fonction des clients avec lesquels on travaille , on est plus dans sa définition traditionnelle dans l'univers de la com corporate et puis beaucoup autour des questions marketing, qui va aussi être assez présente. Encore une fois, on ne manipule pas de définition unique et stabilisée de la réputation, c'est un peu à la main de nos clients* ». [411]

Comme nous le voyons, ces premières propositions de cadrage du concept d'e-réputation mettent déjà en exergue l'existence d'une myriade de définitions. Que les entités qui en proposent la mesure aient une approche plutôt axée sur la communication digitale, le web marketing, l'e-influence ou les relations publiques/ presse 2.0, etc., le concept d'e-réputation prend une forme nouvelle qui ne diffère cependant pas des autres, mais qui lui sont plutôt complémentaires. L'une des approches qu'on retrouve donc à partir de là et de manière plutôt récurrente est celle tournée vers le marketing et la communication corporate (digitale), où le concept d'e-réputation fait référence et renvoie aux questions d'image de marque et de leur gestion, nécessitant une approche pluridisciplinaire comme le souligne l'interviewé 1 :

> *« La réputation c'est quelque chose de totalement, en réalité pour faire de la philosophie c'est quelque chose de totalement intangible. Justement dans la définition même c'est un actif intangible d'une marque donc la valeur de marque dans notre société. Vous me demandiez au début la définition que je pouvais avoir, je pourrais en effet introduire celle-ci, mais ces actifs aujourd'hui d'image, ça fait plus de 50 ans qu'on essaie de les mesurer et évidemment personne n'y est totalement arrivé parce que c'est extrêmement fuyant, pas matériel, et c'est extrêmement compliqué à mesurer donc évidemment ces objets complexes vont nécessiter d'avoir une approche pluridisciplinaire et a priori devrait forcément profiter d'approches multidisciplinaires afin de l'attaquer sous de multiples angles. C'est typiquement ce genre d'objet que vous ne pourrez jamais appréhender d'un seul point de vue. Vous êtes obligé de multiplier les points de vue pour réussir à les circonscrire.* »

[411] Extrait de l'entretien avec l'Interviewé 1

Il est intéressant ici de noter le rapport étroit entre les questions d'e-réputation et d'image que soulignent plusieurs de nos interviewés. En effet, de par les diverses "*inter(actions)*" des entités sur le web (à la fois sur leurs espaces propriétaires, mais aussi et surtout sur le web social de manière plus générale), ces dernières produisent de manière volontaire et/ou involontaire des marqueurs ou attributs identitaires qui contribuent à façonner une forme d'image de marque, image qui n'est autre que le reflet de leur e-réputation.

Lorsqu'on s'intéresse donc à la mesure de l'e-réputation et plus précisément à "l'objet" sur lequel porte la mesure, tous s'accordent sur la notion "d'image" que véhicule l'organisation sur les divers espaces numériques. La mesure de la réputation suppose donc comme le notent les personnes interviewées une prise en compte générale de l'ensemble des avis, notes et données collectées portant sur l'organisation dans sa globalité et plus précisément influant sur la perception de son image auprès de ces divers publics :

> *« Quand les clients demandent des audits, le terme est un peu rare, il est un peu daté maintenant, mais quand les clients le demandent c'est des audits d'image, c'est d'avoir une vue d'ensemble sur l'ensemble des conversations, des contenus éditoriaux des publics accessibles sur le web social et de les hiérarchiser et les ordonner pour se construire une image de ce que cela communique de leur image, de la perception que ces sources ont des marques* »[412]

Cette vue d'ensemble de l'image de l'organisation porte autant sur cette dernière (l'entité), la marque, ses produits/ services et les diverses parties prenantes, et pour paraphraser l'interviewé 1 *« le numérique a apporté ce côté... la réputation aujourd'hui elle est unifiée, rassemblée, donc il n'y plus de distinction c'est à la fois la parole et l'expression des consommateurs, des citoyens, des journalistes, des collaborateurs, des politiques, toutes ces dimensions de la réputation qui sont en fait les perceptions des différentes parties prenantes se retrouvent unifiées, rassemblées aujourd'hui sur le web. Donc quand ils nous demandent d'étudier leur réputation ils ne font pas de distinction, c'est très rare qu'on nous demande un bilan d'image corporate par exemple, on nous demande d'avoir une vue d'ensemble sur tout ce qui est exprimé sur le web et qui est tourné, impacté et influence la perception de l'entreprise dans son ensemble* ».

---

412 Extrait de l'entretien avec l'Interviewé 1

Cette première approche axée communication / marketing se voit rattachée à celle liée au concept de crise et de sa gestion dans une approche plus générale comme nous l'indique l'interviewé 2 :

> « *En général, si je me permets d'avancer un petit peu le propos, on parle d'e-réputation, on pourrait surtout parler d'e-réputation de crise puisqu'en général, les entreprises font appel à un gestionnaire de réputation quand elles ont je ne vais pas dire des casseroles parce que ce n'est pas le mot, mais pour faire..., c'est plus une visée préventive que généralement du suivi sur des points positifs. Je ne sais pas si vous me comprenez, ce que je veux dire... mais c'est plus pour vérifier que personne ne s'attaque à la marque, ne parle mal de la marque ne dénigre la marque ou l'entreprise ou les composantes, ça peut être les personnes qui en font partie, plus que du suivi d'une annonce.* »

On retrouve cette propension à associer e-réputation et gestion de crise chez nombre de nos interviewés pour qui les questions d'e-réputation et leur mesure visent deux principaux objectifs : une anticipative et l'autre à visée curative ou de suivi. Dans sa forme anticipative, la veille en e-réputation vise entre autres (celle-ci pouvant avoir divers objectifs) à détecter d'éventuels signaux faibles pouvant porter atteinte à l'image d'une entité. Il s'agit donc dans ce cas de figure de surveiller, veiller, contrôler et être alerté instantanément via notamment les outils d'alerte (*alerting*)[413] de tout contenu à teneur plus ou moins négative sur l'entité, ses produits / services, dirigeants, parties prenantes, etc. Dès lors que des mentions sont donc faites sur l'entité, elles sont traitées en interne et selon leur pertinence et polarité, envoyées directement via mail à l'entité telle quelle « *En général nous on alerte, on donne un élément de réponse, répondez, ne répondez pas et puis voilà* »[414] ou, et selon la prestation souhaitée, certains acteurs mettent en place une stratégie de suivi avec un tableau de bord (*dashboard*) reprenant les divers éléments de réponse, la temporalité, etc.

---

[413] qui est un des supports utilisé dans la mesure de la réputation et dont nous reviendrons plus en substance dans les livrables

[414] Extrait de l'entretien avec l'interviewé 2

C’est notamment le cas de ceux qui disposent d'outils de modération de contenu leur permettant de gérer automatiquement les publications de leurs clients via les diverses plateformes de réseaux sociaux (social média) : « *Alors nous on a un outil propriétaire [...] en fait la modération (je ne sais pas si vous voyez un peu en quoi ça consiste), mais on reçoit les commentaires de nombreux espaces de certains de nos clients et on est en charge de les publier ou non en fonction de chartes qui sont définies conjointement avec nos clients* ».[415]
Ces outils de modération permettent donc de gérer les contenus provenant à la fois des réseaux sociaux tels Facebook, Linkedin, Youtube, Twitter[416] et grâce aux partenariats médias, de gérer les sites web corporate des entités (et plus spécifiquement dans la section commentaires des divers articles publiés). Le second volet, lui, fait plutôt référence à un suivi régulier via notamment une veille quotidienne afin de pouvoir en cas de crise vérifier les résultats du modèle de suivi déployé dans sa résolution ou identifier dans le cas contraire de nouveaux manques (insights) visant à améliorer l’image de l’entité auprès de ses diverses parties prenantes sur le web.

Cet aspect d’actif intangible se retrouve également dans la propension à associer le concept d’e-réputation à celui d’influence avec notamment la question des liens hypertextes. En effet, les liens (entrants et sortants que nous avons abordés dans le chapitre 1 de la revue de la littérature) entre les diverses pages web et donc les mentions sur celles-ci sont un élément clé dans le degré d’influence qui leur est attribué comme nous l’indique l’interviewé 7 : « *Déjà c'est toute cette notion de liens qu'il va y avoir entre chaque site web et donc de mentions, qui va déterminer la réputation des entreprises, enfin d'une entité. C'est toute la dimension de pouvoir taguer en fait ces acteurs-là pour en comprendre les comportements et également c'est les « rankings » des scores qu'on va aller chercher en fonction des besoins des clients. Si c'est de la réputation sur les réseaux sociaux on sera plutôt sur du « ranking social », et si on est plutôt sur du marché web donc médias, forums, blogs, etc. on sera plutôt sur du « ranking de backlink », de notoriété et de trafic* ».

---

[415] Extrait de l’entretien avec l’interviewée 6

[416] Sur Twitter qui pour la plupart des acteurs est l’une des sources réseaux sociaux principale de collecte des données la modération n’existe pas vraiment comme nous l’ont indiqué certains interviewés car tout simplement elle consiste à masquer les contenus “indésirables” lorsqu’on en est l’auteur

Ce concept d'influence est souvent, pour certains de nos interviewés, substitué à celui de popularité notamment de visibilité sur les moteurs de recherche, avec une approche axée webmarketing. Si, *a priori,* on ne serait pas tenté de les classifier comme relevant des stratégies de réputation, elles répondent à de nouveaux enjeux et besoins des entreprises comme nous l'indique l'interviewé 9 : *« Pour une très grande partie des entreprises, pour 90% des entreprises, je n'ai pas les chiffres exacts pour le dire, mais l'e-réputation, c'est beaucoup plus simple que de l'audit. {...}. Nous, notre positionnement, vous le trouvez original alors qu'il est en réalité raccord avec les besoins basiques des entreprises. Il y a 3 millions d'entreprises en France (je ne sais pas le nombre exact) et elles, ce qu'elles veulent, c'est quand on tape leur nom dans Google, elles veulent avoir une bonne image (de bons résultats). Elles veulent avoir de bonnes notes sur les sites d'avis et éventuellement, si elles peuvent avoir plus de followers sur Twitter, Instagram, Pinterest, Tik Tok, etc., et pourquoi pas des sentiments positifs à travers ce que les gens partagent en fait sur ces marques éventuellement, elles seront contentes* ».

Cette approche qui s'appuie sur des techniques de *SEO*, vise d'une part à positionner une entité dans les premiers résultats des recherches effectuées, mais aussi, et surtout à gérer les contenus négatifs qui émergent et entachent l'image de celle-ci :

« Ce qu'on fait, c'est très simple... Nous on peut faire au choix, selon les clients, on peut améliorer les contenus qui sont référencés, qui vont apparaître rapidement avec des techniques extrêmement légales. Ce n'est pas illégal! Tout ce qu'on fait est légal, mais extrêmement compatible avec les règles imposées par le dictat de Google, ou au contraire, on peut améliorer les avis. Améliorer les avis, ça veut dire qu'on va aller voir des vrais clients satisfaits de notre client, en leur demandant de témoigner et de laisser des avis là où on a besoin de récupérer des avis positifs.

Ou, là pour le coup, on peut créer de faux avis, et des faux avis assez indétectables. Parce qu'en fait, nous on a créé toute une série de faux profils qui ont un vrai historique, qui laissent un avis sur tout, et sur rien... Par exemple, j'ai un faux profil qui donne des avis sur tous les restaurants que je teste. Donc du coup, on peut les relier à de vraies expériences. La plupart du temps, les avis sont positifs, mais je préfère avoir des profils qui disent plutôt des choses positives pour que quand j'ai besoin de dire des choses positives sur un client, que ce soit en accord avec "la ligne éditoriale" du profil. Donc, on peut faire du white ou du black hat pour améliorer soit les avis, soit les pages de résultats {...}. Sachant que quand on diffuse des contenus, ce n'est jamais du mensonge. Même quand on fait du black hat, on ne va pas par exemple pouvoir diffuser des articles sur nos propres profils.

On a créé en 15-20 ans, un réseau de plus de 600 sites. Et si vous tombez dessus, vous n'avez aucune idée qu'ils ont été créés par une agence comme nous dans le but essentiel d'améliorer le référencement des marques de nos clients »[417].

Sur les techniques utilisées, il est intéressant de noter que certaines comme nous le voyons notamment de "*black hat seo*" vont parfois à l'encontre des chartes des moteurs de recherche et peuvent faire l'objet de pénalités lorsqu'elles sont détectées.

Le dernier concept connexe et indissociable du concept d'e-réputation, qu'on retrouve dans nos entretiens à la suite de ceux sus-cités (image, influence, crise), et dont ne nous ferons qu'un retour synthétique l'ayant déjà abordé dans la littérature scientifique est celui de confiance. S'il est une chose essentielle dans le cadrage du concept de réputation en ligne, c'est que la confiance est une conséquence (*output*) de la réputation comme nous l'indique l'extrait suivant :

> *« Elles sont évidemment étroitement liées. C'est-à-dire que la confiance naît de la réputation simplement. C'est-à-dire qu'un déficit, une mauvaise réputation n'appellera pas la confiance. C'est artisanal ce que je dis, mais une mauvaise réputation n'appellera pas un bon niveau de confiance et à l'opposé, une bonne réputation crée et supposera un bon niveau de confiance sachant qu'une bonne réputation c'est une perception a priori. »*[418]

Si ce rapport étroit entre confiance et réputation a déjà fait l'objet de nombreuses études dans la littérature scientifique tant en sociologie, psychologie qu'en gestion, nous retiendrons que la confiance tout comme la réputation se construit à travers le temps et les interactions. C'est un construit social qui est aujourd'hui très affecté par le caractère mimétique des plateformes sociales et l'essor des stratégies d'influence qui font qu'on sera plutôt enclin à "penser, agir" selon une certaine "norme" dictée par une minorité ou majorité (selon le contexte dans lequel on se situe).

---

[417] Extrait de l'entretien avec l'interviewé 9, qui dirige une agence spécialisée dans la gestion des contenus négatifs et l'aide à la construction de la visibilité des entités sur les moteurs de recherche

[418] Extrait de l'entretien avec l'interviewé 1

Après, et pour paraphraser l'interviewé 1, « *une fois qu'on est rentré dans une relation et qu'on a une expérience effectivement, la confiance va devenir de plus en plus indépendante de la réputation*[419]*».*

2.2. L'e-réputation comme transposition de la réputation via l'interface numérique?

Lorsqu'on questionne les interviewés sur l'objet e-réputation, l'une des notions connexes qui ressort, et ce de manière systématique et récurrente, est celle de réputation *traditionnelle*. Ces deux concepts même si nous l'avons déjà noté dans la revue de littérature sont liés peu ou prou, tous s'interrogent sur le lien qui existe entre les deux.

La première approche la plus généralement partagée considère l'e-réputation comme le reflet, via le web social et ses diverses caractéristiques, du concept de réputation traditionnelle comme nous l'indique l'interviewé 6 :

> « *Je pense que l'e-réputation est le reflet de la réputation globale d'une structure parce qu'il y a peu de chances qu'une structure bénéficie d'une mauvaise e-réputation, mais que sa réputation plus classique, enfin globale, soit bonne. Donc ça va vraiment être un reflet de ce qui se passe dans la entre guillemets "vraie vie" même si on voit que plus ça va, plus le web social fait partie de la vraie vie. »*

Ce premier postulat souligne le fait que là où la réputation qui utilisait des canaux traditionnels tel le bouche-à-oreille sera beaucoup moins diffusée, l'e-réputation, elle, prendra une ampleur plus considérable. Ce qui d'une part s'explique par les caractéristiques propres au web social, mais aussi, et surtout par la mutation des techniques et méthodes de collecte de données ou avis comme nous l'explique l'interviewé 5 :

---

[419] Ibid

« Avant en fait, dans un monde ancien, ils faisaient un baromètre d'image ou voilà on interrogeait 150 ou 1000 personnes et on disait " quelles sont les marques que vous connaissez ne serait-ce que de nom, est-ce que celle-là vous la connaissez, etc.", et ensuite on balançait toute une batterie d'image : « est une marque bienveillante, innovante, etc. », et du coup on faisait une cartographie la plus dynamique possible pour pouvoir suivre la réputation et l'associer à des activations qu'on avait pu faire en média hors média, etc. Je pense que quelqu'un qui veut faire du suivi de la e-réputation, il répond... il aspire un peu à ça, mais de façon beaucoup plus dynamique, beaucoup plus immédiate. »

La mesure de l'e-réputation consisterait donc, à partir de ce baromètre de réputation traditionnel, à collecter, en temps réel et dans l'urgence, des données sur une entité de la part de ses diverses parties prenantes, données qui d'une part ne sont plus à aller collecter via des questionnaires par exemple comme on pouvait le faire auparavant, mais avec en sus, la question de la temporalité qui est très prégnante :

« *Il ya une notion de temporalité qui n'est pas la même, mais surtout la chose qui est la plus flagrante c'est que la donnée n'est plus administrée en fait. On est sur de la donnée spontanée en fait, c'est à nous justement de la restructurer derrière*.»[420]

Si pour certains donc l'e-réputation est la transposition vers le numérique de la réputation ce qui suppose une distinction (quoiqu'infime entre les deux concepts), pour d'autres, les deux concepts ne sont en réalité qu'un seul et même objet et certains vont même encore plus loin avec le postulat suivant :

« La e-réputation ça n'existe pas en fait. Il n'y a qu'une seule réputation et l'erreur qu'on a faite en créant le terme e-réputation c'est de considérer qu'il existait une réputation dans la vie réelle (IRL) et une réputation online alors que c'est totalement faux. Il y a une réputation point, et aujourd'hui, celle-ci peut être fractionnée influencée par les prises de parole sur les réseaux sociaux et sur les espaces en ligne »[421].

---

[420] Extrait de l'entretien avec l'Interviewé 5

[421] Extrait de l'entretien avec l'Interviewé 1

Considérer que l'e-réputation est la transposition sous une forme nouvelle de la réputation traditionnelle revient dans une certaine mesure à identifier deux objets distincts, ce qui selon nous n'est pas productif dans la mesure où l'e-réputation se veut générale, globale, avec une approche pluridisciplinaire comme susmentionnée, nécessitant l'utilisation de plusieurs techniques et méthodes de collecte des données. Si on s'intéresse dans une première approche aux sources, méthodes et techniques de collecte de données et plus particulièrement aux baromètres d'images, panels ou questionnaires en face à face visant autrefois à collecter l'avis des usagers sur une entité, une marque, un produit ou service et évoqué par certains de nos interviewés, on se rend compte que ces derniers n'ont pas disparu comme ces derniers pouvaient le penser, mais utilisent aujourd'hui de nouvelles formes.

C'est notamment le cas des sites d'avis clients et autres organismes de confiance qui à l'issue d'une transaction ou interaction avec une entité, ont pour rôle de collecter les évaluations (insights) sur la transaction de manière générale (l'entité, le produit ou service, la qualité du bien ou de l'interaction, les délais de livraison dans le cas d'un produit, le service après-vente, etc.). Avis qui peut donc prendre une forme multiple selon la transaction et l'organisme chargé de sa collecte : si dans certains cas il s'agit d'attribuer simplement une note (le plus souvent sur 5) comme cela peut être le cas pour les enquêtes de satisfaction réalisées par les organismes de confiance et sites d'avis (il est important de noter qu'il existe deux typologies de sites d'avis : ceux vérifiés qui enregistrent les avis issus d'une transaction confirmée et ceux non vérifiés où tout le monde a la possibilité sans avoir eu d'interaction avec la marque, le produit ou le service de laisser un avis : exemple de Trustpilot), d'autres prennent la forme d'un questionnaire simplifié portant sur les éléments clés de la transaction à évaluer (pouvant également inclure une note) : c'est notamment le cas pour ceux provenant des services internes, du service client des organismes ou structures avec lesquelles les internautes interagissent ou ont des transactions.

Certains organismes vont encore plus loin et réalisent des enquêtes plus complètes et font notamment appel à des cabinets d'étude spécialisés sur les questions de réputation (Ifop, Ipsos, etc.).

Ces sites de collecte d'avis de clients, rappelons-le, n'existaient pas encore il y a quelques années, et sont encore balbutiants aujourd'hui, mais ils deviennent des sources incontournables de la mesure de la réputation avec qui les acteurs du marché (en particulier les développeurs de solutions) développent de plus en plus de partenariats : « *Nous sommes justement actuellement en train de finaliser le partenariat avec "avis vérifiés", ce qui nous permettra de connecter la source à notre plateforme afin de pouvoir crawler les avis clients et nous travaillons à développer de nouveaux partenariats* »[422]

Même si notre postulat est que les questions de réputation et d'e-réputation ne sont qu'un seul et même objet nécessitant de facto une approche pluridisciplinaire rappelons le, il n'en demeure pas moins que le numérique a modifié les pratiques de mesure avec la prise en compte aujourd'hui de nombreux facteurs liés à l'instantanéité, à l'urgence, la quantité exponentielle de données disponibles, à l'apparition de nouveaux phénomènes tels les fake news, etc., d'où le recours à une adaptation continuelle comme nous l'indique l'interviewé 6 : « *Au-delà de répondre aux besoins du client, on est obligé de répondre aux évolutions technologiques tout simplement quand on travaille dans une entreprise qui est spécialisée dans le numérique* », adaptation aux nouvelles techniques et méthodes de collecte et d'analyse des données, aujourd'hui indispensable pour tout acteur du marché de la mesure de l'e-réputation et qui ne saurait être un facteur visant à distinguer la réputation de l'e-réputation.

2.3. L'e-réputation, une offre sur mesure

À ce stade, force est de constater à la suite des entretiens qu'il n'existe pas d'offre une et unique dans la mesure de la réputation, celle-ci comme vu précédemment pouvant faire référence à divers objets et donc relative aux besoins des clients. Si pour certains la mesure de la réputation s'adresse et prend la forme de ce qu'on dénomme la mesure du *personal branding* ou de la réputation personnelle (une minorité des personnes interviewées dans le cadre de notre travail de recherche propose cette offre cumulée avec l'offre adressée aux organisations, mais elle existe sur le marché de la mesure de la réputation) la plupart propose

[422] Extrait de l'entretien avec l'interviewé 12

une offre qui s'adresse aux personnes morales (entreprises / organisations, PME en général) et pour les leaders du marché dont les développeurs de solutions logicielles, une offre à destination des grands comptes : *« Nous on est spécialisé dans les grandes marques. La social intelligence pour les grands groupes internationaux. Nous on développe des solutions logicielles. Nos clients, on va à la fois avoir des marques, des institutions, agences, cabinets, institutions. On a donc effectivement tous ces acteurs donc soit des clients qui directement utilisent nos solutions ou soit des clients qui les utilisent pour leurs propres clients.* »[423] Avec pour certains un très large portefeuille clients « *Aujourd'hui on est sur un secteur je vous l'ai dit de 300 clients : ça va de la PME aux très grands comptes donc là on a beaucoup de secteurs. On est présent vraiment partout parce que la veille est nécessaire dans tous les cas que ce soit du public ou du privé […]*»[424].

En effet, lorsqu'on s'intéresse de manière plus générale au marché de la gestion de la réputation en ligne, on constate que celui-ci est fractionné sous une pluralité d'offres allant de la veille en e-réputation ou *social listening*, à la cartographie d'acteurs influents en passant par les stratégies de relations publiques / presse 2.0, d'e-influence ou encore de webmarketing (avec notamment les questions de référencement). Ces offres qui répondent d'une part aux besoins des divers publics, sont aussi et surtout dépendantes du statut des diverses entités du marché (nous proposons une cartographie des acteurs du marché de la réputation qu'on retrouve dans le chapitre 2 de la précédente partie). Tandis que certains proposent seulement une solution complète de mesure de l'e-réputation, d'autres éditeurs disposent eux en sus, d'un pôle *Business Intelligence* (BI) ou *data science* visant à accompagner leurs clients dans le processus général de collecte, analyse et interprétation des résultats. Dans une troisième configuration, l'offre est plutôt spécialisée dans le conseil et l'accompagnement au déploiement des stratégies (grâce à l'utilisation des solutions proposées par les éditeurs).

---

[423] Ibid

[424] Extrait de l'entretien avec l'interviewé 3

Que les clients soient des individus ou organisations, des PME ou grands groupes internationaux, et peu importe le statut des acteurs du marché, le maître mot dans la mesure de l'e-réputation, nous le verrons, est le "sur mesure" : *« En fait nous on est à la fois ad hoc et sur mesure : c'est dire que notre logiciel calcule les mêmes indicateurs pour tout le monde et en fait vu que ce logiciel calcule plus de 100 indicateurs, il ya ensuite la façon dont vous pouvez les agencer, les arranger, les croiser, agréger et tout ça chacun peut faire sa propre recette et en fonction de chaque client on peut faire une recette particulière en fonction des ingrédients que sont l'ensemble des indicateurs que notre logiciel est capable de mesurer.* », Interviewé 1.

Ces indicateurs sont relatifs aux attributs identitaires de l'entité, ses divers publics ou parties prenantes, l'engagement qu'elle génère sur les plateformes social média, son environnement concurrentiel, les acteurs influents qui lui sont associés, etc., tout un ensemble de métadonnées modulable présenté sous forme de cas d'usages comme nous l'explique l'interviewé 5 : « *Il peut y avoir plusieurs cas d'usage, ça peut être : la réputation est un cas d'usage, il peut avoir un cas d'usage crise, il peut avoir un cas d'usage campagne qui active (j'ai une activation social média je veux savoir quelles sont les retombées pour ma marque et comment se positionnent mes concurrents), il peut y avoir un cas d'usage expérience (donc comment les gens réagissent sur un produit, sur un service par exemple quand ils ont utilisé tel produit ou tel service), il peut y avoir un cas d'usage qui peut être plutôt lié à la connaissance de son audience, etc. [...]* ». Le niveau de granularité de l'offre de mesure de la réputation en ligne est donc lié à ces divers cas d'usages et axé sur le client et ses besoins, besoins qui détermineront donc l'offre qui lui sera proposée : « *C'est vrai que le travail d'un chargé de veille ou de consultant en e-réputation varie beaucoup en fonction des besoins et des demandes du client qui sont assez uniques... En réalité, on ne peut pas appliquer un modèle de prestation type, mais il faut toujours qu'on s'adapte à nos clients* »[425]. Adaptation à la fois sur les pratiques, méthodes, mais aussi sur l'utilisation des solutions. En effet, dans une moindre mesure et selon la complexité de l'analyse, l'utilisation de plusieurs solutions peut s'avérer nécessaire comme le montre cet extrait :

[425] Extrait de l'entretien avec l'interviewé 6

« *Donc on a un outil propriétaire qui est vraiment centré modération pour tout ce qui est veille et travail d'e-réputation on se sert plutôt d'outils qui existent sur marché, des outils assez connus dans le milieu type Talkwalker, Visibrain, dans une moindre mesure on se sert de temps en temps de Radarly. Voilà des outils qui se sont largement imposés dans le domaine. Après on va se servir d'outils annexes qui nous servirons sur des petites parties de l'analyse de portée de contenus, de ce genre de choses, de type Buzzsumo qui permet de voir par exemple la diffusion d'URL sur les réseaux sociaux et on se sert aussi d'outils d'alertes d'alerting qui nous permettent d'être avertis en cas d'éléments sensibles, là c'est vraiment pour le côté bad buzz ou signaux faibles [...]. Par exemple, si notre client nous demande la surveillance de presse papier ce qui est possible , il y a un outil qui s'appelle Kantar média qui le fait donc voilà il faudra anticiper en fonction des besoins du client tel outil est plus adapté à sa demande.*»[426].

Cette notion de *sur mesure* est encore plus prégnante et accentuée chez certains acteurs du marché qui proposent ce qu'ils qualifient de *"tickets"* ou *"heures d'astreinte"*, et qui viennent en complément de la mission d'origine comme le souligne l'interviewé 2 :

« *Voilà je peux déjà répondre c'est clairement du sur mesure, c'est -à -dire qu'on a des offres standards [...] pour certains clients, après voilà très honnêtement c'est complètement du sur mesure [...]. On aurait des clients par exemple pour montrer la granularité de la chose, il y a des clients qui… Un client peut choisir soit une veille de chaque jour soit prendre 1 heure de veille soit un montant par exemple vous faites une veille de 30mn par jour soit ... et avec des langues aussi, en quelle langue ça l'intéresse ce que le client est intéressé de savoir, tous les résultats qu'une veille peut produire… Si vous avez un client comme Samsung qui souhaite voir sa visibilité plus celle de ses concurrents en France, mais aussi ailleurs, en France en anglais, mais aussi en chinois par exemple, vous voyez ça aussi c'est encore un niveau de granularité très précis. Il y a des gens qui ont des clients qui vont vouloir avoir des rapports mensuels d'autres semestriels, d'autres pas de rapports. Il y a des personnes qui vont vouloir, et nous on propose des tickets là dessus c'est-à-dire, des heures qu'ils achètent en plus [...]* ».

[426] Ibid.

Cette notion d'accompagnement sur mesure et en temps réel va au-delà de la plateforme mise à disposition des clients, mais concerne aussi et surtout du conseil "humain" et nécessite parfois une disponibilité 7 jours sur 7 comme nous l'indique l'interviewé 6 « *cela demande une surveillance très régulière tout au long de la journée donc forcément des chargés de veille disponibles. On se sert toujours d'outils et ça demande en fait surtout... Nous on travaille en fait 7 jours sur 7 et on est ouvert également les weekends, donc on a tous des weekends différents pour permettre de couvrir l'intégralité de la semaine. Je pense que c'est aussi une question de disponibilité et d'amplitude horaire effectivement malheureusement. Internet ne ferme pas le vendredi soir, mais parfois on aimerait le couper 24h, mais donc voilà ça demande une grande disponibilité* ». Il est donc impossible de définir une offre une, unique et générale dans la mesure de la réputation en ligne, celle-ci (l'offre) variant selon les besoins et attentes du client.

2.4. La mesure de l'e-réputation, une méthode en construction

Même si on suppose, au vu du précédent point que la mesure de la réputation en ligne se veut être une offre sur mesure, il n'en demeure pas moins que dans une approche plus généraliste, tous s'accordent sur une méthode peu ou prou préconçue qui se veut modulable et adaptable selon les objets et besoins de l'analyse, et qu'on peut décliner en trois temps : « *L'idée c'est qu'on fait du 3 en 1 dans la plateforme ... On collecte l'information, on l'analyse et on la diffuse tout ça dans l'outil* ».[427] La première étape, la collecte, vise à *crawler* ou interroger de manière extensive le web à l'image d'un moteur de recherche généraliste, afin d'obtenir le *dataset* ou l'ensemble de données qui sera par la suite traité. Cette phase est essentielle, car du paramétrage de la solution en amont de la collecte dépendra la qualité du travail d'audit. En effet, un des concepts clés à cette étape est le choix des requêtes et le paramétrage qui nécessitent un travail de minutie autant dans l'utilisation des opérateurs booléens de type "*and, or, ni*", des opérateurs de proximité, des troncatures, des sous requêtes ou alias (visant par exemple à regrouper un champ sémantique), de l'ajout ou le retrait de termes afin d'affiner le filtrage, du choix des langues, etc.

---

[427] Extrait de l'entretien avec l'interviewé 3, un des acteurs ayant développé une solution logicielle

Si le choix des requêtes nécessite à minima des connaissances en termes de veille, intelligence sociale ou *social intelligence*, le paramétrage de la solution requiert plutôt du "bon sens", les solutions proposées sur le marché disposant d'une myriade de fonctionnalités à utiliser selon les besoins. L'objectif de ce paramétrage est donc d'éviter "les bruits" et de collecter de la donnée qualitative, précise et pertinente pour l'objet de mesure. Un cas d'école assez récurrent qui ressort de nos entretiens est relatif à l'entité Orange qui étant à la fois un fruit, une couleur et une ville, nécessite un travail de paramétrage minutieux afin d'obtenir de la donnée qualifiée.

Sur les sources utilisées par le *crawler* dans la collecte des données, à l'instar d'un interviewé pour qui les moteurs de recherches généralistes et plus particulièrement Google est la principale source « *Déjà tu peux dire très clairement que google reste quand même numéro 1 et de très très loin dans ces analyses là. En fait, on passe simplement par les moteurs de recherche qui sont une source inépuisable de données. [...]. Après il faut partir du principe en France en tout cas parce que là on est capable de le faire dans toutes les langues dans le monde entier sur n'importe quelle thématique, mais en France en tout cas, 97% des infos sont sur Google même si des fois on va aller interroger Bing, Yahoo, on a essayé Qwant même si franchement ce n'est pas très convaincant en tout cas pour notre métier.*»[428], tous s'accordent sur un ensemble de bouquets de sources indispensables à surveiller : « *on a des millions de sources qui sont référencées en multi langues, donc ça va être des sites de presse internationales/ nationales, des blogs, des forums, des sites d'entreprise, des réseaux sociaux, tout ce qui est accessible on peut le récupérer et en plus, on va pouvoir également capter de l'information premium qui provient de BDD externes* »[429] auxquelles s'ajoutent éventuellement celles issues de partenariats avec d'autres acteurs dont les sites d'avis par exemple (Avis vérifiés, Tripadvisor, etc.), les grands médias traditionnels (la télévision, la radio), les sites de publication scientifique, d'appel d'offres, etc. Le référencement des sources nécessite également quelques fois une part humaine via notamment le "*clipping*" qui vise à " aspirer en un clic le contenu d'un site" et à l'intégrer à la solution utilisée.

---

[428] Extrait de l'entretien avec l'interviewé 7

[429] Extrait de l'entretien avec l'interviewé 3

Chaque acteur dispose donc, en plus de ce qu'on qualifiera de sources primaires utilisées par la quasi-totalité des acteurs du marché (via les solutions logicielles), de partenariats variés selon les besoins et la présence de leurs clients, et certaines plateformes disposent même de méthodes permettant de mesurer la pertinence d'une source pour un acteur ou une recherche donnée :

> « *En fait, on a un outil ici dans la partie source qui s'appelle "supervision des bouquets" pour pouvoir dire justement voilà l'ensemble des bouquets qui m'intéressent [...] toutes mes sources, "qui produit du contenu, à quelle fréquence et éventuellement celles qui produisent peu aussi" parce que c'est une source qui n'est pas active en tant que telle et ça permet de dire aussi voilà celle-ci aussi par rapport à ce que j'ai, je l'ai évaluée et non elle n'est vraiment pas terrible et dans ce cas-là je vais plutôt l'ignorer dans ma requête.* »[430]

À l'issue de cette phase de collecte, il s'agit dans un second temps d'analyser le *dataset* collecté. Pour ce faire, et à partir d'algorithmes de *machine learning* (ML), l'objectif est de sélectionner et classer les données pertinentes pour l'étude.

À ce stade, il est important de noter que « *le moteur apprend et vous propose de nouveaux contenus pertinents au-delà de la requête initiale qui a été créée avec des opérateurs booléens* » comme le note l'interviewé 3, et l'analyse permet de faire ressortir « *les tendances d'un point de vue quantitatif que qualitatif c'est à dire qu'à l'issue, des métadonnées sont calculées par les algorithmes qui permettent [...] de faire apparaître des saillances dans la donnée qui agissent comme des boussoles pour l'analyste qui vont lui permettre d'aller creuser aux endroits spécifiques où les saillances apparaissent. Ensuite, cette analyse, cette exploration humaine permet de prendre contact, de détecter la matière qualitative et à partir de ça, on essaie d'extraire scientifiquement, de sélectionner les spécificités particularités à la fois qualitatives et quantitatives de ce corpus pour donner une image qui sera la plus à la fois exhaustive et sélective de la réputation de l'entreprise* »[431].

---

[430] Extrait de l'entretien avec l'interviewé 3

[431] Extrait de l'entretien avec l'interviewé 1

Une fois les données traitées, une part de travail humain est essentielle afin de piloter le dispositif de manière générale et d'identifier les indicateurs clés aussi bien qualitatifs et relatifs à de l'analyse sémantique, et dont les visées peuvent être multiples et variées selon les objectifs de mesure (*faire ressortir sous forme de nuages de mots et/ ou de verbatims les entités nommées, les groupes nominaux, les concepts qui sont liés à la recherche, etc.*) que quantitatifs axés sur des données démographiques (portant autant sur les clients, consommateurs, influenceurs potentiels que sur les concurrents), un score de réputation, du *sentiment analysis*, etc. Le type de données issues de l'analyse est très varié avec une myriade d'indicateurs à paramétrer et choisir dans la solution de *social listening* selon les objectifs et besoins initiaux (ce qui fait encore appel à des compétences humaines dans le choix des indicateurs pertinents).

Selon le statut de l'acteur du marché de la mesure de la réputation, éditeur de solution avec un pôle Business Intelligence (*BI*), agence généraliste (de marketing ou communication digitale, spécialisée dans l'influence ou RP 2.0) ou *pure player* dans la mesure de la réputation, l'objectif est soit de produire des *insights* « *nous on est vraiment sur la partie insights en amont et après s'en découle la partie BI qui est plus exploitation de la data pour aller transcrire justement… on peut travailler avec d'autres acteurs* »[432] qui seront analysés par les partenaires ou directement envoyés aux analystes des commanditaires, ou, de proposer un accompagnement sur le déploiement d'une stratégie comme nous l'indiquent cet extrait « *en fonction de la demande du client on fait des constats soit on l'accompagne pour faire le pas d'après qui consiste à transformer les constats en conseils.* »[433]. Certains acteurs proposent en sus une cartographie d'acteurs influents « *chez nous, on va soit, selon la demande du client, le spatialiser avec des logiciels tels des algorithmes de spatialisation, ou alors le mettre dans des agrégateurs pour pouvoir identifier des acteurs qui les mentionnent, qui sont proches ou moins proches des marques* »[434].

À l'issue de ce travail d'analyse, les résultats sont envoyés aux commanditaires via plusieurs canaux de diffusion conjointement prédéfinis (nous y reviendrons plus en détail dans le prochain point).

---

[432] Extrait de l'entretien avec l'interviewé 3

[433] Extrait de l'entretien avec l'interviewé 1

[434] Extrait de l'entretien avec l'interviewé 7

Si tous s'accordent sur une méthodologie générale plus ou moins identique et modulable selon les besoins, un seul des acteurs interviewés (un éditeur de solution logicielle) présente et propose un "schéma" de mesure de la réputation, même si tous en disposent sans qu'ils ne soient expressément formalisés, théorisés, eu égard aux méthodologies présentées lors des entretiens. C'est notamment le cas de l'interviewé 1 :

« Par exemple, on a construit un modèle aujourd'hui de réputation chez nous que nous recommandons à certains de nos clients et qui s'appelle ADPR (awareness, durability, proximité, relevance) donc ce sont 4 dimensions qui ont pour objectif d'appréhender les grandes... de mesure de l'image des marques sur le web social. Donc awareness c'est la puissance le fait que la marque soit connue ou pas, discutée ou pas, la durability c'est le fait que la marque soit aimée ou pas la proximité c'est le fait que la marque soit engageante génère de l'engagement ou pas et la relevance c'est la partie la plus innovante, c'est le fait que le bruit social de la marque soit aligné avec les valeurs de la marque donc avec ses valeurs de marque. Donc ce framework après pour certains de nos clients quand ils nous demandent un bilan d'e-réputation on est capable de leur calculer leur score ADPR marché par marché, marque par marque, en comparatif d'1 mois/ semestre/ an sur le suivant. »

Le modèle ADPR proposé par cet acteur s'appuie donc sur l'ensemble des métadonnées disponibles sur l'organisation dont ils mesurent l'e-réputation : « l'*awareness est calculé en fonction des mentions sur leur visibilité potentielle le reach, la durabilité est calculée en fonction des mentions positives, la proximité va être calculée en fonction des actions d'engagement, la relevance en fonction de l'association de la marque à certaines thématiques.* »

Comme on peut le constater dans une approche généraliste, la collecte, l'analyse et l'interprétation des résultats sont les trois étapes clés dans la mesure de la réputation en ligne. Selon la solution choisie, mais aussi et surtout les besoins de l'étude, certaines étapes seront plus ou moins complexes que d'autres : c'est notamment le cas de la phase d'analyse dans laquelle de nouvelles étapes intermédiaires peuvent être ajoutées ou supprimées selon les besoins et la finalité de l'étude.

2.5. Différents types de livrables

L'ultime phase à l'issue des diverses étapes sus-citées dans la mesure de l'e-réputation d'une entité est le livrable qui sera fourni au client, livrable qui comme nous l'avons déjà mentionné dépend encore une fois des objectifs préalablement fixés, et qu'on peut regrouper sous trois (3) principales typologies : les livrables de type " alerte ", les lettres d'information, dans le cas d'un audit complet le rapport.

De manière générale, les résultats issus de l'analyse sont présentés sous forme visuelle dans des tableaux de bord de la réputation. Support le plus plébiscité et intégré aux solutions de veille ou d'écoute sociale (*social listening*), il est composé d'un ensemble de catégories (*widgets*) qui reprend les informations ou données collectées et analysées sous forme d'infographies. Les widgets qui peuvent être paramétrés selon les besoins des commanditaires sont essentiels, car c'est à partir d'eux que sont créés les indicateurs (données démographiques, polarité/ tonalité, score de réputation, etc.) qui serviront par la suite à visualiser les données issues de l'analyse comme nous l'indique l'interviewé 3:

> « *C'est vraiment le client qui va le créer lui, par rapport à ses besoins. Nous on met en place ces outils avec des widgets, chaque client crée ces différents widgets ce qui va l'intéresser [...]. En fait, chacun choisit ses indicateurs, les crée par rapport à ses objectifs et ça permet du coup de partager le contenu derrière et pouvoir dire ce qui va m'intéresser après.* »

À partir de ces tableaux de bord et selon l'objectif de l'étude, les agences proposent un ou plusieurs de ces livrables.

La première typologie l'alerte, consiste, à partir de mots clés préalablement définis avec l'entité, à lui notifier par mail, et ce en temps réel et de manière automatique des alertes sur toute information ou donnée relative à l'organisation. Ce type de livrables est plus couramment utilisé pour pallier les crises ou *bad buzz* éventuels via notamment la réception en temps réel des données sur la totalité des sources surveillées dans la collecte des données dans la mesure de la e -réputation.

Le second type de livrable proposé par les acteurs de la mesure de l'e-réputation est la newsletter. Sur le même principe que l'alerte (*alerting*), la newsletter plus complète est elle plutôt relative à la veille stratégique et concurrentielle. Comme nous l'explique l'interviewé 3, via l'outil d'écoute du web social (*social listening*) :

> « *L'idée c'est que je lui ai défini les grandes catégories et les requêtes donc là je suis vraiment en fonctionnement automatisé je ne vais jamais moi dans le back office pour voir les articles qui sont tombés en fait. Je lui dis voilà j'ai crée plusieurs grandes rubriques "ie actualités" "ie réseaux sociaux", etc., et je lui ai dit voilà tu m'envoies tous les contenus qui correspondent à ma requête tous les 2 jours voilà à 12h... le contenu qui correspond à ma requête sous forme de newsletter et moi ça me permet ensuite de me dire bah tiens ça, ça m'intéresse, ça, ça ne m'intéresse pas et en l'occurrence c'est de la veille commerciale...* »

Il est important de noter que les deux premières typologies de livrables viennent le plus souvent en complément et à la suite des livrables rapport et visent à assurer un suivi sur le long terme de l'entité, et la fréquence d'envoi de la newsletter est décidée de manière conjointe avec l'entité et selon la quantité de données collectées.

Enfin, et comme dernier type de livrable proposé, le rapport classique. S'il est beaucoup moins évoqué par la plupart des interviewés , certains l'utilisent encore comme support final, et il se présente de manière générale sous forme numérique via par exemple des PDF ou PowerPoint. À l'exemple du *dashboard* et selon les objectifs de mesure, il reprend, et ce, de manière complète et exhaustive, les mêmes *insights* ou indicateurs. Tous ces livrables et il est important de le souligner peuvent s'utiliser de manière concomitante et les deux premiers viennent en général en support des livrables plus conséquents avec plusieurs autres formats disponibles, et peuvent varier (dans certains cas) selon le degré de proximité avec le client comme nous l'indique l'interviewé 7 « Nous *on rend un rapport à nos clients et ensuite[...], on va définir en fonction du besoin client des livrables clients donc ça peut être : des fiches pratiques, ça peut être des workshops, ça peut être des "mails automatisés", ça peut être un rapport d'audit assez costaud. Ça dépend vraiment du client, de la proximité aussi qu'on a avec eux. Un client qu'on a depuis 4 ans, il nous connaît, lui il ne veut pas forcément avoir un rapport parce qu'il sait qu'il nous fait confiance, lui il veut surtout les préconisations et l'opérationnel* ».

Cette propension de la part des clients à ne rechercher que des données opérationnelles et des préconisations concises et synthétiques revient de manière assez récurrente chez nos interviewés, dû en partie à la complexité des tableaux de bord, mais aussi et surtout à la situation d'urgence dans laquelle ils se trouvent comme nous l'explique l'interviewé 6 :

> « *Très souvent, on travaille avec des personnes qui sont, il ne faut pas se le cacher, très stressées avec des dirigeants qui ont très peu de temps pour nous lire, donc l'idée c'est vraiment transmettre les éléments qui symbolisent quelque chose. On n'est pas là pour écrire de la grande littérature, on est vraiment là pour transmettre le principal et le plus synthétiquement possible.*»

2.6. La mesure de l'e-réputation, un avenir mitigé?

Ce dernier point évoqué par certains interviewés fait notamment référence aux diverses restrictions en vigueur sur certaines plateformes numériques limitant de fait l'accès à certaines données, et aux questions d'anonymisation, de droit des données et plus généralement de régulation des médias sociaux. Si certains voient aux études d'e-réputation un avenir incertain, d'autres y voient *a contrario* un avenir radieux. Il est intéressant de noter dans une première approche les prévisions plutôt nuancées que font certains des interviewés. Une des zones d'ombres qui questionne les interviewés sur l'avenir des questions liées à l'e-réputation et sur laquelle il est intéressant de se pencher en amont est notamment relative à la difficulté, voire à l'impossibilité parfois d'accéder aux contenus de certaines plateformes. C'est notamment le cas de Facebook par exemple qui a restreint l'accès à ses données, et s'il est encore possible de collecter les publications, les solutions d'écoute du web social disponibles sur le marché ne permettent cependant pas de surveiller les commentaires postés, ce qui est assez problématique dans l'immédiat comme le souligne l'interviewé 6 : « *C'est vrai que certes l'e-réputation se joue dans les contenus qui sont visibles, mais aussi dans la récurrence d'éléments de langage qui peuvent donc revenir parmi les commentaires qui ne seront pas d'une visibilité extraordinaire, mais leur récurrence en fera tout de même une tendance et ça on a du mal à trouver un moyen de pouvoir l'analyser* ».

Ces restrictions font suite au scandale *Cambridge analytica* (2018) sur la fuite des données personnelles de 87 millions d'utilisateurs Facebook, ce qui a poussé les réseaux sociaux de manière plus générale à mettre en place de nouvelles restrictions dans la protection des données des utilisateurs, d'où les craintes de certaines personnes interviewées sur le futur des études de réputation : « *Avec la protection des données et la nouvelle régulation sur les réseaux sociaux, encore plus quand on vient de voir ce qui s'est passé avec la décapitation de Samuel Paty, est-ce que les réseaux sociaux ne vont pas petit à petit fermer leurs données parce que mine de rien qu'on n'aime ou qu'on n'aime pas on scroll, on scanne toute une vie qu'on croit privée?* »[435]. Ces diverses restrictions renvoient donc de facto aux questions d'anonymisation sur les plateformes sociales qui, depuis quelques années, font l'objet d'une remise en cause.

En effet, afin de favoriser l'analyse et la publication des données issues d'un graphe social du profil d'un internaute, la première étape consiste à anonymiser le graphe, afin de supprimer tous les attributs identitaires relatifs à son profil. Même si certains dispositifs (*VPN, TOR* par exemple) permettent l'anonymat sur certaines plateformes numériques et que plusieurs chercheurs ont développé des méthodes et algorithmes d'anonymisation des données (anonymisation de degrés, Liu et Terzi, 2008[436] ; anonymisation de voisinage, Zhou et Pei, 2008[437] ; l'algorithme de K-automorphisme, Zou et al., 2009[438] ; la généralisation du graphe social, Hay, et al. et Kampan et Truta, 2008[439] ; anonymisation par perturbation aléatoire, Bronchi et al., 2011[440]), il est important de noter que l'anonymat total et/ ou définitif n'existe pas réellement dans les espaces numériques.

---

435 Extrait de l'entretien avec l'interviewé 2

436 Kun Liu and Evimaria Terzi. Towards identity anonymization on graphs. In SIGMOD Conference, pages 93–106, 2008

437 Bin Zhou and Jian Pei. Preserving privacy in social networks against neighborhood attacks. In *ICDE*, pages 506–515, 2008

438 Lei Zou, Lei Chen, and M. Tamer O¨zsu. K-automorphism : A general framework for privacy preserving network publication. *PVLDB*, 2(1) :946– 957, 2009.

439 Michael Hay, Gerome Miklau, David Jensen, Donald F. Towsley, and Phi- lipp Weis. Resisting structural re-identification in anonymized social net- works. *PVLDB*, 1(1) :102–114, 2008. via l'algorithme GraphGen / Alina Campan and Traian Marius Truta. Data and structural k- anonymity in social networks. In PinKDD, pages 33–54, 2008. via l'algorithme SaNGreeA (Social Network Greedy Anonymization)

440 Francesco Bonchi, Aristides Gionis, and Tamir Tassa. Identity obfuscation in graphs through the information theoretic lens. In *ICDE*, pages 924–935, 2011.

On parlera plutôt “d’anonymisation naïve” ou de “pseudonymat” qui vise à dissimuler son identité via un pseudonyme, procédure qui ne garantit pas la préservation de la vie privée ou protection des données et qui peut permettre l’identification des internautes sur les plateformes sociales numériques sur lesquelles ils s’inscrivent. Par ailleurs, depuis 2019, le gouvernement français a entrepris, suite au déferlement de haine et de harcèlement en ligne, des réformes avec un nouveau plan d’action visant à interdire l'anonymat pour certains des usages sur les médias sociaux (et plus particulièrement les réseaux sociaux tels Facebook, Twitter, etc., les forums ...). Face à ce constat, certains interviewés s’inquiètent de la réaction des utilisateurs sur ces plateformes comme le résume cet extrait de l’interviewé 2 :

> « *Maintenant, quand on entend l'anonymat, l'anonymisation pardon remise en question sur internet, donc les gens, est-ce qu'ils prendront plus la parole en leur nom? Vous savez, chaque jour pour nos clients, on a parfois des menaces directes de mort. Quand il s'agit parfois de banques, voilà est -ce que la personne écrira la même chose? Il y a même des gens qui travaillent dans certaines institutions de certains de nos clients qui vont se permettre de parler, qu'on a identifié parce qu'ils ont fait des erreurs, mais qui se permettent de parler et s'ils avaient leur vrai nom ils ne diraient pas la même chose. Donc l'anonymisation c'est un débat qui rend ça un peu plus difficile*».

Cette première assertion quelque peu critique sur l’avenir des études de réputation, se voit contrebalancée par un second courant pour qui, ces questions de réputation existant depuis la nuit des temps et avec les clients / consommateurs qui sont passés d’un statut plus ou moins passif à celui de « *consom’acteurs* », de parties prenantes intégrales dans les stratégies des organisations, il faudrait plutôt s’attendre à un essor de pratiques et méthodes (nouvelles peut être) afin de mieux cerner le concept, ses mutations et de l’appréhender. C’est notamment le cas pour l’interviewé 1 :

> *« Des études de réputation, d'image, je leur vois un avenir extrêmement florissant. Après la façon de faire ça a pas mal changé, mais des études de réputation je leur vois un avenir florissant pourquoi parce que simplement nous sommes dans une économie de la conscience de l'attention qui est la valeur première qui est issue des groupes primitifs et à ce titre là, les questions réputationnelles, d'image sont fondamentales dans cette économie de la conscience, de l'attention et dans ce capitalisme cognitif donc je pense que les questions d'image n'ont jamais été aussi*

fondamentales et que les questions de réputation n'ont jamais été aussi importantes[...]. Les études d'e-réputation vont, si derrière on les appelle si on considère que ce sont les outils, méthodologies qui permettent de modéliser, suivre, analyser et traduire la perception de marque de produits/ services, ça va aller à fond avec le temps et devenir de plus en plus critique».

En effet, s'il est une chose sur laquelle tous s'accordent relativement, c'est que peu importe l'entité (personne physique ou morale) ou son statut, tout le monde aujourd'hui souhaite savoir ce qui se dit de lui. Qu'il s'agisse de se « googliser » sur les moteurs de recherche génériques ou par l'utilisation de solutions plus sophistiquées de veille ou d'écoute du web social (*social listening*) à visée de gestion de sa réputation en ligne, ces pratiques font désormais partie intégrante de notre quotidien et connaîtront de fait un certain essor tant sur les pratiques, les méthodes que les outils : « *J'en vois un développement en tout cas certain et le développement il est certain il n'y a pas de doute tout le monde va vouloir savoir ce qu'on dit sur soi aujourd'hui et je trouve même que certaines sociétés sont un peu en retard parce que ça donne, en plus d'insights sur soi même, ça en donne sur ses concurrents, sur les consommateurs [...]. C'est quelque chose qui va se développer bien sûr à mesure que le web se développe. Encore plus s'il y a des notions de ... peut être un jour de réalité virtuelle [...] il peut y avoir pleins de choses qui peuvent arriver même au niveau de la vidéo parfois quand on utilisera une marque pour faire un jour un deep fake [...]. Il y a un développement certain, toute société a intérêt à faire de l'e-réputation [...] là on a vraiment de la donnée clé et ça nous donne des insights à la fois sur nous et sur des résultats concrets. Donc c'est hyper important et ça va se développer fois 1000. En plus, il ya tellement de micro-influenceurs, d'influenceurs qui se développent aussi, donc ça va se développer bien sûr* »[441].

Il est par ailleurs intéressant de noter et ce, peu importe le positionnement qu'on a sur la question, que la mesure de la réputation suppose l'utilisation uniquement de données publiques, c'est-à-dire de contenus accessibles par tous sur le web social. Cette question de données publiques est essentielle, car c'est de leur agrégation que naît l'e-réputation d'une entité.

[441] Extrait de l'entretien avec l'interviewé 2

Dans l'optique où certains contenus seraient importants pour une entité et qu'ils seraient en outre difficiles, voire impossibles à collecter (sur des plateformes telles Facebook par exemple), l'outil ou la solution ne fait pas tout comme nous le soulignent plusieurs des interviewés. Les solutions viennent en support afin de faciliter l'accès aux contenus. Cependant, et pour plus d'exhaustivité, la plupart ont recours à de la surveillance manuelle afin de compléter ou étoffer celle automatique via les solutions logicielles :

> « *Comme je vous l'ai dit, tant qu'on peut encore avoir l'accès à des pages publiques et c'est dans l'intérêt de tout le monde, le format d'e-réputation, les techniques vont peut être juste changer on va juste embaucher quelqu'un qui va juste aller voir les pages ou les sites internet. Par exemple, typiquement il y a des choses qu'on a du mal à scanner ce sont les avis par exemple Google my business et les forums de consommateurs dans ce cas-là nous on se met des listes de cibles à surveiller [...] au fait c'est de la veille complètement manuelle qui, voilà, on a dû s'adapter parce que l'outil ne le fait pas et on doit aller voir. Linkedin typiquement on n'a pas parlé de Linkedin, mais Linkedin il n'y a rien qui est monitoré par aucune, j'entends des plateformes. Typiquement, on doit y aller à la main pour voir si la campagne de communication elle a marché, non? Il faut aller voir, donc on saura toujours s'adapter [...], mais on va devoir mettre un peu plus la main à la patte* »[442].

L'aptitude des solutions du marché à développer de nouveaux partenariats avec d'autres acteurs (sites d'avis, nouveaux réseaux : Tik tok par exemple) et à avoir une couverture la plus exhaustive possible, mais aussi et surtout la capacité d'adaptation des consultants en e-réputation à fouiller manuellement (en support à la veille automatique) les pages inaccessibles de prime abord sont donc un enjeu futur majeur dans la mesure de l'e-réputation et les acteurs du marché rivalisent donc de méthodes ou approches :

> « *On est obligé de s'adapter effectivement aux nouvelles tendances et aux nouveaux outils. Nous on commence petit à petit à faire de la veille sur Tik Tok aussi par exemple, un réseau social qui n'est pas encore tout à fait maîtrisé en termes de plateforme social média, mais on commence à trouver quelques astuces pour regarder les contenus qui nous intéressent, qui sont diffusés donc voilà, il faut sans cesse nous renouveler et nous intéresser aux nouvelles habitudes d'utilisation de consommation des plus jeunes et donc effectivement évoluer en même temps que l'entreprise* »[443].

---

[442] Ibid

[443] Extrait de l'entretien avec l'interviewé 6

Si l'adaptation et le renouveau sont essentiels pour cerner les questions de réputation en ligne, il est une chose qui reste et restera inchangée et qui profitera toujours à l'e-réputation c'est le concept de crise. Pour paraphraser l'interviewé 2 « *la crise fait vivre la réputation* ». En effet, comme il le souligne avec ironie:

> « *J'ai envie de vous dire qu' avant le covid on était entre guillemets égal aux autres, mais c'est vrai que le covid a complètement bouleversé le pôle. De toute façon, toute crise fait vivre le pôle e-réputation. Pour vous dire, ce n'est pas à cacher ou quoi que ce soit, parmi l'ensemble ou l'intégralité des pôles, tous ont plus ou moins été mis en chômage partiel donc certains membres ou un pôle complet ça dépend quelle équipe ou quoi et on est les seuls qui au contraire et ça c'est pas vraiment à dire juste pour vous le tourner autrement parce que ça, ce n'est pas à moi de le dire, mais on les seuls qui avons eu des primes parce qu'on travaillait 2 fois 3 fois 4 fois plus que d'habitude. On travaillait jusqu'à très tard le soir parce que tout le monde voulait avoir des retours soit sur le télétravail, l'impact du covid, le progrès de la pandémie, vraiment... je ne pourrai pas vous dire le pourcentage que ça représente... mais juste vous dire que les crises et du coup le covid a eu un impact fondamental ça veut dire que même bah ils m'ont recruté c'est vous dire que ça pèse dans la balance et qu'on continue à recruter dans mon pôle et que chez les autres bah ça stagne quoi. Donc, je remercie la covid.* »

À partir de ces éléments et comme le suggère l'analyse inductive, nous proposons une structure schématique reprenant les diverses catégories d'analyse :

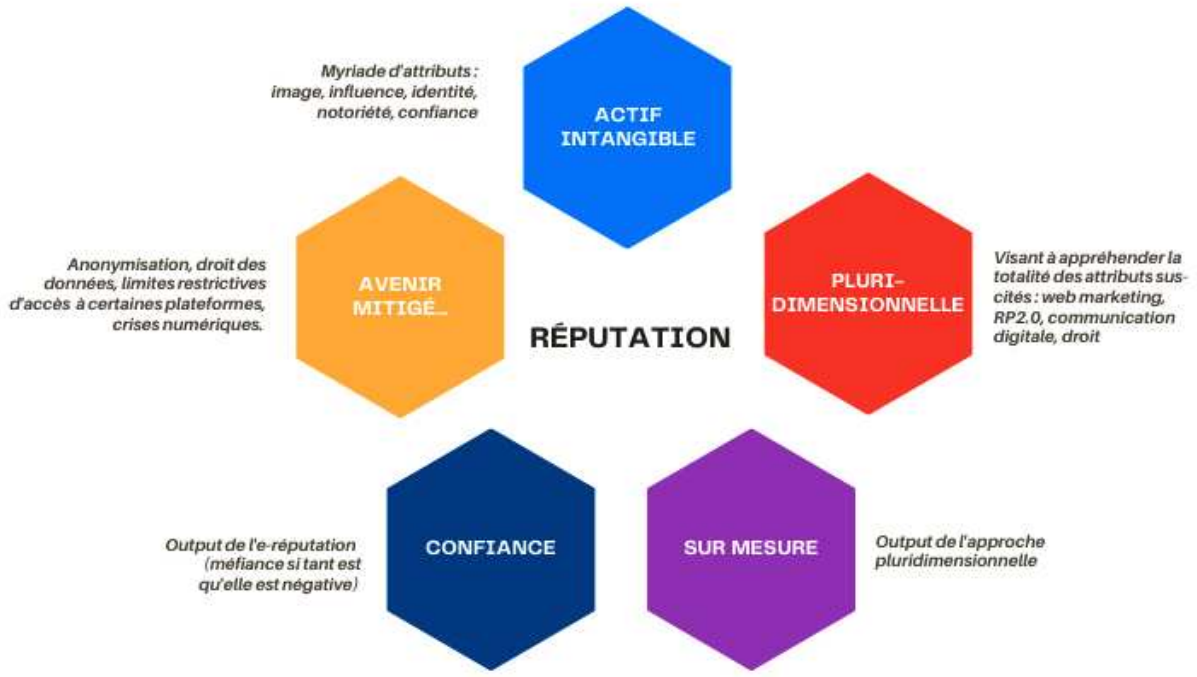

Figure n°33: Catégorisation des unités de sens, réalisée à partir des entrevues

Au centre de la structure du modèle, on retrouve l'objet réputation pour rappeler l'assertion selon laquelle il n'existe qu'une seule et même réputation pour toute entité (ici perçue au sens d'individualité, organisation, produit ou service). Cette réputation utilise aujourd'hui dans sa gestion et plus particulièrement sa mesure une approche différente eu égard aux nouvelles méthodes et pratiques mises en avant par le web social et ses applications. L'un des concepts clés qui ressort également dans notre modélisation est la notion d'actif intangible qui n'est pas sans rappeler l'approche pluridisciplinaire nécessaire, voire indispensable, dans la captation des divers signaux (au sens de données numériques) disséminés sur la toile et servant d'éléments de mesure. Ces signaux sont axés sur la totalité des attributs d'une entité, mais aussi et surtout de son environnement (incluant ses diverses parties prenantes, les concurrents, etc.).

Ces métadonnées renvoient *de facto* à l'ensemble des attributs (environnement concurrentiel, image ou identité, influence, tonalité ou polarité, données de référencement) sur lesquels porte la mesure (tout le caractère intangible). Afin de les rendre intelligibles, le recours à une approche sur mesure s'avère être la méthode la plus adéquate (du moins celle prônée par les acteurs du marché). S'il est une question qui porte quelque peu à confusion c'est bien celle relative à l'avenir de ces questions de réputation qui somme toute continueront à exister avec le développement de nouvelles formes de pratiques et mesures même si certaines plateformes restreignent l'accès aux données, le fondement même de la mesure de la réputation étant l'utilisation de la donnée publique. Et enfin, le concept de confiance abordé dans la modélisation théorique de l'objet réputation, n'est autre que la finalité (*l'output),* la résultante sur le long terme d'une bonne réputation.

De la restitution de ces entrevues et leur analyse, il convient de retenir plusieurs éléments :

- Les questions de réputation existent depuis toujours, et les questions de réputation et d'e-réputation font référence à un seul et même objet, c'est-à-dire que la réputation reste la réputation, qu'elle se mesure en ligne ou hors ligne (*in real life ou off line*) avec cependant une mutation des pratiques et des méthodes qui prennent aujourd'hui en compte une certaine forme de technicité, et l'apparition de problématiques nouvelles relatives aux caractéristiques intrinsèques du web social;

- La mesure de la réputation n'est pas une science exacte et aucun des acteurs ne s'appuie sur une définition claire et concise du concept, et au-delà de cet état de fait, certains prestataires proposent une approche commerciale dénuée de toute scientificité (en prétendant vendre une offre de mesure qui en réalité n'en est pas une lorsqu'on s'y intéresse de plus près);

- Ces questions d'image qui auparavant et via notamment les baromètres prenaient une forme et utilisaient une approche plutôt qualitative, se déclinent aujourd'hui sous une approche quantitative (eu égard notamment à la quantité exponentielle de données disponibles sur le web social);

- L’effet de scientificité du nombre dans le traitement, mais aussi et surtout la mise en forme visuelle (dans les tableaux de bord) des données quantitatives collectées, satisfait aujourd’hui les clients.

L’objectif principal de ces entrevues est, rappelons-le, d’identifier des phénomènes récurrents perçus par les interviewés (eu égard à leurs expériences) et relatifs à notre objet d’étude, afin de déceler un éventuel rapport entre les savoirs académiques et les pratiques professionnelles dans la mesure de la réputation en ligne. À partir de ces premiers constats et si on postule que les objets réputation et e-réputation ne font qu’un et que seules les pratiques et méthodes ont mué, que la mesure de la réputation suppose parfois une approche et rhétorique plus commerciale que scientifique, et qu’il n’existe pas de rationalisation claire sur l’essence même (il est question ici de la définition) du concept et ses diverses déclinaisons, il serait interessant de questionner dès lors, le rapport qui existe entre les ancrages théoriques observés et les diverses pratiques et méthodes professionnelles?

3. Interprétation des résultats et discussion

À partir des différents éléments d'analyse et constats, il convient de revenir sur les questions et problématique autour desquelles se sont articulés nos travaux de recherche.
Notre problématique principale se présente comme suit : les professionnels de la e-réputation ont-ils recours dans leurs pratiques professionnelles de mesure, à des savoirs académiques?
Afin d'apporter des éléments de réponse visant à étayer notre problématique centrale, nous nous appuierons sur un certain nombre d'éléments qui se rapportent à nos questions de recherche.

3.1. Les modèles théoriques sur lesquels sont construites les méthodes de mesure : des questions de recherche à la problématique…

Questionner l'apport des savoirs académiques dans les pratiques professionnelles dans le processus de mesure de l'e-réputation suppose, nous l'avons évoqué, un processus en trois phases.
D'abord, il s'agit d'identifier les composantes et attributs clés autour desquels sont construits les modèles théoriques. Ensuite, il convient d'analyser les méthodes de mesure proposées par les professionnels qui évoluent au sein du marché de la réputation.
Et enfin, à partir des deux matrices issues des phases d'analyse, déterminer un éventuel rapport entre les deux approches.

À ce stade de nos travaux, trois éléments de réponse peuvent être apportés à notre problématique centrale.

- ***Même s'il n'y a pas de rationalisation claire sur les pratiques, certaines approches de mesure de la réputation proposées sur le marché se sont inspirées dans une certaine mesure de trois (3) des catégories de modèles préexistants***:

En effet, si on reprend la cartographie des acteurs qui composent le marché de l'e-réputation, on constate qu'il existe autant d'offres de services que de besoins avec six (6) principales catégories de prestataires. Ce manque de rationalisation claire dont l'origine vient de la difficulté même à proposer un cadre théorique empirique sur l'objet réputation, amène les agences à proposer une offre *sur mesure*, offre qui pour certaines d'entre elles se sont inspirées d'approches académiques : modèles sommaires, ad hoc et génériques.

Si on se réfère à notre classification des modèles théoriques de mesure de la réputation, quatre grandes familles d'approches se distinguent : les modèles sophistiqués, ceux Ad hoc, les approches génériques et enfin les modèles sommaires. Les approches génériques sont construites sur le principe modulaire, c'est-à-dire avec une finalité d'utilisation pluridimensionnelle. Sur les approches ad hoc, elles visent entre autres à évaluer un "aspect" spécifique de l'objet réputation, tandis que les modèles sommaires eux, s'appuient sur la collecte de données simples sans l'utilisation spécifique pour la plupart d'approches algorithmiques de mesure.

À partir de là, notre analyse est la suivante : ***certaines méthodes de mesures utilisées par les éditeurs de solutions logicielles et les pure players s'appuient sur une approche générique; certains modèles Ad hoc ont inspiré les méthodes proposées par les agences Ad hoc, celles généralistes et certaines applications gratuites; et enfin, les sites d'avis utilisent dans les faits, des éléments des modèles sommaires.***

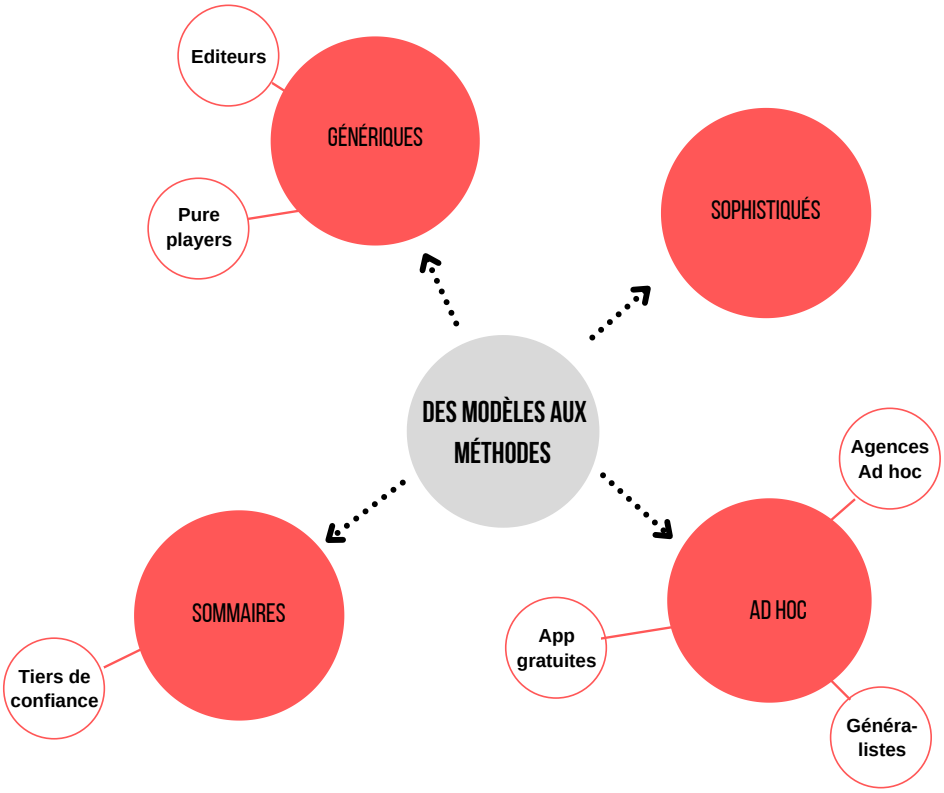

Figure n°34 : Les méthodes de mesure et les approches théoriques qui les ont inspirées

À partir de notre analyse des méthodes de mesure proposées par les acteurs, il ressort que les éditeurs de solutions logicielles et les *pure players* notamment se sont inspirés des modèles génériques (ou modulaires). En effet, le principe des approches modulaires comme l'indique leur nom est simple : proposer sous forme modulable un processus de mesure de la réputation adaptable à divers contextes, environnements ou applications.

Les éditeurs de solutions de veille, audit en e-réputation ou *social listening* (Talkwalker, Digimind, Linkfluence, Sindup, Synthesio, Meltwater, Visibrain, par exemple) mettent à disposition des pures players et agences notamment, une plateforme visant à collecter un ensemble de données (*inputs)* génériques pour répondre aux besoins des clients (suivi de mentions social média, détection d'influenceurs, analyse de sentiment, détection de co-occurence de termes à titre d'exemple).

Les *pure player* quant à eux, qui sont les agences spécialisées dans la mesure de la réputation, proposent, au même titre que les éditeurs de solutions et notamment par l'utilisation de leurs plateformes, une offre complète construite sur le principe du "sur mesure" et donc modulaire. Cette offre complète s'appuie sur de la mesure mais aussi, et à l'issue du processus, de la mise en place d'un processus de gestion de la réputation sur les diverses plateformes socionumériques : modération de contenus, *content marketing*, *Community management*, déréférencement, *flooding.* C'est notamment le cas des agences Net reputation, Atchik, Netino, Human to Human, Laudans, Reputation Squad par exemple.

À la différence de ces derniers dont l'objectif est de proposer un ensemble de services qui cumulés composent l'offre de gestion de la réputation, l'offre des agences généralistes, ad hoc elle et de certains des outils gratuits, est construite sur le principe des modèles *Ad hoc*. Si les agences *Ad hoc* se réfèrent en certains points aux *pure players,* contrairement à ces derniers, elles ne proposent pas d'offre générale ou générique de mesure, mais s'appuient sur un aspect singulier du processus. C'est notamment le cas par exemple des agences de réputation qui ne proposent qu'un service de gestion de crise ou des avis négatifs (Rispote, Sitbon James, Netwash, La Frenchcom), des cabinets d'avocats spécialisés dans le recours légal en cas d'atteinte à la réputation (Haas Avocat, Pivocshi, Bensoussan, Porlon) ou des agences spécialisées dans le domaine d'activité spécifique (Taste dans l'agroalimentaire ; WKDO ou ma réputation en ligne dans la santé ; My hôtel réputation dans le domaine hôtelier par exemple).

Sur le même principe, par agences généralistes, nous entendons l'ensemble des agences de communication digitale, marketing digital ou relations publiques numériques qui, eu égard à l'essor des questions de réputation, se sont adaptées avec la mise en place d'actions spécifiques selon leur domaine d'expertise. Certaines, au même titre que les agences ad hoc, proposent donc une offre spécifique relative soit à la construction ou à la gestion de la réputation. L'offre qui vise à accroître la visibilité notamment sur les moteurs de recherche[444] est proposée par les agences webmarketing (Maliboo, La nantaise du web, Kel référencement, Eskimoz, Drift digital, Digimood, Smart web Group par exemple) : référencement ou déréférencement, *flooding* ou encore *black Hat SEO* pour certains acteurs.

[444] Avec pour finalité de construire la réputation ou d'accroître la popularité

Sur les applications gratuites, il est important de souligner qu'il en existe plusieurs catégories selon les usages spécifiques. À titre d'exemple, la plateforme de *personal branding* Webmii propose un score de réputation personnel avec un tableau de bord des mentions, Social bearing et Twazzup permettent de faire de la veille sur le microblogging Twitter, Crowdtangle met à disposition une extension permettant de suivre le partage des liens sur les médias sociaux.

Enfin, et dernière catégorie, les méthodes qui s'appuient sur les approches sommaires et notamment relatives aux sites d'avis. En effet, si à l'origine les questions de réputation étaient l'apanage des sites d'avis clients via notamment le e-commerce, ces derniers (les *recommander sites*) sont aujourd'hui considérés comme des tiers de confiance et une des sources de collecte des données à laquelle fait de plus en plus appel les éditeurs de logiciels via notamment des partenariats[445]. On retrouve à titre d'exemples les plateformes, Guest suite, Trust folio, qui proposent une recension des avis des internautes sur une entité et de les mettre en avant sur les plateformes corporate. Certaines, comme l'agence Ad-works par exemple, proposent en sus, la suppression des avis négatifs. Si la méthodologie des tiers de confiance est relativement simple (d'où l'approche sommaire), ils offrent l'avantage pour certains (Avis vérifiés par exemple)[446], de collecter de la donnée vérifiée et émanant d'acteurs ayant « réellement » interagi avec l'entité dont on mesure la réputation. Ce qui questionne notamment sur l'introduction de mécanismes de confiance dans la mesure de la réputation, questionnement appréhendé par la littérature et à l'origine (en partie) des modèles sophistiqués.

Notre processus de catégorisation des modèles théoriques a permis d'identifier, rappelons-le, quatre (4) catégories de modèles théoriques. Si les modèles ad hoc, génériques et sommaires ont inspiré dans une certaine mesure les pratiques professionnelles, quid des modèles sophistiqués?

---

[445] C'est le cas par exemple de Talkwalker qui a développé récemment (en 2020) un partenariat avec Net Reviews afin de crawler les données issues du site d'avis

[446] En l'occurrence ceux qui fonctionnent selon une approche centralisée

- **Notre seconde analyse est qu'a*ucune des approches de mesure proposée par les professionnels de la réputation en ligne ne s'appuie sur des modèles sophistiqués :***

Le principe des modèles sophistiqués de mesure de la réputation issus de la littérature scientifique peut se résumer comme suit : proposer, à partir de techniques algorithmiques liées aux approches en intelligence artificielle, un processus pour évaluer plusieurs métriques de la réputation (l'influence ou la popularité, la confiance, l'identité ou l'image), avec l'utilisation de mécanismes sécuritaires notamment de détection de biais ou de données frauduleuses. S'il est un constat clé qui ressort de notre analyse, c'est que les modèles sophistiqués sont, pour la plupart, des modèles hybrides, c'est-à-dire de confiance et de réputation. En effet, dans la conception des modèles hybrides, la réputation et la confiance sont deux phénomènes intrinsèquement liés (avec la confiance comme *output* de la réputation), et mesurer la réputation suppose une approche bidirectionnelle : la réputation est ici perçue autant au sens de l'émetteur que du récepteur. Ce qui suppose, que le processus de mesure consiste d'une part à évaluer une entité à partir d'un ensemble de données collectées mais aussi et surtout, à évaluer les agents à l'origine des informations collectées.

Les modèles sophistiqués s'appuient sur le postulat selon lequel la réputation devrait émaner d'agents ayant réellement interagi de manière directe ou indirecte (d'où la collecte d'informations issues de sources directes et indirectes) avec l'entité ou l'objet évalué. Ce qui suppose de fait, une évaluation de l'agent qui évalue. Si l'interaction directe utilise dans certains modèles (notamment dans leur conception) l'intégration de procédés issus de ceux des tiers de confiance comme preuve, l'évaluation des sources indirectes suppose la mise en place de mécanismes sécuritaires et de détection des biais. C'est le cas par exemple des modèles BDI+REPAGE de Pinyol & al. (2012), REPAGE de Sabater & al. (2006), TRAVOS de Teacy & al. (2006), LIAR de Muller & Vercouter (2006), PEERTRUST de Xiong & Liu (2004), FIRE de Huynh & al. (2004) et du modèle REGRET de Sabater & Sierra (2002).

En effet, dans leur modèle REGRET, Sabater & Sierra utilisent, comme nous l'avons vu, autant les données explicites (avis, commentaires, notes) que les informations de nature implicites (appartenance à un groupe par exemple) afin d'évaluer la réputation. Dans le modèle TRAVOS, les auteurs ont implémenté un mécanisme sécuritaire visant à punir les agents qui fournissent des informations fausses ou biaisées.

L'objectif visé par les approches sophistiquées, nous le voyons, est de garantir dans une certaine mesure l'intégrité, la fiabilité des données de réputation avec, dans le processus d'évaluation général, la mesure de la réputation de l'agent qui évalue. Ce qui, dans l'absolu, nous amène à la conclusion que ces approches ne sont pas encore totalement intégrées dans le processus de mesure proposé par les professionnels même si, certains éditeurs sont aux prémisses des modèles sophistiqués.

- ***Si certaines offres proposées par les professionnels s'appuient peu ou prou sur les savoirs académiques, il n'en demeure pas moins que beaucoup d'acteurs proposent en outre des prestations avec effet de scientificité, qui ne sont construites sur aucune heuristique ou méthodologie scientifique spécifique et ne s'appuient de fait sur aucun ancrage théorique*** :

En effet, à partir de notre analyse documentaire des sites corporate et de l'ensemble des documents (articles, livres blancs, rapports) proposés par les professionnels de la réputation, il nous est apparu que plusieurs d'entre eux s'appuyaient sur une approche commerciale avec effet de scientificité. C'est notamment le cas d'une majorité des agences *Ad hoc* (surtout généralistes) et de certains *pure players.* La raison est très simple : le marché de la réputation ayant connu un boom à partir des années 2009, les agences dans l'optique de "*s'adapter aux tendances*" et "*répondre aux besoins du marché*" comme l'ont notifié certains des interviewés lors des entrevues, se sont vues dans l'obligation de développer une offre liée à la réputation afin d'adresser de nouveaux segments de marché. Ces dernières, les agences, proposent de fait sur leurs pages corporate pour certains : des approches très sommaires en termes d'offres ou de méthodologie, des articles de blogs commerciaux avec l'utilisation de requêtes liées au concept sans offre précise présentée, plusieurs sites web pour la même entité ou offre pour d'autres, des données de contact dans les mentions légales inexistantes ou factices, ou encore l'utilisation de contenus dupliqués (*duplicate content*) par exemple.

Conclusion du chapitre 7

Ce chapitre visait, à la suite des approches méthodologiques utilisées dans nos travaux, à présenter d'une part le terrain sur lequel nous nous sommes appuyée dans notre phase de collecte du corpus, et d'autre part, à proposer des éléments de réponses à nos questions de recherche. Dans la première section, il s'agissait de présenter le processus d'échantillonnage à l'issue duquel nous avons identifié vingt (20) entités auprès desquelles nous avons réalisé nos entrevues. À partir de notre cartographie des acteurs du marché de la réputation et afin de garantir une certaine forme de représentativité, les entrevues ont été réalisées auprès d'éditeurs de solutions logicielles, de *pure players,* d'agences ad hoc et d'agences généralistes.

À partir des six catégories d'analyse issues de notre processus de réduction des données tel que le prévoit l'analyse inductive, nous avons restitué les résultats dans les catégories correspondantes.

L'analyse des résultats issus de nos travaux à la fois sur les approches théoriques et professionnelles de mesure de l'e-réputation nous ont par la suite permis de répondre à notre problématique centrale.

Conclusion de la partie III

Cette dernière partie visait essentiellement à répondre à notre problématique de départ. Si la problématique centrale ayant conduit à nos travaux était d'identifier l'utilisation de manière implicite ou non d'approches théoriques par les praticiens dans leurs stratégies de mesure de la réputation, notre approche méthodologique basée sur l'approche inductive générale s'est déclinée sous deux principales formes. Dans la première, il s'agissait de déterminer l'approche heuristique utilisée dans la conception théorique des modèles. La seconde section elle, consistait à déterminer via notre terrain, les approches méthodologiques mobilisées par les professionnels de la réputation.

Afin de proposer des éléments de réponse, nous nous sommes donc d'une part appuyée sur les données de terrain circonscrites via notamment les catégories d'analyse prédéfinies. Ces dernières font notamment référence à plusieurs thématiques : la perception de l'objet e-réputation comme actif intangible, la corrélation entre les questions de réputation traditionnelle et celle d'e-réputation, l'offre de mesure qui s'articule autour du "sur mesure", la / les méthode (s) de mesure en perpétuelle construction, les *output*s, et les perspectives futures des problématiques et questions liées à la réputation. Cette phase nous a permis, et à partir de notre postulat selon lequel l'e-réputation et la réputation représentent le même objet avec des approches en certains points différents, d'identifier l'ensemble des éléments impactés par la mutation, et de déterminer les méthodes de mesure proposées par les acteurs du marché.

Au vu de ces résultats, et pour répondre à notre problématique, si certaines méthodes utilisées par une minorité de praticiens s'appuient de manière implicite dans une certaine mesure sur des modèles théoriques (les modèles ad hoc, génériques et sommaires), plusieurs des offres sont construites autour de la recherche d'un effet de scientificité, faisant fi d'ancrages ou de savoirs académiques.
Quid de nos hypothèses de travail? Sont-elles vérifiées, infirmées? Quelles sont nos perspectives futures de recherche? Nous reviendrons sur ces éléments dans la conclusion générale de nos travaux de recherche.

CONCLUSION GÉNÉRALE ET PERSPECTIVES

Nous proposons dans cette conclusion, une analyse synthétique de nos travaux, analyse qui s'appuie sur trois principaux axes : nos hypothèses de recherche, les éléments à retenir et les perspectives de recherches futures. S'il est de coutume de présenter les réponses aux hypothèses dans le "*corps*" de la thèse, nous avons cependant choisi de les exposer dans cette partie pour nous appuyer sur ces dernières afin de proposer des axes de réflexion personnels qui pourront faire l'objet de travaux futurs.

1. Quelles réponses avons-nous pu apporter à nos hypothèses ?

Au début de nos travaux, nous exposions, à partir de notre analyse de la littérature scientifique et de notre analyse exploratoire des approches professionnelles sur l'objet e-réputation, certaines hypothèses de travail, sur lesquelles nous nous proposons de revenir.

Hypothèse 1 :

On peut postuler que l'e-réputation est l'évaluation d'un « objet », qu'il s'agisse d'une personne au sens d'individualité ou d'organisation ou encore d'un produit ou service, à partir notamment des "traces" laissées de manière intentionnelle ou non. L'intérêt dont elle fait l'objet, en tant que bien immatériel désormais monnayable économiquement, expliquerait l'engouement des agences à en proposer des mesures.

Afin d'apporter des éléments de réponse à cette hypothèse, il nous apparaît intéressant de nous appuyer sur le Chapitre 1 de la première partie de nos travaux. Dans ce dernier, nous avons proposé une analyse comparative des divers courants de pensée qui ont étudié le concept et ce, afin de définir un cadre de l'objet réputation. Si elle est perçue par le prisme de la sociologie comme construit social; comme risque par celui de la psychologie; comme actif intangible, signal ou garantie de contrat par les économistes; comme représentation ou jugement par les gestionnaires ou encore comme signal, entre autres, en philosophie. Les SIC au sein desquelles s'inscrivent nos travaux l'appréhendent plutôt comme évaluation, mesure portant sur les opinions et informations.

En effet, à partir des interactions et actions de communication des internautes sur le web social et ses diverses plateformes, sont produites des données sous forme de "traces" qui sont diffusées, remixées, documentarisées, et qui, dès lors qu'elles sont collectées et traitées notamment par le biais de méthodes algorithmiques, produisent des indicateurs réputationnels sur l'entité en question. Si les questions de réputation existent depuis toujours, force est de constater qu'elles ont pris une ampleur significative depuis quelques années avec la création du concept d'e-réputation.

Une offre très importante s'est dès lors développée en France: *social listening*, relations publiques 2.0, *influencer marketing, social media management, brand content*, *community management*, *content marketing*, *sentiment analysis*, déréférencement, gestion de crise 2.0, bad buzz, *flooding, SEO, SEA, SEM, etc*.

Nous avons étayé l'idée d'un développement fulgurant du marché de la réputation, que nous avons décrit dans les premiers chapitres, avec des agences à la fois spécialisées et généralistes sur la question, ce qui s'explique par le caractère intangible, nouveau, mais aussi et surtout économiquement monnayable du concept.

Hypothèse 2 :

Avec l'essor du web social, la production et la diffusion exponentielle de "données" semble avoir favorisé le développement de nouvelles méthodes et pratiques de la part des professionnels de la réputation, tant dans la collecte que dans l'analyse des informations collectées. On peut supposer une forme d'opportunisme dans l'utilisation des métriques rendues possibles par les technologies et la recherche d'un effet de "scientificité" liées à l'utilisation de mesures quantitatives nombreuses.

Si à l'origine l'évaluation de la réputation, qui utilisait notamment les baromètres d'image et de notoriété, s'appuyait sur des approches plus ou moins rudimentaires, aujourd'hui, on assiste à une recrudescence d'outils et de solutions de collecte et de mesures, techniques favorisées par l'usage d'internet. En effet, lorsqu'on s'intéresse à la cartographie des acteurs du marché de la réputation, on constate la présence d'un noyau central composé notamment des éditeurs de solutions logicielles autour duquel gravitent les offres du marché proposées par les agences. Ces dernières rivalisent d'ingéniosité afin de proposer des offres basées sur des approches techniques (via l'utilisation des solutions logicielles) et construites autour d'une myriade de métriques et d'indicateurs.

D'une manière générale, les méthodes de collecte et d'analyse des données de réputation s'articulent autour d'approches algorithmiques d'intelligence artificielle (à mi-chemin entre les réseaux de neurones, la logique mathématique et les sciences de l'informatique) et plus précisément d'apprentissage automatique (*machine learning*).

Le principe est simple : développer un modèle et à partir d'un certain nombre de tests via des algorithmes (de classification, de *clustering*, de régression, etc.) et un jeu de données, apprendre à la machine à résoudre une problématique spécifique. En termes de mesure de la réputation, l'algorithme implémenté à la solution (d'audit, de veille ou *social listening*) aura comme tâche selon la finalité souhaitée, de *crawler* le web afin de collecter l'ensemble des métadonnées se référant à l'objet d'étude, de les classifier puis de procéder à leur analyse afin, de produire un ensemble de métriques à la fois quantitatives (données démographiques, nombre de mentions, portée, etc.) et qualitatives (analyse de sentiment, entités nommées, mots clés, verbatims, etc.) se rapportant à la réputation de l'entité dont il est question.
Les offres étudiées dans ce travail montrent une grande hétérogéneïté dans la qualité et la pertinence des prestations fournies, certaines très complètes, d'autres très basiques. Dans tous les cas, les métriques utilisées et plus ou moins justifiées, mises en scène dans des infographies ou des tableaux de bord débordant de chiffres nous permettent de dire qu'un effet de scientificité est recherché, de même qu'un effet d'expertise, quand de nombreuses offres se disent « sur mesure ».

Hypothèse 3 :

Dans la littérature scientifique, surtout anglo saxonne, il existe une myriade de modèles théoriques de mesure de la réputation et on peut supposer que ces derniers ont évolué à travers le temps, et pourraient intrinsèquement être liés à plusieurs phénomènes clés dont les notions de « confiance», « image » et « risque », concepts difficiles à appréhender et maîtriser sur la toile. On peut également supposer que les définitions et dynamiques de la « confiance » dans des environnements économiques et socio-politiques différents peuvent expliquer des variations dans les modèles théoriques existants.

Cette hypothèse soulève deux principaux questionnements : d'une part les attributs et caractéristiques utilisés dans la conception des modèles théoriques d'une part et d'autre part les relations de connexité entre les concepts de réputation et confiance, et de manière plus ou moins implicite de risque et d'image.

S’il est un constat qui au premier abord semble évident, c’est le lien étroit entre les deux premiers phénomènes, lien sur lequel nous avons par ailleurs proposé une analyse exhaustive dans le chapitre 4 et sur lequel trois principaux courants de pensées s’entrecroisent. Pour résumer nos propos et les lectures exposées dans notre travail, la mise en place d’une relation de confiance nécessitant à minima une interaction (au sens large du terme) qui fera l’objet d’une évaluation, la confiance est une conséquence (*output*) de la réputation. Ceci signifie que la confiance naît en aval de la réputation : une bonne réputation sera gage de confiance tandis qu'une mauvaise réputation entraînera la méfiance. Il est par ailleurs important de souligner, dans une seconde approche, que la confiance au fil du temps et des interactions tend à "s'émanciper" de la réputation et devient dans certains cas un *input* de l’e-réputation. Elle sera dès lors perçue non plus comme une expérience personnelle mais comme une “référence” au sens de recommandation qu’on ferait par exemple à un tiers sur la base de sa propre expérience.

Plusieurs modèles théoriques ont intégré cet attribut dans leur conception de l’e-réputation et certains sont même allés au-delà en développant des modèles hybrides de confiance et de réputation qui visent d’une part à évaluer la réputation mais aussi et surtout le degré de confiance attribué à une entité. C’est le cas par exemple des modèles de Sabater & Sierra (2002), Mui & al. (2002), Wang & Vassileva (2003), Teacy & al. (2006), Zhou & al. (2007), Fan & al. (2011).

Sur la question de la variation et de l’évolution des modèles, nous pouvons dire que les modèles conceptuels pris en compte dans nos travaux dont le premier date de 1998 (PAGERANK de Page et Brin) et le dernier de 2016 (Li & al. : DDOA), ont évolué dans le temps avec, d’une part, des attributs ou caractéristiques différents pris en compte dans leur conception, et d’autre part des domaines d'application variés. Sur l’ensemble des attributs, les modèles qu’on retrouve dans notre taxonomie s’appuient dans un premier temps essentiellement sur les concepts d’influence et de popularité, puis par la suite les modèles ont intégré de nouvelles métriques de la réputation telles que l’identité, l’image, et la confiance.

Sur le plan applicatif, on est passé de l'usage des moteurs de recherche, aux plateformes de e-commerce, en passant par les sites de recommandation, le partage de documents, les medias sociaux ou encore et plus récemment les domaines de l'intelligence artificielle (*cloud computing, smart grid*, *Vanet*, *Wireless sensor network*, etc.). Il semble que c'est avant tout l'évolution du domaine technologique et l'émergence des technologies liées au web et à l'intelligence artificielle, qui ont impacté les modèles d'affaire et ont fait évoluer les modèles théoriques étudiés.

Hypothèse 4 :

L'*offre proposée par les professionnels de la réputation et qui s'articule autour du maître mot "sur mesure", semble différer d'un praticien à l'autre et cela, selon plusieurs phénomènes : leur statut, les besoins et attentes du marché au sein duquel ils évoluent et notamment ceux des commanditaires, l'offre proposée par les concurrents, etc., ce qui nous amène à questionner l'existence d'une éventuelle méthode "universelle" de mesure de l'e-réputation. On peut supposer une mobilisation différenciée et stratégique des ressources théoriques, technologiques et d'accès aux données des acteurs de terrain dans la recherche d'un modèle économique commercialisable, plus qu'une recherche de validité des résultats, au regard des mesures convoquées dans la mesure de l'e-réputation.*

Notre analyse exploratoire du marché de la réputation a mis en exergue l'existence de six principaux acteurs : les éditeurs de solutions, les *pure players*, les agences *ad hoc*, les agences généralistes, les tiers de confiance et les applications gratuites. Pour chacun d'entre eux, l'offre de service proposée diffère : si certains *pure players* au même titre que certains éditeurs proposent une approche plus ou moins générique de mesure de la réputation, elle varie considérablement d'une organisation à l'autre. Certaines agences *ad hoc* et généralistes proposent en outre une offre partielle portant sur une des métriques de la réputation sans évaluer cette dernière dans son intégralité. Selon le statut de ces dernières, les offres auront plutôt une visée de mesure (audit, *social listening* par exemple), de gestion avec une approche légale par exemple (aide à la gestion des droits des données numériques, droit à l'oubli, etc.) ou de construction de la réputation (approche de marketing digital: indexation

dans les moteurs de recherche par exemple). Cette approche pluridimensionnelle rend *a priori* impossible l'identification d'une méthode unique et universelle de la mesure de la réputation. La diversité est la règle, d'où la généralisation d'offres "sur mesure" prônées par les prestataires ce qui, pour nous, est la marque d'un champ professionnel en recherche de maturité, plus attaché à la mise au point d'un modèle économique "commercial" efficace, et à la création d'un effet de scientificité qu'à la mise au point d'une mesure fiable, valide et reproductible de la réputation.

2. Quelles premières conclusions?

Dans cette partie, nous souhaitons revenir sur nos principaux apports et les contributions de ces travaux. Si on reprend notre problématique de départ, qui a jalonné l'ensemble des travaux et à partir de laquelle ont découlé les principaux questionnements clés et les hypothèses de travail, elle s'articulait comme suit:

« ***Analyser les modèles théoriques d'e-réputation existants dans la littérature scientifique, phase préliminaire qui par la suite permettra de questionner un éventuel rapport entre ces derniers et les méthodes d'audit comprises dans l'offre proposée par les professionnels de l' e-réputation ».***

Le premier constat qui ressort, avant de nous intéresser aux résultats, est relatif au concept "*d'audit*" employé au début de nos travaux de recherche. Par audit, terminologie proposée dans le champ de la communication organisationnelle (Odiorne (1954 ; Jacobson et Seashore, 1951; Davis, 1953 ; Greenbaum, 1974 ; Strenski, 1978 ; Gildea et Rosenberg, 1979 ; Wirtz, 1981 ; Willett, 1994 ; Booth, 1986 ; Downs, 1988), nous entendons l'approche d'analyse comparative visant à déceler d'éventuels écarts sur un phénomène donné, d'en identifier les origines afin de proposer des pistes d'amélioration objectives. Si au début de notre recherche (en 2015) le terme audit était encore employé afin d'appréhender l'objet réputation, force est de constater qu'au fil des années, le terme est devenu désuet avec depuis, l'emploi d'autres terminologies pour définir la pratique : étude d'image, veille en e-réputation, *social listening*, *social media management*, cartographie des acteurs et gestion de la réputation de manière plus générale.

Ce cadre posé, quels sont de fait nos apports sur la question ? Nos contributions se résument comme suit :

- ***La proposition d'un cadrage des concepts de réputation et d'e-réputation***

S'il est un constat clé à retenir de notre analyse, c'est que contrairement aux assertions selon lesquelles les concepts de réputation (au sens traditionnel) et d'e-réputation sont relatifs à deux objets distincts, les deux notions font, pour nous, référence à un seul et même objet. En effet, aujourd'hui, avec les nouveaux enjeux que posent les questions de réputation, le marché de l'e-réputation s'est adapté afin de répondre au mieux aux problématiques actuelles de la digitalisation, tout en conservant un objectif similaire.

Comme mentionné dans la première section du chapitre 5, les questions de réputation ont longtemps été réduites aux problématiques d'image au sens traditionnel (image perçue *versus* voulue). Aujourd'hui les études utilisent de nombreux marqueurs : des marqueurs liés aux questions légales (atteinte à la réputation par exemple), de visibilité et d'indexation, de notoriété et de *ranking,* d'engagement (capacité à faire parler de soi) et d'influence. La collecte et la mesure des données s'appuient dorénavant sur des méthodes algorithmiques, parfois issues des techniques d'intelligence artificielle et particulièrement de *machine learning*, technicité qui marque une rupture dans les approches.

- ***La définition pour plusieurs champs disciplinaires, du ou des objet (s) sur le(s)quel(s) porte la mesure de la réputation***

Il nous est apparu essentiel, et ce à partir du cadrage des diverses conceptions du concept de réputation à travers la littérature scientifique, de définir l'attribut ou objet sur lequel porte l'évaluation. Pour les économistes et gestionnaires, la mesure de la réputation s'appuie sur les actions passées, les attributs identitaires de l'entité et l'offre générale (produit / service) qu'elle propose. La sociologie, elle, conçoit la réputation à travers l'interaction sociale et le contexte (normes / attentes) au sein duquel elle s'effectue. En SIC, la mesure de la réputation suppose la prise en compte des opinions (au sens d'opinion publique de D'Almeida) et des informations (au sens de signaux informationnels pour Alloing par exemple). Enfin, si la psychologie essaie de mesurer les états mentaux et la confiance dans le processus d'évaluation, la philosophie, elle, pense la réputation au sens de traces ou indices.

À partir de ce constat, nous avons proposé (dans la section 5 du chapitre 1) un cadre des différentes composantes nécessaires à la construction de la réputation (il est important ici de souligner que ce cadre ne vise en aucun cas à définir l'ensemble des étapes prises en compte dans la construction de l'e réputation[447]). Ces attributs font référence au contexte, aux agents en présence, à l'objet de mesure et à l'historique sur lequel s'appuiera l'évaluation.

- ***La proposition d'une taxonomie des modèles de mesure qu'on peut regrouper selon quatre grandes catégories***

Si comme nous l'avons vu (dans le chapitre 1 de la partie 2) la littérature scientifique anglo-saxonne sur les modèles théoriques est foisonnante, les critères de classification utilisés par les chercheurs sont divers et variés et s'appuient entre autres sur les catégories d'interactions, le type et les sources de données, le type de visibilité, la granularité, les questions de temporalité, l'approche mathématique de mesure utilisée, le processus de stockage, l'objet sur lequel porte la mesure, la valeur finale de la réputation. Réunir l'ensemble des modèles produits de 1998 à 2016 a été un travail lourd, désormais à la disposition des chercheurs qui pourront utiliser ces sources pour de nouvelles études.

On notera que les tentatives de classification des modèles que nous avons étudiées intègrent trois catégories : les modèles de confiance, les modèles de réputation et les modèles hybrides qu'on retrouve sous l'appellation "*Trust and reputation models*". Notre approche s'appuie sur les modèles de réputation et les modèles hybrides (auxquels nous avons uniquement emprunté les approches utilisées dans l'évaluation de la réputation). À partir de là, nous avons défini de nouvelles dimensions de classification des modèles qu'on peut regrouper selon nous en quatre catégories : les modèles sommaires, ad hoc, modulaires et les approches sophistiquées.

À l'issue de cette phase d'analyse, nous avons proposé (en annexe 1) un tableau de classification des modèles théoriques de mesure de la réputation, tableau qui s'appuie sur les quatre catégories définies étayées par plusieurs critères que nous avons développés.

---

[447] Processus de construction abordé par plusieurs auteurs dont notamment Camille Alloing dans sa thèse. Camille Alloing. Processus de veille par infomédiation sociale pour construire l'e-réputation d'une organisation. Approche par agents-facilitateurs appliquée à la DSIC de La Poste. Sciences de l'information et de la communication. Université de Poitiers, 2013. Français. NNT : 526 . tel- 00915004

- ***La proposition d'une nouvelle cartographie des acteurs du marché de la réputation***

Si depuis 2009 plusieurs auteurs s'attellent à proposer des cartographies afin de circonscrire les acteurs qui évoluent et proposent une offre de service en réputation, l'évolution du marché en perpétuel questionnement et construction rend la proposition d'une taxonomie unique et standard quelque peu complexe. Cette complexité, encore accrue aujourd'hui, s'explique tout simplement par l'absence d'un consensus sur ce qu'est la réputation et ses nombreuses déclinaisons. À partir des précédents travaux auxquels nous avons emprunté certaines terminologies, nous avons proposé une nouvelle catégorisation des acteurs avec l'essor depuis peu d'un nouvel entrant.

Les prestataires qui composent le marché de la réputation en ligne peuvent être regroupés selon six catégories : les *pure players,* les agences ad hoc, les agences généralistes, les éditeurs de solutions logicielles, les fournisseurs d'applications gratuites et les tiers de confiance. Cette dernière catégorie, même si elle n'est pas encore totalement intégrée à l'environnement professionnel et qu'elle en est encore à ses débuts fait partie des acteurs majeurs, permettant *de facto* d'établir un lien entre les questions de réputation et de confiance. L'annexe n°2 propose une classification de l'ensemble des acteurs du marché selon la catégorie à laquelle ils appartiennent.

Eu égard à ces constats, il convient de retenir de nos travaux que si certains des acteurs (une minorité) s'appuient peu ou prou dans leurs méthodes sur des savoirs académiques, la plupart surfe sur une forme d'opportunisme dans la proposition de leur offre construite sur un modèle économique commercial avec une recherche d'un effet de scientificité, effet de scientificité construit autour de l'utilisation d'un nombre important de données chiffrées ce qui, à ce stade, semble satisfaire les clients. Le qualificatif "*sur mesure*" accolé à de nombreuses offres ne répond à aucun ancrage méthodologique défini et spécifié.

3. Quelles sont les limites de nos travaux et les perspectives futures?

Si nos ambitions au début de ces travaux étaient d'une part de réaliser une analyse exhaustive de l'ensemble de la littérature scientifique sur les modèles développés et d'autre part de conceptualiser les méthodes de mesure proposées par les professionnels, nous nous sommes très rapidement retrouvée confrontée à certaines difficultés qui constituent autant de limites à nos travaux.

En effet, l'une des caractéristiques clés autour desquelles sont construits les modèles de mesure de la réputation est relative au modèle mathématique utilisé, qui s'appuie comme nous l'avons vu sur plusieurs approches : probabilistes, sur la logique subjective, sur les approches floues ou encore sur celles basées sur les flux. Par ailleurs, l'e-réputation, et notamment son évaluation s'appuie désormais sur une grande technicité rendue possible par l'utilisation des approches en intelligence artificielle. Le *machine learning* qui est aujourd'hui devenu incontournable, s'appuie sur les concepts et techniques précédemment évoqués : le *sentiment analysis*, la fouille d'opinion (*opinion mining*) ou *data mining* ou encore la *data science* (science des données). Si nous avons évoqué ces questions, elles mériteraient cependant un approfondissement pluridisciplinaire réalisé en collaboration avec des spécialistes de ces domaines pour déconstruire les processus utilisés, et mesurer la pertinence des calculs qui manquent singulièrement de transparence, rendant leur évaluation souvent impossible.

Une seconde limite de nos travaux est relative à l'aspect managérial, au terrain et notamment à la difficulté voire l'impossibilité d'obtenir certaines données. Les interviewés, de par leur poste (fondateur, chargé de réputation, consultant *social listening*, consultant e-réputation et influence, *account manager*, entre autres) devaient respecter des clauses de confidentialité et les données collectées étaient parfois très généralistes, à la limite du superficiel. Ce qui nous a amené, entre autres, à effectuer un travail d'analyse de l'ensemble des documents (site web et notamment blogs, livres blancs et *ebook*) disponibles, et à participer à un certain nombre d'événements numériques organisés par ces acteurs, pour approfondir nos informations.

Concernant les perspectives et ouvertures que nos travaux ont permis d'envisager, nous aimerions en évoquer quelques unes.

Nous avons constaté une pluralité de métiers ou intitulés de poste des professionnels interrogés (*social media manager*, *online reputation manager*, *community manager*, analyste des données ou *data analyst*, chargé d'études, consultant e-réputation et *SEO* par exemple). Il pourrait être intéressant de questionner l'ensemble de ces intitulés, pour proposer une terminologie spécifique aux métiers de la réputation.

Le marché de la réputation, comme nous l'avons vu, est en construction avec un développement fulgurant des offres qui sont diverses et variées. Si comme nous l'avons souligné, seulement quelques acteurs (pour la plupart les éditeurs, les *pures players* et les agences généralistes) s'appuient sur un ancrage théorique avec une méthode d'évaluation quelque peu scientifique, comment dès lors aider à détecter les offres commerciales peu sérieuses? Cette question pourra également faire l'objet de travaux ultérieurs.

Enfin, la dernière piste pour poursuivre notre étude concerne les métriques prises en compte dans la mesure de la réputation. Sur la question de l'influence numérique, la plupart des acteurs l'ont intégrée à leur pratique. En ce qui concerne la confiance, si la littérature a complètement intégré le concept avec le développement de modèles hybrides et de confiance (Marsh, 1994[448]; Castelfranchi & Falcone, 1998[449]; Schillo et al., 2000[450]; Abdul Rahman et Hailes, 2000-2005[451]; Esfandiari and Chandrasekharan, 2001[452]; N. Griffiths, 2006 par exemple), les approches professionnelles sont encore à la traîne sur le sujet. Comme nous

---

[448] S. P. Marsh. *Formalising Trust as a Computational Concept*. PhD thesis, Department of Mathematics and Computer Science, University of Stirling, Scotland, UK, 1994

[449] Castelfranchi, Cristiano, et Rino Falcone. « Principles of trust for MAS: Cognitive anatomy, social importance, and quantification ». In *Proceedings International Conference on Multi Agent Systems (Cat. No. 98EX160)*, 72-79. IEEE, 1998.

[450] Michael Schillo, Petra Funk & Michael Rovatsos (2000) *Using trust for detecting deceitful agents in artificial societies*, Applied Artificial Intelligence, 14:8, 825-848, DOI: 10.1080/08839510050127579

[451] Abdul-Rahman, Alfarez, et Stephen Hailes. « *Supporting trust in virtual communities* ». In *Proceedings of the 33rd annual Hawaii international conference on system sciences*, 9 pp. vol. 1. IEEE, 2000.

[452] Esfandiari, B. & Chandrasekharan, S. (2001). *On How Agents Make friends: Mechanisms for Trust Acquisitio*n. In: *Proceedings of the Fourth Workshop on Deception, Fraud and Trust in Agent Societies, Montreal, Canada*

l'avons constaté, l'approche utilisée par les tiers de confiance n'est pas encore complètement intégrée à l'écosystème de la réputation. On pourrait se demander comment les professionnels peuvent intégrer dans leurs offres, un mécanisme afin de garantir notamment l'intégrité des données collectées?

Par ailleurs, afin d'avancer dans la compréhension du phénomène d'e-réputation, on pourrait également questionner plus précisément le poids respectif des items ou attributs mesurés dans les processus d'évaluation et la production des « scores » d'e-réputation. En effet, si on part du postulat que la mesure de la réputation en ligne suppose une compréhension de l'ensemble des objets qui lui sont associés (l'influence ou la popularité, la confiance, l'image et l'identité), comment dès lors hiérarchiser ces items? Ce processus d'ordonnancement nécessiterait, lui aussi, une approche pluridisciplinaire, faisant se côtoyer les SIC, mais aussi la psychosociologie, la psychologie sociale et la mobilisation de compétences approfondies en statistiques, pour valider, éventuellement de manière expérimentale, la validité et la fiabilité des scores de réputation produits.

BIBLIOGRAPHIE

A

Abdul-Rahman, A., & Hailes, S. (2000). Supporting trust in virtual communities. *Proceedings of the 33rd annual Hawaii international conference on system sciences*, 9 pp. vol. 1.

Aboulwafa S. and Bahgat R. 2010. DiReCT: Dirichlet-based reputation and credential trust management. *Informatics and Systems (INFOS)*, 2010 The 7th International Conference on, pp. 1–8, 28–30 March 2010.

Adamopoulou, A. A., & Symeonidis, A. L. (2014). A simulation testbed for analyzing trust and reputation mechanisms in unreliable online markets. *Electronic Commerce Research and Applications*, *13*(5), 368-386.

Agarwal, J., Osiyevskyy, O., & Feldman, P. M. (2015). Corporate Reputation Measurement : Alternative Factor Structures, Nomological Validity, and Organizational Outcomes. *Journal of Business Ethics*, *130*(2), 485-506.

Aggarwal, R., & Gupta, L. (2017). A Hybrid Approach for Sentiment Analysis using Classification Algorithm. *IJCSMC*, Vol. 6, Issue. 6.

Aljazzaf Z. M., Perry M., and Capretz M. A. M. (2010). Online trust: Definition and principles. *Proceedings of the Computing in the Global Information Technology (ICCGI)*. 20–25 Sept. Valencia, Spain: IEEE, 163–168.

Alloing, C. (2013). *Processus de veille par infomédiation sociale pour construire l'e-réputation d'une organisation. Approche par agents-facilitateurs appliquée à la DSIC de La Poste*. Sciences de l'information et de la communication. Université de Poitiers. Français. NNT : 526 . tel- 00915004.

Alloing, C. (2010). De la mauvaise appréhension du risque image sur le web, au vol d'informations grises par des techniques de social engineering « 2.0 ». *Sécurité et stratégie*, *3*(1), 36-42.

Alloing, C. (2012). De surveiller à « prendre soin » : Comment repenser la veille sur les réseaux sociaux numériques en termes de management de réseaux d'acteurs ? *Revue internationale d'intelligence economique*, *Vol 4*(1), 55-70.

Alloing, C. (2013). La médiation documentaire sur les réseaux socionumériques comme vecteur de l'e-réputation ? *Communication et organisation*, *43*, 73-84.

Alloing C. (2014). (e)Réputations et identités numériques : Vers une nouvelle forme d'autorité ?

Alloing, C. (2015). « Votre entreprise plus nette sur le Net ». Comment les professionnels fabriquent-ils les publics et les imaginaires de l'e-réputation ? *Communication. Information médias théories pratiques*, *vol. 33/2*.

Alloing, C. (2016). La fabrique des réputations selon Google. *Communication & langages*, *188*, 101-122.

Alloing, C. (2017). La réputation pour questionner l'autorité informationnelle : Vers une. *Quaderni*, *93*, 33-41.

Alloing, C., & Moinet, N. (2010). Des réseaux d'experts à l'expertise 2.0. *Les Cahiers du numérique*, *Vol. 6*(1), 35-53.

Alloing, C., Pierre, J. (2012). Construire un cadre d'analyse avec les SIC pour comprendre les pratiques et les enjeux de la réputation en ligne (des individus et des organisations). hal-00704007.

Andonova Y. et Vacher B., « Nouvelles formes de visibilité des individus en entreprise : technologie et temporalité », *Communication et organisation* [En ligne], 44 | 2013, consulté le 15 mars 2017.

Arnaud, M. (2013). Démonétiser l'identité numérique. *Médium*, *16-17*, 191-199.

Ashri R., Ramchurn S. D., Sabater J., Luck M., and Jennings N. R. (2005). Trust evaluation through relationship analysis. *In Proceedings of the 4th International Joint Conference on Autonomous Agents and MultiAgent Systems*, 2005, pp. 1005–1012.

Aula, P. (2010). Social media, reputation risk and ambient publicity management. *Strategy and Leadership*, *38*(6).

Ayday E., (2011). Iterative algorithms for trust and reputation management and recommender systems, PhD Thesis, *School of Electrical and Computer Engineering*, Georgia Institute of Technology.

B

Barbosa, Hugo & Morais, Telmo. (2016). Study on computational trust and reputation models. In *High Performance Network Studies* (pp.58-66)

Barnett, M. L., Jermier, J. M., & Lafferty, B. A. (2006). Corporate Reputation : The Definitional Landscape. *Corporate Reputation Review*, *9*(1), 26-38.

Beauvisage, T., & Mellet, K. (2016). Travailleurs du like, faussaires de l'e-réputation. *Réseaux*, *197-198*, 69-108.

Becker, K., & Nobre, H. (2014). Social Network Reputation Management : An International Study. *Journal of Promotion Management*, *20*(4), 436-451.

Belleguic, T., Coutard, J., & Doueihi, M. (2011). Les enjeux de la réputation à l'ère du numérique. *Sécurité et stratégie*, *6*(2), 22-27.

Bertocco C., and Ferrari C. (2008). Context-dependent reputation management for soft security in multi agent systems. *In IEEE/WIC/ACM International Conference on Web Intelligence and Intelligence Agent Technology* (WI-IAT'08), 9–12 Dec. 2008, Sydney, NSW, Australia, Vol. 3, pp. 77–81.

Beuscart, J.-S., Chauvin, P.-M., Jourdain, A., & Naulin, S. (2015). La réputation et ses dispositifs. *Terrains & travaux*, *N° 26*(1), 5-22.

Beuscart, J.-S., & Mellet, K. (2015). La conversion de la notoriété en ligne. *Terrains & travaux*, *N° 26*(1), 83-104.

Bing L., (2012). *Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies*, Lecture #16. Series Editor: Graeme Hirst, University of Toronto.

Boistel, P. (2008). La réputation d'entreprise : Un impact majeur sur les ressources de l'entreprise. *Management Avenir*, *n° 17*(3), 9-25.

Boistel, P. (2013). Web 2.0. *Communication. Information médias théories pratiques*, *Vol. 31/1*.

Boistel, P. (2014). Réputation : Un concept à définir. *Communication et organisation*, *46*, 211-224.

Boistel, P. (2016). La réputation d'entreprise n'est-elle qu'un concept de réception ? *Sens-Dessous*, *N° 18*(2), 29-37.

Boullier, D., & Lohard, A. (2012). Opinion mining et Sentiment analysis : Méthodes et outils. *OpenEdition Press*.

Boullier, D., & Lohard, A. (2015). Médiologie des réputations numériques. *Terrains & travaux*, *N° 26*(1), 105-125.

Boullier, D., Crépel, M. & Jacomy, M. (2016). Zoomer n'est pas explorer: Spatialiser les graphes, catégoriser et (dé)construire les réseaux. *Réseaux*, 195, 131-161.

Boutin, E., Liu, P., & Buisson, L. (2008). Veille d'image sur Internet : Enjeux, méthodes, limites. *Communication et organisation*, *34*, 98-114.

Boyd, J. (2002). In community we trust : Online security communication at eBay. *Journal of Computer-Mediated Communication*, *7*(3), JCMC736.

Braga, D. D. S., Niemann, M., Hellingrath, B., & Neto, F. B. D. L. (2018). Survey on Computational Trust and Reputation Models. *ACM Computing Surveys*, *51*(5), 1-40.

Brin S. et Page L. (1998). The PageRank Citation Ranking : Bringing Order to the Web*, Computer Science Department,* Stanford University.

Brin S. et Page L., The Anatomy of a Large-Scale Hypertextual Web Search Engine, *Computer Science Department*, Stanford University,

Brodin, O., & Magnier, L. (2013). Le développement d'un index d'exposition de soi dans les médias sociaux : Phase exploratoire d'identification des indicateurs constitutifs. *Management & Avenir*, *58*, 144-168.

Bromley, D. B. (1993). Reputation, image and impression management. John Wiley & Sons. Ltd, 312p, *ISBN* 0471938696.

Bromley, D. B. (2000), Psychological Aspects of Corporate Identity, Image and Reputation, *Corporate Reputation Review*, 3(3), 240-252.

Buche, A., Chandak, D. M. B., & Zadgaonkar, A. (2013). Opinion Mining and Analysis : A survey. *International Journal on Natural Language Computing*, *2*(3), 39-48.

Buchegger S., Le Boudec, J. Y., (2004). A Robust Reputation System for P2P and Mobile Ad-hoc Networks, *In Proceedings of P2PEcon 2004*, Harvard University, Cambridge MA, U.S.A.

Buskens, V. (1998). Social Networks and the Effect of Reputation on Cooperation. *Proceedings of the 6th International Conference on Social Dilemmas*.

C

Cadel P. (2010). Le marché de l'e-réputation : Du positionnement fonctionnel aux enjeux technologiques*, Les cahiers du numérique,* 2010/4 Vol. 6 | pages 111 à 121.

Carayol, V. (1994). L'impossible audit de communication interne. *Communication et organisation*, *5*.

Carayol V.et Gramaccia G., « Modèles et modélisations, pour quels usages », *Communication et organisation* [En ligne], 30 | 2006, mis en ligne le 21 juin 2012, consulté le 5 octobre 2016. URL : http://journals.openedition.org/communicationorganisation/3441

Carayol V. et Gramaccia G., « Recherches émergentes en communication des organisations », *Communication et organisation* [En ligne], 47 | 2015, mis en ligne le 18 avril 2019, consulté le 24 août 2019. URL : http://journals.openedition.org/communicationorganisation/4866

Carayol V., Lépine V., Morillon L. (2020). Le côté obscur de la communication des organisations. Maison des Sciences de l'Homme d'Aquitaine (MSHA), 234 p., 978-2-85892-605-3. ⟨hal-02525979⟩

Carbo, J., Molina, J., & Davila, J. (2002). Trust management through fuzzy reputation. *International Journal of Coopeative Information Systems* 12, 1, 135–155

Carbo, J., Molina, J., & Davila, J. (2002a). Comparing predictions of sporas vs. A fuzzy reputation agent system. 3rd International conference on fuzzy sets and fuzzy systems, Interlaken, 147-153.

Carbo, J., Molina, J. M., & Davila, J. (2002b). Introduction of newcomers into a fuzzy reputation agent system. *Applied informatics proceedings*, 314-319.

Cardon, D. (2009a). Le design de la visibilité. *Réseaux*, *n° 152*(6), 93-137.

Cardon, D. (2009b). L'identité comme stratégie relationnelle. *Hermès, La Revue*, *n° 53*(1), 61-66.

Cardon, D. (2013). Du lien au like sur Internet. *Communications*, *n° 93*(2), 173-186.

Cardon, D., Fouetillou, G., & Roth, C. (2015). Topographie de la renommée en ligne. *Réseaux*, *n° 188*(6), 85-120.

Carter, J., Bitting, E., & Ghorbani, A. A. (2002). Reputation formalization for an information–sharing multi–agent system. *Computational Intelligence*, *18*(4), 515-534.

Caruana, A., & Chircop, S. (2000). Measuring Corporate Reputation : A Case Example. *Corporate Reputation Review*, *3*(1), 43-57.

Casare, S., & Sichman, J. (2005). Towards a functional ontology of reputation. *Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems - AAMAS '05*, 505.

Castelfranchi, C., & Falcone, R. (1998). Principles of trust for MAS: Cognitive anatomy, social importance, and quantification. *Proceedings International Conference on Multi Agent Systems (Cat. No. 98EX160)*, 72-79.

Castelfranchi, C., & Falcone, R. (2001). Social trust: A cognitive approach. *In Trust and Deception in Virtual Societies*, pp. 55–90. Kluwer Academic Publishers.

Castelfranchi C., Falcone R., and Pezzulo G. (2003). Trust in information sources as a source for trust: A Fuzzy approach. *In Proceedings of the Second International Joint Conference on Automonous Agents and Multiagent Systems (AAMAS)*, ACM Press, New York, USA, pp. 89–96

Castellano S., Dutot V. (2013). *Une analyse de l'e-réputation par analogie ou contraste avec la réputation : Une approche par les médias sociaux*.

Castellano, S., Khelladi, I., Chipaux, A., & Kupferminc, C. (2014). The influence of social networks on e-reputation : How sportspersons manage the relationship with their online community. *ResearchGate*, *10*(4), 65-79.

Center, E. B. I. (2011). TREMA: A Tree-based Reputation Management Solution for P2P Systems. *International Journal on Advances in Security,* Volume 4, Number 3 & 4, 2011.

Ceruzzi, P. E. (2012). Aux origines américaines de l'Internet : Projets militaires, intérêts commerciaux, désirs de communauté. *Le Temps des médias*, *n° 18*(1), 15-28.

Charest, F., Alcantara, C., Lavigne, A., & Moumouni, C. (2017). *E-réputation et influenceurs dans les médias sociaux : Nouveaux enjeux pour les organisations*. *PUQ*.

Charest, F., & Bouffard, J. (2015). The characteristics of the e-influence of Community Managers : Issues for the e-reputation of organizations. *Public Relations Review*, *41*(2), 302-304.

Charest, F., & Doucet, A. (2014). Les animateurs de communauté et les professionnels en relations publiques : Enjeux sur la e-réputation des organisations. *Communiquer. Revue de communication sociale et publique*, *11*, 63-80.

Charest, F., Gauthier, A.-M., & Grenon, F. (2013). Appropriation et stratégies d'intégration des médias sociaux par les professionnels de la communication. *Communication et organisation*, *43*, 269-280.

Charest F., Lavigne A. et Moumouni C. (2015), Médias sociaux et relations publiques, Québec, Presses universitaires du Québec. *Communication. Information médias théories pratiques*, *vol. 34/1*.

Chauvin, P.-M. (2013). La sociologie des réputations. *Communications*, *n° 93*(2), 131-145.

Chun, R. (2005b). Corporate reputation : Meaning and measurement. *International Journal of Management Reviews*, *7*(2), 91-109.

Chun, R., & Davies, G. (2001). E-reputation : The role of mission and vision statements in positioning strategy. *ResearchGate*, *8*(4), 315-333.

Cieliebak, M., Dürr, O., & Uzdilli, F. (2013). Potential and Limitations of Commercial Sentiment Detection Tools. Conference: *ESSEM 2013*, *Emotion and Sentiment in Social and Expressive Media*.

Cismaru, D.-M. (2012). A Model for Assessing the Online Reputation of Public Actors.

Clardy, A. (2012). Organizational Reputation : Issues in Conceptualization and Measurement. *Corporate Reputation Review*, *15*(4), 285-303.

Colapinto, C., & Benecchi, E. (2014). The presentation of celebrity personas in everyday twittering : Managing online reputations throughout a communication crisis. *Media, Culture & Society*, *36*(2), 219-233.

Collomb C. (2013). *«* Penser la trace numérique : Quels paradigmes, quelles conséquences ? », séminaire sur le paradigme de la trace, Homo Ludens, UQAM.

Colombo, F., & Jeanneret, Y. (2014). Contrôle, identité, parrhèsia : Une approche foucaldienne du Web 2.0. *Communication & langages*, *180*, 7-24.

Conte, R., & Paolucci, M. (2002). Reputation in artificial societies : Social beliefs for social order (Vol. 6). *Springer Science & Business Media*.

Cormerais, F., Le Deuff, O., Lakel, A., & Pucheu, D. (2016). Les SIC à l'épreuve du digital et des Humanités : Des origines, des concepts, des méthodes et des outils. *Revue française des sciences de l'information et de la communication*.

Cormerais F. & Lakel A.(2019), « Recherches digitales et production des données, bouleversement des agencements pour le chercheur en SIC », *Etudes Digitales*, 6, Garnier classiques.

Costagliola, G., Fuccella, V. & Pascuccio, A. (2014). Towards a Trust, Reputation and Recommendation Meta Model. *Journal of Visual Languages & Computing*. 10.1016/ j.jvlc.2014.10.001.

Coutant, A. (2013). Quelle place pour l'innovation dans les médias sociaux ? *Communication et organisation*, *43*, 123-134.

Coutant, A., & Stenger, T. (2010). Processus identitaire et ordre de l'interaction sur les réseaux socionumériques. *Les Enjeux de l'information et de la communication*, *Volume 2010*(1), 45-64.

Coutant, A., & Stenger, T. (2011). Production et gestion d'attributs identitaires. *Les Cahiers du numérique*, *Vol. 7*(1), 61-74.

Coutant, A., & Stenger, T. (2012). Les médias sociaux : Une histoire de participation. *Le Temps des médias*, *18*, 76-86.

Cravens, K., Goad Oliver, E., & Ramamoorti, S. (2003). The Reputation Index: *European Management Journal*, *21*(2), 201-212.

D

D'Almeida, N. & Andonova, Y. (2014). La communication des organisations. Dans : Stéphane Olivesi éd., *Sciences de l'information et de la communication* (pp. 137-153). FONTAINE, France: Presses universitaires de Grenoble.

D'Almeida N., « L'opinion publique », *Hermès, La Revue*, 2014/3 (n° 70), p. 88-92. DOI : 10.3917/herm.070.0088.

D'Almeida N., « L'opinion publique : solution démocratique, problème épistémologique », *Hermès, La Revue*, 2018/1 (n° 80), p. 209-213. DOI : 10.3917/herm.080.0209. URL :

Damiani, E., di Vimercati, D. C., Paraboschi, S., Samarati, P., & Violante, F. (2002). A reputation-based approach for choosing reliable resources in peer-to-peer networks. *Proceedings of the 9th ACM Conference on Computer and Communications Security - CCS '02*, 207.

Danek A., Urbano J., Rocha A. & Oliveira E. (2010). Engaging the Dynamics of Trust in Computational Trust and Reputation Systems. *Researchgate*.

Dawn G. (2010), Organizational image construction in a fragmented online media environment. *Journal of Public Relations Research*.

Davies, G. (2015). Corporate Reputation. In *Wiley Encyclopedia of Management* (p. 1-4). American Cancer Society.

Davis, J. L. (2016). Curation : A theoretical treatment. *Information, Communication & Society*, *0*(0), 1-14.

Delavoët D-A., Dupui-Castérès A., Benatia L. (2011) « Valeur de marque et réputation : un enjeu majeur de la nouvelle gouvernance », *Sécurité et stratégie*, 2011/2 (6), p. 9-20.

Dellarocas, C., « Immunizing online reputation reporting systems against unfair ratings and discriminatory behavior », *Proceedings of the 2nd ACM conference on Electronic commerce*, pp. 150- 157, 2000.

Dellarocas, C. (2003). The Digitization of Word of Mouth : Promise and Challenges of Online Feedback Mechanisms. *Management Science*, *49*(10), 1407-1424.

Dellarocas, C., « Reputation Mechanism Design in Online Trading Environments with Pure Moral Hazard », *Information Systems Research* June, Vol. 16, pp.209-230, 2005

Derbas G., Kayssi A., Artail H., and Chehab A. A. (2004). TRUMMAR - A trust model for mobile agent systems based on reputation. *In IEEE/ACS International Conference on Pervasive Services (ICPS)*, pp. 113–120

Desmoulins, L., Alloing, C. & Mohli, V. (2018). L'influence n'est-elle que donnée(s) ? Médiations et négociations dans les agences de communication « influenceurs ». *Communication & Organisation*, 54, 29-40.

Dey, L., & Haque, S. M. (2009). Opinion mining from noisy text data. *International Journal on Document Analysis and Recognition (IJDAR)*, *12*(3), 205-226.

Dhokrat, A., Khillare, S., & Mahender, C. N. (2015). Review on Techniques and Tools used for Opinion Mining. *International Journal of Computer Applications Technology and Research*, *4*(6), 6.

Diminescu, D. (2016). Traces numériques. *Plein droit*, *110*, 3-6.

Domenget, J.-C. (2014). Formes de déconnexion volontaire et temporalités de Twitter. *Réseaux*, *186*, 77-103.

Domenget, J.-C. (2015). Usages professionnels et figures d'usager des médias socionumériques. *Revue française des sciences de l'information et de la communication*, *6*.

Domenget, J.-C. (2015). *Construire son e-réputation sur Twitter*. In C. Alcantara (Éd.), E-réputation : Regards croisés sur une notion émergente (p. 135-143). Lextenso éditions.

Dong, R., O'Mahony, M. P., Schaal, M., McCarthy, K., & Smyth, B. (2016). Combining similarity and sentiment in opinion mining for product recommendation. *Journal of Intelligent Information Systems*, *46*(2), 285-312.

Dowling, G. R. (2016). Defining and Measuring Corporate Reputations : Corporate Reputations. *European Management Review*, *13*(3), 207-223.

Dutot, V., & Castellano, S. (2015). Designing a Measurement Scale for E-Reputation. *Corporate Reputation Review*, *18*(4), 294-313.

E

Elgohary N. E., Elfetouh A. A., and Barakat S. I. (2010). Developing a reputation model for electronic markets. *International Journal of Electrical and Computer Science (IJECS-IJENS)* 10, 6

Emler, N. (2013). La réputation comme instrument social. *Communications*, 93, 85-99.

Ertzscheid, O. (2009). L'homme, un document comme les autres. *Hermès, La Revue*, *n° 53*(1), 33-40.

Ertzscheid, O (2013). Qu'est-ce que l'identité numérique ? Enjeux, outils, méthodologies. Nouvelle édition [en ligne]. Marseille : *OpenEdition Press*.

Esfandiari, B., & Chandrasekharan, S. (2001). On how agents make friends : Mechanisms for trust acquisition. *Proceedings of the fourth workshop on deception, fraud and trust in agent societies*.

F

Fang, H. Zhang J., and Thalmann N. M. (2013). A trust model stemmed from the diffusion theory for opinion evaluation. *In Proceedings of the 12th International Joint Conference on Autonomous Agents and MultiAgent Systems (AAMAS'2013)*, St. Paul, MN, pp. 805–812.

Fang, H. Zhang J., and Thalmann N. M. (2012). Sarc: Subjectivity alignment for reputation computation. *In Proceedings of the 11th International Joint Conference on Autonomous Agents and MultiAgent Systems*. 1365–1366

Finin, T., & Joshi, A. (2002). Agents, trust, and information access on the semantic web. *ACM SIGMOD Record*, *31*(4), 30.

Fombrun, C. J., Gardberg, N. A., & Sever, J. M. (2000). The Reputation QuotientSM : A multi-stakeholder measure of corporate reputation. *Journal of Brand Management*, *7*(4), 241-255.

Fombrun, C., & van Riel, C. (1997). The Reputational Landscape. *Corporate Reputation Review*, *1*(2), 5-13.

Fombrun, C.J., & Van Riel C. B.M. (2003). Fame and fortune: How Successful Companies Build Winning Reputations, *Financial Times/Prentice Hall*.

Fudenberg, D., Levine, D. K., Celentani, M., & Pessendorfer, W. (1996). Maintaining a Reputation Against a Long-Lived Opponent. *Econometrica*,.

Fueyo C., (2015). *E-réputation corporate : influence de la voix digitale des employés via l'e-contenu de leur « Profil » sur les réseaux sociaux professionnels ; application au secteur automobile*. Thèse de Doctorat, Université Toulouse 1 Capitole

Fullam K., Klos T. B., Muller G., Sabater J., Scholosser A., Topol Z., Barber K. S., Rosenshein J. S., Vercouter L. , and Voss M. (2005). A specification of the agent and reputation trust (art) testbed: Experimentation and competition for trust in agent societies. *In Proceedings of the 4th International Joint Conference on Autonomous Agents and MultiAgent Systems (AAMAS'05)*, pp. 512–518. ACM Press, July 2005

G

Galinon-Mélénec B. (2011), L'Homme trace, Perspectives anthropologiques des traces contemporaines, Paris, *CNRS Éditions*, 2011a., 2011b, p. 171-189, 2011c, p. 191-212.

Galinon-Mélénec B., Zlitni S., (2013). L'Homme-trace, producteur de traces numériques. *CNRS éditions*, pp.7-19, 2013, CNRS Alpha, 978-2-271-07239-9.

Gambetta, D. (2000). *Can we trust trust ? In Trust* : *Making and breaking cooperative relations*, Basil Blackwell, New York, pp. 213–237

Garnier, A., Guérin, G., Deuff, O. L., Deschamps, C., Henrotte, G., Blas, F., Alloing, C., & Delcroix, É. (2012). À l'échelle des organisations. *Documentaliste-Sciences de l'Information*, *49*(1), 46-59.

Gaultier-Gaillard, S., & Morlon, X. (2011). Le risque d'atteinte à la réputation : Le cas de Buffalo Grill. *Sécurité et stratégie*, *6*(2), 29-37.

Gaultier-Gaillard, S., & Pratlong, F. (2012). Le risque de réputation : Le cas du secteur bancaire. *Management & Avenir*, *n° 48*(8), 272-288.

Georges, F. (2009). Représentation de soi et identité numérique. *Réseaux*, *n° 154*(2), 165-193.

Georges, F. (2010). Pratiques informationnelles et identités numériques. *Études de communication. langages, information, médiations*, *35*, 105-120.

Georges, F. (2011). L'identité numérique sous emprise culturelle. *Les Cahiers du numérique*, *Vol. 7*(1), 31-48.

Georges, F. (2013). Le spiritisme en ligne. *Les Cahiers du numérique*, *Vol. 9*(3), 211-240.

Georgy, C. (2017). *Visibilité numérique et recrutement. Une sociologie de l'évaluation des compétences sur Internet*. Sociologie. Université Paris-Saclay. Français.

Giardini, F., Conte, R. & Paolucci, M.. (2017). Reputation for Complex Societies. 10.1007/978-3-319-66948-9_18.

Gomez Marmol, Felix & Martinez Perez, Gregorio. (2011). Trust and Reputation Models Comparison. *Internet Research*.

Gomez-Mejia, G., Nicey, J., & Vaezi, S. (2016). Questionner les pouvoirs éditoriaux de Google. *Communication & langages*, *188*, 23-43.

Gong, L., & Wang, H. (2018). When Sentiment Analysis Meets Social Network : A Holistic User Behavior Modeling in Opinionated Data. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining - KDD '18*, 1455-1464.

Granatyr J., Botelho V., Lessing O. R., Scalabrin E. E., Barthes J-P., and Enembreck F. (2015). Trust and reputation models for multiagent systems. *ACM Comput*. Surv. 48, 2,

Grandison, T., & Sloman, M. (2000). A survey of trust in internet applications. *IEEE Communications Surveys & Tutorials*, *3*(4), 2-16.

Granjon, F. (2011). De quelques pathologies sociales de l'individualité numérique. *Réseaux*, *n° 167*(3), 75-103.

Griffiths N. and Luck M. (2003). Coalition formation through motivation and trust. *In Proceedings of the 2nd International Conference on Autonomous Agents and MultiAgent Systems (AAMAS'03)*. 17–24.

Griffiths N. (2005). Task delegation using experience-based multi-dimensional trust. *In Proceedings of the 4th International Joint Conference on Autonomous Agents and MultiAgent Systems (AAMAS'05)*, ACM Press, pp. 489–496.

Grignon, T. (2016). Quand Google fait école. Une prétention pédagogique en question. *Communication & langages*, *188*, 123-139.

Grishchenko V. S. (2004). A fuzzy model for context-dependent reputation. *In Proceedings of Trust, Security and Reputation Workshop at ISWC 2004*, Hiroshima, Japan.

Guha, S., Joshi, A., & Varma, V. (2015). SIEL : Aspect Based Sentiment Analysis in Reviews. *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, 759-766.

Gui, J., & Deng, X. (2008). An Adaptive Topology-Based Reputation Model for Unstructured P2P Networks. 1505-1510.

Guilain, Y. (2015). Le « Seeking Mantra » au service de l'e-réputation. *I2D – Information, données & documents*, *Volume 52*(2), 46-48.

Gurini, D. F., Gasparetti, F., Micarelli, A., & Sansonetti, G. (2013). A Sentiment-Based Approach to Twitter User Recommendation. *CEUR Workshop Proceedings*.

Gutowska A. and Buckley K. (2008). Computing reputation metric in multi-agent e-commerce reputation system. *In IEEE International Conference on Distributed Computing Systems Workshops*, 255–260. EEE, Los Alamitos

H

Haghpanah Y. & DesJardins M. (2012). *PRep: A Probabilistic Reputation Model for Biased Societies.* 315-322. *In Proceedings of the 11th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS'12)*. 315–322

Hamdi, S. (2016). *Computational models of trust and reputation in online social networks. Modeling and Simulation*. Université Paris-Saclay, Université de Tunis. Faculté des sciences de Tunis. English.

Hendler, J., Shadbolt, N., Hall, W., Berners-Lee, T., & Weitzner, D. (2008). Web science : An interdisciplinary approach to understanding the web. *Communications of the ACM*, *51*(7), 60.

Hossein Shad Manaman, S. J. (2016). Online reputation measurement of companies based on user-generated content in online social networks. *Computers in Human Behavior*, *54*, 94-100.

Hourlay, L., Chalon, P. X., Rignanese, M., & Briat, G. (2014). Comment mettre en place une veille d'e-réputation en utilisant des outils gratuits ? *Cahiers de la documentation / Bladen voor documentatie*. 2013(3):18-2

Houser, D., & Wooders, J. (2006). Reputation in Auctions : Theory, and Evidence from eBay. *Journal of Economics Management Strategy*, *15*(2), 353-369.

Hu, M., & Liu, B. (2004). Mining Opinion Features in Customer Reviews. *Proceedings of AAAI*.

Huang H., Zhu G. , and Jin A. (2008). Revisiting trust and reputation in multi-agent systems. *In ISECS International Colloquium on Computing, Communication, Control, and Management*, Guangzhou, China, pp. 424–429

Hugo B. & Telmo M. (2016). Study on computational trust and reputation models. In *High Performance Network Studies (pp.58-66)*. Consulté 22 janvier 2017

Hung, Y.-H., Huang, T.-L., Hsieh, J.-C., Tsuei, H.-J., Cheng, C.-C., & Tzeng, G.-H. (2012). Online reputation management for improving marketing by using a hybrid MCDM model. *Knowledge-Based Systems*, *35*, 87-93.

Hussain F. K., Chang E. E., and Hussain O. (2008). A robust methodology for prediction of trust and reputation values. *In Proceedings of the ACM Conference on Computer and Communications Security*. 97–108.

Huynh, T. D., Jennings, N. R., & Shadbolt, N. R. (2004). FIRE: An integrated trust and reputation model for open multi-agent systems. *In Proceedings of the 16th European Conference on Artificial Intelligence (ECAI)*. 18–22.

Huynh, T. D., Jennings, N. R., & Shadbolt, N. R. (2006). Certified reputation : How an agent can trust a stranger. *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems - AAMAS '06*, 1217.

Huynh, T. D., Jennings, N. R., Shadbolt, N. R. (2006). An integrated trust and reputation model for open multi-agent systems. *Autonomous Agents and Multi-Agent Systems.* 13(2), 119-154

J

Jazaieri, H., Allison, M., Campos, B., Young, R., & Keltner, D. (2018). Content, structure, and dynamics of personal reputation: The role of trust and status potential within social networks. *Group Processes & Intergroup Relations*. 22. 136843021880605. 10.1177/1368430218806056.

Jha, V., Ramu, S., Shenoy, P. D., & Venugopal, K. R. (2017). Reputation Systems : Evaluating Reputation Among All Good Sellers. *Data-Enabled Discovery and Applications*, *1*(1).

Jiang, L., Ding, L., Liu, J., & Chen, J. (2012). Reputation rating modeling for open environment lack of communication by using online social cognition. *Journal of Network and Computer Applications*, *35*(3), 881-891.

Jones, A. J. (2002). On the concept of trust. *Decision Support Systems*, *33*(3), 225-232.

Jøsang, A., Ismail, R. (2002). The beta reputation system. *In Proceedings of the 15th Bled Electronic Commerce Conference*, June 2002.

Josang A. (1996). The right type of trust for distributed systems. *In Proceeding of the 1996 Workshop on New Security Paradigms*, Lake Arrowhead, USA, 16–19 September 1996, pp. 119–131.

Jøsang, A., Ismail, R., & Boyd, C. (2007). A survey of trust and reputation systems for online service provision. *Decision Support Systems*, *43*(2), 618-644.

Jurca R. and Faltings B. (2003). An incentive compatible reputation mechanism. *In Proceedings of the 6th International Workshop on Deception Fraud and Trust in Agent Societies (at AAMAS'03)*, ACM

Jurek, A., Mulvenna, M. D., & Bi, Y. (2015). Improved lexicon-based sentiment analysis for social media analytics. *Security Informatics*, *4*(1).

K

Kalepu, S., Krishnaswamy, S., & Seng Wai Loke. (2004). Reputation = f(user ranking, compliance, verity). 200-207.

Kamvar, S. D., Schlosser, M. T., & Garcia-Molina, H. (2003). The Eigentrust algorithm for reputation management in P2P networks. *Proceedings of the Twelfth International Conference on World Wide Web - WWW '03*, 640.

Karlins, M., & Abelson, H. I. (1970). Persuasion : How opinions and attitudes are changed. (2nd ed.). *Springer*

Kasavana, M.L., Nusair, K. and Teodosic, K. (2010), "Online social networking: redefining the human web", *Journal of Hospitality and Tourism Technology*, Vol. 1 No. 1, pp. 68-82.

Katz E., & Lazarsfeld P.F. (1955). Personal influence, *The Free Press*.

Keung S. N. L. C. and Griffiths N. (2008). Using recency and relevance to assess trust and reputation. *In Proceedings of AISB 2008 Symposium on Behaviour Regulation in Multi-Agent Systems*. The Society for the Study of Artificial Intelligence and Simulation of Behaviour 4, pp. 13–18.

Khatib, K. C., Kamble, T. D., Chendake, B. R., & Sonavane, G. N. (2016). Social Media Data Mining For Sentiment Analysis. *03*(04). *International Research Journal of Engineering and Technology (IRJET)*.

Khosravifar B., Bentahar J., Gomrokchi M., and Alam R. (2012). CRM: An efficient trust and reputation model for agent computing. *Knowledge-Based Systems*, in press.

Kotras, B. (2015a). L'opinion autorisée. *Revue d'anthropologie des connaissances*, *Vol. 9, n° 3*(3), 311-329.

Kotras, B. (2015b). Du bruit au signal. *Terrains travaux*, *N° 26*(1), 127-145.

Kotras, B. (2018). Le tout plutôt que la partie: Big data et pluralité des mesures de l'opinion sur le web. *Revue française de sociologie*, 59, 451-474.

Koutrouli, E., & Tsalgatidou, A. (2015). Reputation systems evaluation survey. *ACM Comput*. Surv. 48, 3, Article 35 (December 2015), 28 pages.

Koutrouli, E., Daskalakis, C., & Tsalgatidou, A. (2018). Finding Topic-Specific Trends and Influential Users in Social Networks. In L. Soldatova, J. Vanschoren, G. Papadopoulos, & M. Ceci (Éds.), *Discovery Science* (Vol. 11198, p. 405-420). Springer International Publishing.

Kugler, M. (2013). Recension de Proulx S., Millette M., et Heaton L., (dir.) (2011), Médias sociaux, enjeux pour la communication. *Communication. Information médias théories pratiques*, *Vol. 31/2*.

Kwei-Jay Lin, Haiyin Lu, Tao Yu, & Chia-en Tai. (2005). A Reputation and Trust Management Broker Framework for Web Applications. 262-269. *IEEE International Conference on e-Technology, e-Commerce and e-Service*, pp. 262-269.

L

Lakel, A., Massit-Folléa, F., & Robert, P. (2009). Imaginaire(s) des technologies d'information et de communication. *Les Editions de la MSH*

Lakel, A., & Le Deuff, O. (2017). « *À* quoi peut bien servir l'analyse du web ? », *Les Cahiers du numérique*, 13 (3), 39-62.

Lakel A., (2019). « Prises de positions et influences sur le web : le cas de l'information de santé », *Revue française des sciences de l'information et de la communication.*

Lakel, A. (2021). My web intelligence : un outil pour l'analyse du web et des réseaux. *I2D - Information, données & documents*, 1, 96-103.

Landivar, D., Monnin, A., & Ramillien, E. (2015). Cartographier l'ontologie d'un territoire sur le web. Netcom. *Réseaux, communication et territoires*, *29-3/4*, 297-324.

Lardellier, P., & Bryon-Portet, C. (2010). « Ego 2.0 ». *Les Cahiers du numérique*, *Vol. 6*(1), 13-34.

Lark, J. (2017). *Construction semi-automatique de ressources pour la fouille d'opinion. Traitement du texte et du document*. Université de Nantes, Faculté des sciences et des techniques, Français

Le Moënne, C. (2006). Quelques remarques sur la portée et les limites des modèles de communication organisationnelle. *Communication et organisation*, *30*, 48-76.

Le Deuff, O. L. (2011). Contrôle des métadonnées et contrôle de soi. *Études de communication. langages, information, médiations*, *36*, 23-38.

Le Deuff O. L,. (2014). Identité numérique. Vers une formation aux écritures et au souci de soi ? *Hermès Science Publication*. Enseignement, préservation et diffusion des identités numériques, lavoisier, 978-2746239173.

Le Deuff O., L. (2010). La formation à la présence numérique. *Documentaliste - Sciences de l'Information*, ADBS, 2010, 47 (01), p.42-44.

Le Deuff, O. (2014). Réseaux sociaux, entre médias et médiations, des espaces à méditer plutôt qu'à médire. *Communication & Organisation*, *n° 43*(1), 5-12.

Le Deuff, O. (2009) Le Ka documentarisé et la culture de l'information.. Actes de la deuxième conférence Document numérique et Société, Feb 2009, France. pp.431-444.

Leitzelman, M. (2010). La veille 2.0. *Les Cahiers du numérique*, *Vol. 6*(1), 119-133.

Leteinturier, C. (2012). Recension de Stenger T. et Coutant T. (dir.) (2011), « Ces réseaux numériques dits sociaux ». *Communication. Information médias théories pratiques*, *Vol. 30/2*.

Li, J., & Hovy, E. (2017). *Reflections on Sentiment/Opinion Analysis*. In E. Cambria, D. Das, S. Bandyopadhyay, & A. Feraco (Éds.), A Practical Guide to Sentiment Analysis (Vol. 5, p. 41-59). Springer International Publishing.

Li Xiong & Ling Liu. (2004). PeerTrust : Supporting Reputation-Based Trust for Peer-to-Peer Electronic Communities. *IEEE Transactions on Knowledge and Data Engineering*, *16*(07), 843-857.

Li Z., Fan Y., Jiang B., Lei T., Liu W. (2018). A survey on sentiment analysis and opinion mining for social multimedia, *Springer Science+Business Media, LLC, part of Springer Nature 2018*.

Lim C. K. S., Keung C. & Griffiths N. (2008). Using Recency and Relevance to Assess Trust and Reputation.

Lim C. K. S., Keung C. & Griffiths N. (2008). Towards Improved Partner Selection Using Recommendations and Trust.

Liu, B., Hu M., (2006). Opinion Extraction and Summarization on the Web Mining, *American Association for Artificial Intelligence*

Liu, B. (2010). Sentiment Analysis and Subjectivity, *Handbook of Natural Language Processing*, Second Edition, (editors: N. Indurkhya and F. J. Damerau).

Liu, B(2012). *Sentiment Analysis and Opinion Mining*, Morgan & Claypool Publishers.

Liu, B. (2015). Sentiment Analysis : Mining Opinions, Sentiments, and Emotions. *Cambridge University Press*.

Liu, B., & Zhang, L.J. (2012). A Survey of Opinion Mining and Sentiment Analysis. Mining Text Data.

Liu, L., & Munro, M. (2012). Systematic analysis of centralized online reputation systems. *Decision Support Systems*, *52*(2), 438-449.

Liu S., Yu H., Miao C., and Kot A. C. (2013a). A fuzzy logic based reputation model against unfair ratings. *In Proceedings of the 2013 International Conference on Autonomous Agents and Multi-agent Systems*. 821–828

Liu X., Datta A., and Rzadca K. (2013b). Trust beyond reputation: A computational trust model based on stereotypes. *Journal of Electronic Commerce Research and Applications* 12, 24–39.

Liu S., Zhang J., Miao C., Theng Y. L., and AlexK. C. (2011). iCLUB: An integrated clustering-based approach to improve the robustness of reputation systems (extended abstract). *In Proceedings of the Tenth International Joint Conference on Autonomous Agents and Multiagent Systems*. 1151–1152.

M

Mahon, J., & Wartick, S. L. (2012). Corporate social performance profiling : Using multiple stakeholder perceptions to assess a corporate reputation: Assessing corporate reputation. *Journal of Public Affairs*, *12*(1), 12-28.

Maimon, O., & Rokach, L. (2005). *Introduction to Knowledge Discovery in Databases*. In O. Maimon & L. Rokach (2005), Data Mining and Knowledge Discovery Handbook (p. 1-17). Springer-Verlag.

Malik, Z., & Bouguettaya, A. (2009a). Reputation Bootstrapping for Trust Establishment among Web Services. *IEEE Internet Computing*, *13*(1), 40-47.

Malik, Z., & Bouguettaya, A. (2009b). RATEWeb : Reputation Assessment for Trust Establishment among Web services. *The VLDB Journal*, *18*(4), 885-911.

Marrakchi, M. E., Bensaid, H., & Bellafkih, M. (2017). E-Reputation Prediction Model in Online Social Networks. National Institute of Posts and Telecommunications, STRS Lab, Rabat, Morocco, *International Journal of Intelligent Systems and Applications*, *9*(11), 17-25.

Mark Bunting & Roy Lipski. (2001). Drowned out? Rethinking corporate reputation management for the Internet. *Journal of Communication Management*, *5*(2), 170-178.

Marsh, S. P. (1994). Formalising trust as a computational concept. PhD thesis, University of Stirling, *Department of Computer Science and Mathematics.*

Marsh, S. P. & Dibben, M. (2005). The Role of Trust in Information Science and Technology.

Mathias, P. (2009). Note introductive aux identités numériques. *Cités*, *n° 39*(3), 59-66.

Maximilien, E. M., & Singh, M. P. (2002). Conceptual model of web service reputation. *ACM SIGMOD Record*, *31*(4), 36.

McDonald, C. G., & Slawson, V. C. (2002). Reputation in an Internet Auction Market. *Economic Inquiry*, *40*(4), 633-650.

Mcknight, D. H., Carter, M., Thatcher, J. B., & Clay, P. F. (2011). Trust in a specific technology : An investigation of its components and measures. *ACM Transactions on Management Information Systems (TMIS)*, *2*(2), 12.

McKnight, D. H., & Chervany, N. L. (2001). Trust and distrust definitions : One bite at a time. In *Trust in Cyber-societies* (p. 27-54). Springer.

McKnight, D. H., & Chervany, N. L. (2006). Reflections on an initial trust-building model. *Handbook of trust research*, 29-51.

Medić, A. (2012). Survey of computer trust and reputation models–the literature overview. *International journal of information and communication technology research*, *2*(3).

Meguebli, Y. (2015). Leveraging User-Generated Content for Enhancing and Personalizing News Recommendation.

Melnik, M. I., & Alm, J. (2003). Does a Seller's eCommerce Reputation Matter? Evidence from eBay Auctions. *The Journal of Industrial Economics*, *50*(3), 337-349.

Melnikov, A., Lee, J., Rivera, V., Mazzara, M., & Longo, L. (2018). Towards Dynamic Interaction-Based Reputation Models. *(:unav)*.

Mercklé, P. (2016). *Les origines de l'analyse des réseaux sociaux*. 20. *Mercklé Pierre, « II.* L'analyse des réseaux sociaux, une méthodologie quantitative *»,* dans : Pierre Mercklé éd., La sociologie des réseaux sociaux. Paris, La Découverte, « Repères », p. 21-36.

Merklé p., (2004), *Sociologie des réseaux sociaux*, Paris, Éd. La Découverte, coll. Repères, 128 p

Mericskay, B. (2016). Cartographie à l'ère du web : Réflexion sémantique et sémiologique sur les POI.

Mericskay, B., & Roche, S. (2011). Cartographie 2.0 : Le grand public, producteur de contenus et de savoirs géographiques avec le web 2.0. *Cybergeo : European Journal of Geography*.

Merzeau, L. (2009a). Du signe à la trace : L'information sur mesure. *Hermès, La Revue*, *n° 53*(1), 21-29.

Merzeau, L. (2009b). Les données post mortem. *Hermès, La Revue*, *n° 53*(1), 30-31.

Merzeau, L. (2009c). De la surveillance à la veille. *Cités*, *n° 39*(3), 67-80.

Merzeau, L. (2009d). Présence numérique : Les médiations de l'identité. *Les Enjeux de l'information et de la communication*, *Volume 2009*(1), 79-91.

Merzeau L. (2010). La présence plutôt que l'identité. *Documentaliste - Sciences de l'Information*, ADBS, 2010, 47 (1), pp.32-33.

Merzeau L.(2013). *Traces numériques et recrutement : du symptôme au cheminement*. CNRS éditions. Traces numériques: de la production à l'interprétation, *CNRS éditions*, pp.35-53. halshs- 01071357.

Merzeau, L. (2012). La médiation identitaire. *Revue française des sciences de l'information et de la communication*, *1*.

Merzeau L. (2013). L'intelligence des traces. *Intellectica - La revue de l'Association pour la Recherche sur les sciences de la Cognition (ARCo), Association pour la Recherche sur la Cognition,* 1 (59), p.115-135. halshs-01071211.

Mesguich, V., Diallo, A., Jdey, A., Bergeret, C., Dumas, S., Séménéri, M., Asselin, C., Besson, B., Sala, J., & Remize, M. (2008). Où va la veille ? *Documentaliste-Sciences de l'Information*, *Vol. 45*(4), 58-69.

Millette, L. (2013). Information et médias sociaux, les défis de la qualité. *Éthique publique. Revue internationale d'éthique sociétale et gouvernementale*, *15*(1).

Millette, M. (2013). Pratiques transplateformes et convergence dans les usages des médias sociaux. *Communication et organisation*, *43*, 47-58.

Mitic, P. (2017). Standardised Reputation Measurement. arXiv:1705.09955 [q-fin].

Mittal, N (2016). *Data, learning and privacy in recommendation systems. Machine Learning* [cs.LG]. Université Rennes 1, 2016. English. NNT : 2016REN1S084 . tel-01477420.

Moinet, N., Frion, P., & Samier, H. (2007). Le phénomène de googlelisation dans les recherches d'informations sur internet.

Mokhtari E., Noorian Z., Ladani B. T., and Nematbakhsh M. A. (2011). A context-aware reputation-based model of trust for open multi-agent environments. *In Advances in Artificial Intelligence*, pp. 301–312. Springer Berlin Heidelberg

Mondoux, A. (2011). Identité numérique et surveillance. *Les Cahiers du numérique*, *Vol. 7*(1), 49-59.

Mondoux, A. (2012). À propos du social dans les médias sociaux. *Terminal. Technologie de l'information, culture & société*, *111*, 69-79.

Morzy, M., Wojciechowski, M., & Zakrzewicz, M. (2005). *Intelligent Reputation Assessment for Participants of Web-Based Customer-to-Customer Auctions*. In P. S. Szczepaniak, J. Kacprzyk, & A. Niewiadomski (Éds.), Advances in Web Intelligence (Vol. 3528, p. 320-326). Springer Berlin Heidelberg.

Mui, L., Halberstadt, A., & Mohtashemi, M. (2003). *Evaluating Reputation in Multi-agents Systems*. In R. Falcone, S. Barber, L. Korba, & M. Singh (Éds.), Trust, Reputation, and Security : Theories and Practice (p. 123-137). Springer Berlin Heidelberg.

Mui, L., Mohtashemi, M., & Halberstadt, A. (2002). A computational model of trust and reputation. *Proceedings of the 35th Annual Hawaii International Conference on System Sciences*, 2431-2439.

Mui, L., Mohtashemi, M., & Halberstadt, A. (2002). Notions of reputation in multi-agents systems: A review. *Proceedings of the First International Joint Conference on Autonomous Agents and Multi-Agent System (AAMAS 2002)*, Bologna, Italy, 15–19 July 2002, pp. 280–287

Muller G. and Vercounter L. (2005). Decentralized monitoring of agent communications with a reputation model (LIAR). *In Trusting Agents for trusting Electronic Societies, Lecture Notes in Computer Sciense*, 144–161.

N

Nkosi N. T., Adigun M. O., and Emuoyibofarhe J. O. (2007). Agent-to-agent reputation-based trust management. *IADIS International Conference Applied Computing*, pp. 143–150

O

Origgi, G. (2013). Un certain regard. *Communications*, *n° 93*(2), 101-120.

Osman, N., Provetti, A., Riggi, V. & Sierra, C.. (2015). MORE: Merged Opinions Reputation Model. 10.1007/978-3-319-17130-2_5.

Ouakrat, A., & Mésangeau, J. (2016). Resocialiser les traces d'activités numériques : Une proposition qualitative pour les SIC. *Revue française des sciences de l'information et de la communication*.

P

Padovan B., Sackmann S., Eymann T., and Pippow I. (2001). A prototype for an agent-based secure electronic marketplace including reputation-tracking mechanisms. *In Proceedings of the 34th Annual Hawaii International Conference on System Sciences (HICSS-34*), 2235–2244, Maui, Hawaii – Track 7, January 3–6 2001

Pan, Y. (2011). Eliminating the cyber 'lemons' problem with the e-reputation in e-commerce market : Theoretical model and practice. *International Journal of Networking and Virtual Organisations*, *8*(3-4), 182-191.

Papadopoulos, S., Kompatsiaris, Y., Vakali, A., & Spyridonos, P. (2012). Community detection in Social Media : Performance and application considerations. *Data Mining and Knowledge Discovery*, *24*(3), 515-554.

Pappas, N., & Popescu-Belis, A. (2013). Sentiment analysis of user comments for one-class collaborative filtering over ted talks. *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '13*, 773.

Paquerot, M., Queffelec, A., Sueur, I. & Biot-Paquerot, G. (2011). L'e-réputation ou le renforcement de la gouvernance par le marché de l'hôtellerie ? *Management & Avenir*, 45, 280-296. Consulté 1 février 2016

Perea, F. (2010). L'identité numérique : De la cité à l'écran. Quelques aspects de la représentation de soi dans l'espace numérique. *Les Enjeux de l'information et de la communication*, *Volume 2010*(1), 144-159.

Perriault, J. (2009). Traces numériques personnelles, incertitude et lien social. *Hermès, La Revue*, *n° 53*(1), 13-20.

Pierre, J. (2011). Génétique de l'identité numérique. *Les Cahiers du numérique*, *Vol. 7*(1), 15-29.

Pierre J. (2014). De l'identité numérique à l'individu transmédiatique.

Ping Xu, Ji Gao, & Hang Guo. (2005). Rating Reputation : A Necessary Consideration in Reputation Mechanism. 182-187.

Pinyol, I., & Sabater-Mir, J. (2013). Computational trust and reputation models for open multi-agent systems : A review. *Artificial Intelligence Review*, *40*(1), 1-25. DOI: 10.1007/s10462-011-9277-z

Pinyol, I., & Sabater-Mir, J. (2007). Arguing about reputation : The lrep language. *International Workshop on Engineering Societies in the Agents World*, 284-299.

Pinyol, I., Sabater-Mir, J., & Cuní, G. (2007). How to talk about reputation using a common ontology : From definition to implementation. *Proceedings of the Ninth Workshop on Trust in Agent Societies. Hawaii, USA*, 90-101.

Pinyol, I., Sabater-Mir, J., Dellunde P., and Paolucci M. (2012). Reputation-based decisions for logic-based cognitive agents. *Autonomous Agents and Multi-Agent Systems* 24, 1, 175–216

Pirolli, F. (2011). Pratiques d'indexation sociale et démarches de veille informationnelle. *Études de communication. langages, information, médiations*, *36*, 53-66.

Plantin, J.-C. (2013). Qu'y a-t-il à côté d'un graphe de sites web ? *Communication et organisation*, *43*, 59-70.

Player, C. & Griffiths, N. (2017). Bootstrapping Trust and Stereotypes with Tags.

Podnar, K., & Golob, U. (2017). The Quest for the Corporate Reputation Definition : Lessons from the Interconnection Model of Identity, Image, and Reputation. *Corporate Reputation Review*, *20*(3-4), 186-192.

Ponzi, L. J., Fombrun, C. J., & Gardberg, N. A. (2011). RepTrak™ Pulse : Conceptualizing and Validating a Short-Form Measure of Corporate Reputation. *Corporate Reputation Review*, *14*(1), 15-35.

Portmann, E., Meier, A., Cudré-Mauroux, P., & Pedrycz, W. (2015). FORA – A fuzzy set based framework for online reputation management. *Fuzzy Sets and Systems*, *269*, 90-114.

R

Ramchurn S. D., Jennings N. R., Sierra C., and Godo. L. (2004a). Devising a trust model for multi-agent interactions using confidence and reputation. *Applied Artificial Intelligence*, 18, 9–10, 833–852.

Rchika, T. (2018). Les tweets comme traces numériques. Anticiper les enjeux réputationnels des publications sur Twitter. *Revue française des sciences de l'information et de la communication*, *13*.

Regan, K., & Cohen, R. (2005). A model of indirect reputation assessment for adaptive buying agents in electronic markets.

Regan K., Poupart P., and Cohen R. (2006). Bayesian reputation modeling in e-marketplaces sensitive to subjectivity, deception and change. *In Proceedings of the Conference on Artificial Intelligence* (AAAI'06). 1206

Renaud, C. (2016). Topogram, outil web de cartographie et de visualisation spatio-temporelle des réseaux.

Resnick, P., Zeckhauser, R., Friedman, E., & Kuwabara, K. (2000). Reputation Systems. *Commun. ACM*. 43. 45-48.

Reuber, A. R., & Fischer, E. (2009). Signalling reputation in international online markets. *Strategic Entrepreneurship Journal*, *3*(4), 369-386.

Rieder, B. (2011). Pratiques informationnelles et analyse des traces numériques : De la représentation à l'intervention. *Études de communication*, *35*, 91-104.

Rossi, M. (2017). Graph Mining for Influence Maximization in Social Networks. *Artificial Intelligence [cs.AI]. Université Paris-Saclay. English.*

S

Sabater, J., & Sierra, C. (2001a). Social regret, a reputation model based on social relations. *ACM SIGecom Exchanges*, *3*(1), 44-56.

Sabater, J., & Sierra, C. (2001b). REGRET: reputation in gregarious societies. *In AGENTS '01: Proceedings of the fifth international conference on Autonomous agents, pages 194–195*.

Sabater, J., & Sierra, C. (2002). Reputation and social network analysis in multi-agent systems. *Proceedings of the first international joint conference on Autonomous agents and multiagent systems: Part 1*, 475-482. Bologna, Italy.

Sabater, J., & Sierra, C. (2005). Review on Computational Trust and Reputation Models. *Artificial Intelligence Review*, *24*(1), 33-60.

Sabater J., Paolucci M., and Conte R. M. (2006). Repage: REPutation and ImAGE among limited autonomous partners. *Journal of Artificial Societies and Social Simulation* 9, 2.

Sabater-Mir, J., & Paolucci, M. (2007). On representation and aggregation of social evaluations in computational trust and reputation models. *International Journal of Approximate Reasoning*, *46*(3), 458-483.

Salehi-Abari A. and White T. (2012). DART: A distributed analysis of reputation and trust framework. *Computational Intelligence*, 28, 642–682. DOI: 10.1111/j.1467-8640.2012.00453

Sajus, B., Cardon, D., Levrel, J., Vatant, B., Bermès, E., Oury, C., & Sussan, R. (2009). Web 2.0, et après ? *Documentaliste-Sciences de l'Information*, *46*(1), 54-66.

Sänger, J., Richthammer, C., & Pernul, G. (2015). Reusable components for online reputation systems. *Journal of Trust Management*, *2*(1), 1-21.

Sarstedt, M., Wilczynski, P., & Melewar, T. C. (2013). Measuring reputation in global markets—A comparison of reputation measures' convergent and criterion validities. *Journal of World Business*, *48*(3), 329-339.

Schafer, V., & Thierry, B. G. (2013). Qui a inventé Internet ? Une vraie « fausse question ». *Le Temps des médias*, *n° 20*(1), 223-235.

Schiffner, S., Clauß, S., & Steinbrecher, S. (2010). *Privacy and Liveliness for Reputation Systems*. In F. Martinelli & B. Preneel (Éds.), Public Key Infrastructures, Services and Applications (p. 209-224). Springer Berlin Heidelberg.

Sharma N. K., Gaur V., and Muttoo S. K. (2012). A dynamic reputation system with built-in attack resilience to safeguard buyers in e-market. *ACM SIGSOFT Software Engineering Notes* 37, 4, 1–19.

Sharma, D. M., & Baig, M. M. (2013). Sentiment Analysis on Social Networking : A Literature Review. *International Journal on Recent and Innovation Trends in Computing and Communication*, *3*(2), 6.

Sen, S. & Sajja, N.. (2002). Robustness of reputation-based trust: Boolean Case. *Proceedings of the International Conference on Autonomous Agents (AAMAS). 288–293, Bologna, Italy*.

Serrano E., Rovatsos M., and Botia J. (2012). A qualitative reputation system for multiagent systems with protocol-based communication. *In Proceedings of the 11th Conference on Autonomous Agents and Multi- agent Systems* 1, pp. 307–314

Severo M. (2012). La cartographie du Web : le lien social sur le Net. 2012. halshs-00678768.

Severo, M. & Venturini, T.. (2016). Enjeux topologiques et topographiques de la cartographie du Web: Le cas du patrimoine culturel immatériel français. *Réseaux*. n° 195.

Singh Jandail, R. R. (2014). A proposed Novel Approach for Sentiment Analysis and Opinion Mining. *International Journal of UbiComp*, *5*(2), 1-10.

Sire, G. (2016). Le pouvoir normatif de Google. Analyse de l'influence du moteur sur les pratiques des éditeurs. *Communication & langages*, *188*, 85-99.

Škorić, B., Hoogh, S. J. A. de, & Zannone, N. (2015). Flow-based reputation with uncertainty : Evidence-based subjective logic. *International Journal of Information Security*, 1-22.

Soldano, Henry & Santini, Guillaume & Bouthinon, Dominique & Lazega, Emmanuel. (2017). Hub-Authority Cores and Attributed Directed Network Mining.

Sonnek, J. D., & Weissman, J. B. (2005). A quantitative comparison of reputation systems in the grid. 8 pp.

Stenger, T. (2014). Social media and online reputation management as practice : First steps towards social CRM? In *International Journal of Technology and Human Interaction*, *10*(4), 49-64.

Stenger, T., & Coutant, A. (2013). Les sites de réseautage professionnels au service du marketing : Le cas Viadeo. Entrevue avec Olivier Fecherolle. *Communication et organisation*, *43*, 283-289.

Stenger T., & Coutant A. (2013). Médias sociaux : clarification et cartographie - Pour une approche sociotechnique ». *Décisions Marketing*, 70, avril-juin, 107-117.

Swamynathan, G., Almeroth, K. C., & Zhao, B. Y. (2010). The design of a reliable reputation system. *Electronic Commerce Research*, *10*(3-4), 239-270.

T

Tan, C., Lee, L., Tang, J., Jiang, L., Zhou, M., & Li, P. (2011). User-level sentiment analysis incorporating social networks. *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '11*, 1397.

Taylor, P., Barakat, L., Miles, S., & Griffiths, N. (2018). Reputation assessment : A review and unifying abstraction. *The Knowledge Engineering Review*, *33*.

Teacy L. T. W., Luke M., Rogers A., and Jennings R. N. (2012). An efficient and versatile approach to trust and reputation using hierarchical bayesian modeling. *Artificial Intelligence* 193, 149–185

Teacy W. T. L., Patel J., Jennings N. R., and Luck. M. (2006). TRAVOS: Trust and reputation in the context of inaccurate information sources. *Autonomous Agents and Multiagent Systems* 12, 2, 183–198.

Tian J., Li J., and Yanga L. (2009). Reputation-based multi-agent model for network resource selection. *International J. Communications, Network and System Sciences*, 2, 764–774, DOI:10.4236/ijcns.2009.28089

Tran T. and Cohen R. (2004). Improving user satisfaction in agent based electronic marketplaces by reputation modelling and adjustable product quality. *In Proceedings of the Third International Joint Conference on Autonomous Agents and Mutiagent Systems (AAMAS'04)*. 828–835, New York, USA

Trček, D. (2018). *Computational Trust and Reputation Management*. In D. Trček, Trust and Reputation Management Systems (p. 21-54). Springer International Publishing.

U

Urzica A. (2010). Distributed reputation extraction in multi-agent systems. I*n Proceedings of the 12th European Agent Systems Summer School Student Session*. 5–11

Urzica A., Mogos H. A., and Florea M. A. (2011). A reputation based negotiation model for barter transactions between software agents. *International Journal on Artificial Intelligence Tools* 20, 1001.

Utz, S., Matzat, U., & Snijders, C. (2009). On-line Reputation Systems : The Effects of Feedback Comments and Reactions on Building and Rebuilding Trust in On-line Auctions. *International Journal of Electronic Commerce*, *13*(3), 95-118.

V

Versailles, D. W., Dutot, V., & Lacalle Galvez, E. (2016). CSR communications strategies through social media and influence on e-reputation : An exploratory study. *Management Decision*, *54*(2), 363-389.

Vogiatzis G., MacGillivray I., and Chli M. (2010). A probabilistic model for trust and reputation. *In Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems.* 225–232

W

Walker, K. (2010). A Systematic Review of the Corporate Reputation Literature : Definition, Measurement, and Theory. *Corporate Reputation Review*, *12*(4), 357-387.

Wang, Y., & Vassileva, J. (2003). Trust and reputation model in peer-to-peer networks. *Proceedings Third International Conference on Peer-to-Peer Computing (P2P2003)*, Linkoping, Sweden; 1–3 September 2003, pp. 150-157.

Wartick, S. L. (2002). Measuring Corporate Reputation : Definition and Data. *Business & Society*, *41*(4), 371-392.

Watts D., Dodds P. S. (2007). Influentials, Networks, and Public Opinion Formation, *Journal Consumer Research,* 34, 441-458.

Whitby, A., Jøsang, A., & Indulska, J. (2005). Filtering Out Unfair Ratings in Bayesian Reputation Systems. *The Icfain Journal of Management Research*.

Windley, P. J., Daley, D., Cutler, B., & Tew, K. (2007). Using reputation to augment explicit authorization. *Proceedings of the 2007 ACM workshop on Digital identity management*, 72-81.

Windley, P. J., Tew, K., & Daley, D. (2007). A Framework for Building Reputation Systems. Dept. of Computer Science Brigham Young University Provo, UT.

Wooldridge M. (2009). An Introduction to MultiAgent Systems. Department of Computer Science, University of Liverpool, UK. John Wiley & Sons, LTD

X

Xiao, S., & Dong, M. (2015). Hidden semi-Markov model-based reputation management system for online to offline (O2O) e-commerce markets. *Decision Support Systems*, *77*, 87-99.

Y

Yan, S.-R., Zheng, X.-L., Wang, Y., Song, W. W., & Zhang, W.-Y. (2015). A graph-based comprehensive reputation model : Exploiting the social context of opinions to enhance trust in social commerce. *Information Sciences*, *318*, 51-72.

Ye, S., Gao, Gordon G., & Viswanathan, S. (2013). Strategic Behavior in Online Reputation Systems : Evidence from Revoking on eBay. *Social Science Research Network*. Consulté le 3 février 2016

Yoganarasimhan, H. (2013). The Value of Reputation in an Online Freelance Marketplace. *Marketing Science*, *32*(6), 860-891.

You L. (2007). *An adaptative reputation-based trust model for intelligent agents in e-market place*. Doctoral Dissertation, University of Texas at Arlington, USA

You, L., & Sikora, R. (2013). Performance of online reputation mechanisms under the influence of different types of biases. *Information Systems and E-Business Management*, *12*(3), 417-442.

Yu B. and Singh M. P. (2000). A social mechanism of reputation management in electronic communities. *In Proceedings of the 4th International Workshop on Cooperative Information Agents* IV, pages 154–165.

Yu, B., & Singh, M. P. (2002). An evidential model of distributed reputation management. *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems Part 1 - AAMAS '02*, 294.

Yu B. and Singh M. P. (2003). Detecting deception in reputation management. *In Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. ACM, 73–80

Z

Zacharia, G. (1999). Collaborative reputation mechanisms for online communities. Massachusetts Institute of Technology.

Zacharia, G., Maes, P. (2000). Trust management through reputation mechanisms. *Applied Artificial Intelligence* 14(9)

Zhang, K., Downey, D., Chen, Z., Xie, Y., Cheng, Y., Agrawal, A., Liao, W., & Choudhary, A. (2013). A probabilistic graphical model for brand reputation assessment in social networks. *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining - ASONAM '13*, 223-230.

Zhang, Y., Li, Q. & Lin Z. (2010). A Novel Reputation Computing Model. *Conference: Proceedings of the Second international conference on Trusted Systems*. Consulté 21 janvier 2017 DOI:10.1007/978-3-642-25283-9_21

Zhou, M., Dresner, M., & Windle, R. J. (2008). Online reputation systems : Design and strategic practices. *Decision Support Systems*, *44*(4), 785-797.

Zhou Q., Wang C., and Xie J. (2009). Core: A trust model for agent coalition formation. *Fifth International Conference on Natural Computation (ICNC'09)* 5, 541–545

ANNEXES

Annexe 1 : Tableau de classification des modèles théoriques de mesure de l'e-réputation

Avant de présenter les modèles théoriques issus de notre analyse de la littérature scientifique surtout anglo-saxonne, il nous semble important de spécifier certains des items utilisés. Sur les neuf (9) grandes thématiques (**D, CIT, C, DBI, MMA, MREP, SRD, TYPE** et **CAT**), vous trouverez dans le tableau n°3, un index qui reprend l'ensemble des critères.

Sur les critères spécifiques par dimension, on retrouve:

- Pour la dimension **DBI** : l'attribut ***(+S)*** est relatif à la mise en place dans la conception du modèle d'un mécanisme sécuritaire visant à filtrer les agents frauduleux;

- Pour la dimension **MMA** : sur le modèle mathématique, cinq (5) principales approches sont utilisées dans les modèles : les modèles construits autour de la ***somme ou moyenne (S / A)***, les modèles qui s'appuient sur les ***probabilités (P)***, les approches liées à la ***logique subjective (SL)***, les modèles qui utilisent la logique floue ou ***fuzzy logic (FL)*** et enfin les approches basées sur les ***flux ou flow base (FB);***

- Pour la dimension **MREP** relative aux métriques de réputation utilisées de manière concrète ou implicite, on retrouve : la ***notoriété (N), l'influence (I) ou la popularité (P), l'identité (ID) ou l'image (IM)*** et enfin la ***confiance (C)*** souvent associée au ***risque (R);***

- Pour la dimension **SRD** : sur les sources de données utilisées dans les modèles de mesure de la réputation, on distingue les informations issues de sources directes ou d***irect informations (DI)***, les données relatives aux observations directes ou ***direct observation (DO)***, les ***witness informations (WI)*** ou témoignages de tiers, les ***sociological informations (SI)*** ou ***rules (R)*** dans certains modèles, les préjugés ou ***prejudice informations (P)***, la réputation certifiée ou ***certified reputation (CR);***

- Pour la dimension **TYPE** , on identifie deux principaux types de modèles : les ***modèles de réputation (Rep)*** et les ***modèles hybrides*** de confiance et de réputation ***(H);***

- La dernière dimension **CAT** que nous avons identifiée à l'issue de notre travail d'analyse est relative à la catégorie à laquelle appartiennent les modèles. On retrouve dans cette dernière dimension, les ***modèles ad hoc centralisés (C)***, les ***modèles ad hoc décentralisés (D)***, les ***approches génériques (G),*** *les* ***modèles sommaires (SO)*** et les ***modèles sophistiqués (S)***.

MODÈLE	D	CIT	C	DBI	MMA	MREP	SRD	TYPE	CAT
Page & al. (PAGE RANK)	1998	-	Moteurs de recherche (Google)	0	FL	N	DI+WI	Rep	**D**
Zacharia & Maes (SPORAS & HISTOS)	2000	818	E commerce	0	S	-	DI, WI	Rep	**SO**
Yu & Singh	2001 /3	892	P2p	0	SL	-	DI+WI	Rep	**D**
Sabater et Sierra (REGRET)	2002	965	MAS / Général	2	FL	C	DI+WI+ SI+ P	H	**S**
Mui et al.	2002	1254	E commerce	0	P	C	DI+WI	H	**D**
Sen & Sajja	2002	218	MAS / Général	2	P	-	DI+WI+ DO	Rep	**S**
Carbo et al. (AFRAS)	2002	180	E commerce	2	FL	C	DI+WI	H	**D**
Carter & al.	2002	173	-	0	S	ID	RL+WI	Rep	**C**
Josang & Ismail (BRS)	2002	1770	E commerce	0	SL /B	-	DI+WI	Rep	**C**

MODÈLE	D	CIT	C	DBI	MMA	MREP	SRD	TYPE	CAT
Michiardi & Molva (CORE)	2002	2117	P2p mobile network (manet)	1	A	-	DI+WI	Rep	D
Padovan et al. (AVALAN-CHE)	2002	105	Places de marché	2	A	-	-	Rep	D
Damiani & al. (XREP)	2002	-	File sharing	-	S / A	C	DI+WI	H	D
Wang & Vassileva	2003	700	File sharing	0	B	C	DI,WI	H	D
Buchegger & Le Boudec (CONFI-DANT)	2003	1152	P2p mobile network (manet)	-	FL	-	-	Rep	D
Jurca & Faltings	2003	319	MAS / Général	2+ S	P	ID	DI+WI	Rep	S
Kamvar & al. (EIGEN-TRUST)	2003	5058	P2p Général	-	FB	C	-	H	S
Derbas et al. (TRUMMAR)	2004	52	MAS / Général	0	S	C	DI+WI +SI	H	G
Gupta & Somani	2004	66	P2p file sharing	2+ S	FB / A	ID	-	Rep	D

MODÈLE	D	CIT	C	DBI	MMA	MREP	SRD	TYPE	CAT
Huynh et al. (FIRE)	2004/ 20 06	851	MAS / Général	2+S	S / A	C	DI+WI +CR +RL	H	S
Tran & Cohen	2004	101	Places de marché	2	FB	C	DI	H	D
Xiong & Liu (PEER-TRUST)	2004	2443	Général	2	FB	C	-	H	S
Song & al (FUZZY-TRUST)	2005	-	P2p places de marché	-	FL	-	-	Rep	D
Muller & Vercouter (LIAR)	2005	55	Général/ C & D	2	FB	C	DI+WI +SI	H	S
Fullan & al. (ART)	2005	260	Places de marché	0	FL / S	C	DI+DO	H	G
Teacy & al. (TRAVOS)	2006	669	Général	2	P	C	DI+WI	H	S
Sabater & al. (REPAGE)	2006	213	MAS	2	FL	IM	DI+WI	Rep	S
Regan & Cohen	2005	21	Places de marché	0	FB	IM	DI+WI	Rep	D

MODÈLE	D	CIT	C	DBI	MMA	MREP	SRD	TYPE	CAT
Regan et al. (BLADE)	2006	150	Places de marché	2	P	-	DI+WI	Rep	D
Tian et al. (GROUP-REP)	2006	43	-	2	FB / A	IM / R	-	Rep	S
Zhang et al. (DIRECT)	2006	17	Places de marché	2	S / A	C	DI+WI	H	D
Chen & al. (CUTEAID)	2006	61	Q&A / SNA	0	S	I	DI+WI	Rep	D
Chen & al. (CUBOID TRUST)	2007	41	P2p file sharing	1	A	C	DI+WI	H	D
Li et al. (WSREP)	2007	20	Général (MAS)	0	P	-	DI+WI +SI	Rep	D
Tajeddine & al. (PATROL)	2007	31	Général	-	FL	I / P / C	-	H	S
Zhou & al. (POWER-TRUST)	2007	923	P2p grid/ file sharing	1	FB	C	-	H	D
Zhang & al. (RRM)	2008	15	Général	1	S / A	-	DI+WI	Rep	G
Zhao & Li (H TRUST)	2008	46	P2p grid/ file sharing	0	FL	C	-	H	D

MODÈLE	D	CIT	C	DBI	MMA	MREP	SRD	TYPE	CAT
Bertocco & ferrari	2008	4	Général (MAS)	0+S	FB / A	-	DI, WI	Rep	C
Gutowska & Buckley	2008	18	E commerce market-place	0	FL / A	-	DI+WI	Rep	D
Da Silva et al.	2009	5	Général	0	FL	-	RL	Rep	D
Tian & al. (RMNRS)	2009	1	MAS / Général	0	FL	ID	DI+WI	Rep	D
Li & Gui (TREE-TRUST)	2009	17	-	-	SL	C	-	H	C
Liu & Issarny	2009	79	P2p mobile network (manet)	1	FL	-	DI,WI	Rep	D
Kerr & cohen (TREET)	2009	52	Places de marché	1	-	C	-	H	G
Nguyen & al.	2010	124	MAS / Général (web services)	0	P (B)	C	DI+WI	H	D
Vogiatzis & al.	2010	78	E commerce	0	P	C	DI+WI	H	C
Almulla et al. (TRUST-PERT)	2011		P2p mobile network (manet)	1	FL	-	DI+WI +DO	Rep	D

MODÈLE	D	CIT	C	DBI	MMA	MREP	SRD	TYPE	CAT
Urzica & al.	2011	2	MAS	1	FB	C	DI+WI +RL	H	C
Tian & Yang (R2 TRUST)	2011	47	File sharing	2	P	C	-	H	D
Wang et al. (LAGA)	2011	145	MAS / Général	0	P	-	-	Rep	D
Qureshi & al. (M-TRUST)	2011	52	P2p mobile network	1	A	C	DI+WI	H	D
Fan & al. (DUAL EIGEN-REP)	2011	31	File sharing	1	FB	C	DI+DO	H	D
Haghpanah & Desjardins (PREP)	2012	2	MAS / Général	1	P	-	DI, WI	Rep	G
Fang & al. (SARC)	2012	24	E commerce	1	P (B)	-	-	Rep	C
Teacy & al. (HABIT)	2012	120	MAS / Général	2	P	C	DO+ DI+ WI	H	G
Khosravifar & al. (CRM)	2012	81	MAS / Général	2	P	C	DI+WI	H	D
Jung	2012	28	wiki / social media	0	SL	-	DI	Rep	D

MODÈLE	D	CIT	C	DBI	MMA	MREP	SRD	TYPE	CAT
Pinyol & al. (BDI+ REPAGE)	2012	58	MAS / Général	2	SL /P	I / IM / C	DI+WI	H	S
Serrano & al.	2012	21	Général	0	SL	-	DI+WI+ DO	Rep	D
Salehi-Abari & White (DART)	2012	18	MAS / Général	2	FL	C	DI+WI+ DO	H	S
Li & al. (TBM)	2013	39	Sites d'avis	2	FL	-	-	Rep	D
Tuna & al.	2013	3	Robotique sensor network	2+S	FL	-	DI+WI	Rep	D
Qureshi & al. (FIRE+)	2013	11	P2p mobile network (manet)	2	A	C	DI+WI	H	D
Kussul & al.	2013	21	Grid computing	3+S	FL	-	-	Rep	D
Wang & al. (ADVS)	2013	87	VoIP networks	2	-	ID / C	DI+WI	H	D
Javanmardi & al. (FR TRUST)	2014	96	MAS / Général	1	FL	C	-	H	G

MODÈLE	D	CIT	C	DBI	MMA	MREP	SRD	TYPE	CAT
Ransi and Kobti	2014	3	E commerce	0	FL	C	DI+WI	H	C
Yu & al. (ARM)	2014	26	MAS / Général	0	FL	-	-	Rep	G
Marmol & Perez (TRIP)	2012/ 14	213	Vanet	2	FL	C	DI+WI +RL	H	D
Fang & al. (BTRES)	2015	72	Wireless sensor network	2+S	P	C	DI+WI	H	D
Osman & al. (MORE)	2015	6	P2p Général	0	P	I	DI+WI	Rep	D
Lin & al. (RMTAC)	2015	19	Mobile cloud computin g	2+S	S	C	DI+WI	H	D
Yan et al.	2015	60	E commerce & social media	2+S	FL	-	DI, WI	Rep	G
Portmann & al. (FORA)	2015	30	Bookmark & twitter	0	FL	-	-	Rep	D
Abdel-Hafez et al. (BETADR)	2015	6	E commerce / sites d’avis	0	S / A	-	-	Rep	C

MODÈLE	D	CIT	C	DBI	MMA	MREP	SRD	TYPE	CAT
Majd and Balakrishnan (R-D-C)	2015	8	MAS / e commerce	0	S / A	C	-	H	D
Zhang & al.	2015	13	Supply chain/ e commerce	2+S	FB	I / P	-	Rep	D
Kravari & Bassiliades (DISARM)	2016	7	MAS / Général (+ social networks)		SL	-	DI+WI	Rep	G
Wang & al. (RPREP)	2016	29	Vanet	2+S	FL	-	DI	Rep	C
Chandran et al.	2016	20	Cloud computing	2+S	FL	C	-	H	D
Škorić et al.	2016	12	P2p file sharing	-	FB	-	-	Rep	D
Lu & al.	2016	18	P2p file sharing	2+S	FL	-	-	Rep	D
Acampora & al.	2016	26	E commerce	-	FL	-	-	Rep	C
Li & al (DDOA)	2016	54	IOT (smart grid/ electricity)	2+S	P	-	-	Rep	C

* Dernière mise à jour du nombre de citations le 20 octobre 2020.
* Les (-) signifient que les caractéristiques n'ont pas clairement été spécifiées dans les articles scientifiques des modèles

Annexe 2 : Le marché de l'e-réputation

Les acteurs qui composent le marché de l'e-réputation peuvent être regroupés selon **6 principales catégories** : les *pure players*, les agences généralistes, les agences ad hoc, les éditeurs de solutions logicielles, les applications gratuites et les tiers de confiance.

AGENCE	CATÉGORIE	OFFRE	PAGE CORPORATE
ACTI	Généraliste	Gestion e-réputation	https://www.acti.fr
ACTIVIS	Généraliste (Webmarketing)	Veille, Référencement, Content Marketing	https://www.activis.net
ACTULLIGENCE	Pure player	Veille (e-réputation, intelligence économique, social listening)	www.actulligence.com
ADCOM	Généraliste	Social listening & gestion e-réputation	https://www.adcom.fr
ADIT	Généraliste	Intelligence stratégique (e-réputation)	https://www.adit.fr/fr/
ADVERIS	Pure player	Gestion e-réputation	https://www.adveris.fr
ADGENCY 360	Généraliste (Communication digitale)	Gestion de crise en e-réputation	https://www.adgency360.com
AD-WORKS	Tiers de confiance	Collecte & gestion avis clients	https://www.ad-works.fr
AGENCE KALI	Généraliste	Audit éco-systèmes digitaux, stratégie social media	http://www.agencekali.fr/
AGENCE 1M30	Généraliste (Webmarketing)	Audit de la présence digitale	https://www.1min30.com/
AGENCE 148	Généraliste	Veille & gestion e-réputation	https://148.fr
AGENCE 404	Généraliste	Audit (diagnostic) & stratégie de la présence digitale	https://www.agence404.com
AGENCE 1789	Généraliste (Webmarketing)	Audit & gestion e-réputation	https://www.1789.fr
AGENCE 33 DEGRE	Généraliste (Communication digitale)	Gestion e-réputation	https://www.agence33degres.com

AGENCE 90	Généraliste (Webmarketing)	Gestion e-réputation	https://www.agence90.fr
AGENCE AGORALYS	Généraliste	Accompagnement stratégie digitale (gestion e-réputation)	https://www.agoralys.com
AGENCE POINT	Généraliste	Gestion e-réputation	https://agencepoint.com
AGORA PULSE	Editeur	Social Media management	https://www.agorapulse.com/fr/
AKIO	Généraliste	Akio Insights (veille e-réputation & analyse sémantique)	https://www.akio.com/fr
ALERTI	Editeur / App gratuite	-	https://fr.alerti.com
ALTYVEILLE	Pure player	Veille stratégique & e-réputation	http://www.altyveille.com
ALIOZE	Généraliste	Veille & gestion e-réputation	https://www.alioze.com/
ALTERBUZZ	Pure player	-	https://www.alterbuzz.com
ANTIDOX	Généraliste	Social listening & personal branding	http://www.antidox.fr
APP ASO	Généraliste (Webmarketing)	Veille & gestion e-réputation	https://www.appaso.fr
ARTWAI	Généraliste	Formation e-réputation	https://www.artwai.com/
ATCHIK	Pure player	-	http://www.atchik.com/
AVIS PARTNERS	Généraliste	Veille & étude d'image	https://www.avisa-partners.com
AVOCATS	Ad hoc	Droit de l'e-réputation	https://www.avocats-picovschi.com
AWARIO	Editeur	-	https://awario.com
AXELLESCOM	Généraliste	Audit, benchmark, e-réputation	www.axellescom.com
AXECIBLES	Généraliste (Webmarketing)	Audit e-réputation	https://www.axecibles.com
BLUEBOAT	Pure player	-	http://www.blueboat.fr
BOOST YOUR WEB	Généraliste (Webmarketing)	Audit & stratégie e-réputation	www.boostyourweb.fr
BRAND 24	Editeur	-	https://brand24.com

BRAND MENTIONS	Editeur	-	https://brandmentions.com/#
BRANDWATCH	Editeur	-	www.brandwatch.com/fr/
BRANDYOURSELF	Editeur	Social listening & protection de la réputation	https://brandyourself.com
BUFFER	Editeur	Social Media management	https://buffer.com
BURSON COHN & WOLF (MARSTELER)	Généraliste	Audit & gestion e-réputation (crise)	https://bcw-global.com
BUZZ SUMO	Editeur	-	https://buzzsumo.com
CARMA	Editeur	-	https://carma.com
CHURCHILL	Généraliste	Veille & gestion e-réputation	https://www.agence-churchill.fr
CISON	Généraliste	Social listening & gestion e-réputation	https://www.cision.fr
CLEVER AGE	Généraliste	Gestion e-réputation	https://www.clever-age.com/fr/
COEXEL	Editeur	Veille	https://www.coexel.com
COMON	Généraliste (Communication digitale)	Gestion e-réputation	https://comon-solution.fr
COM AND HEALTH	Généraliste (Communication digitale)	Veille & gestion e-réputation	www.comandhealth.com/
COMPTOIRS	Généraliste	Audit, veille & gestion e-réputation	https://www.comptoirs.co/
CONVERSATIONNEL	Généraliste	Social listening	www.conversationnel.fr
CROWDTANGLE	App gratuite	Tracking partage des liens sur les medias sociaux	https://www.crowdtangle.com
C3M	Généraliste (Communication digitale)	Audit e-réputation	http://agence-c3m.paris
CWM CONSULTING	Généraliste (Webmarketing)	Etude e-réputation	www.cwm-consulting.com
DATA OBSERVER	Editeur	Veille en e-réputation	https://data-observer.com/do-search/
DIGIMIND	Editeur	-	https://www.digimind.com

DIGIMOOD	Généraliste (Webmarketing)	Veille & gestion e-réputation	https://www.digimood.com
DIGITAL KEYS	Généraliste (Webmarketing)	Social listening	https://www.digitalkeys.fr
DINGBAT	Généraliste (Communication digitale)	Veille & gestion e-réputation	https://dingbat.win
DRIFT DIGITAL	Généraliste (Webmarketing)	Veille & gestion e-réputation	https://www.driftdigital.fr
EDEP CONSEIL	Généraliste	Veille & gestion e-réputation	https://www.edep-conseil.fr
ED INSTITUT	Généraliste	Veille & gestion e-réputation	https://www.edinstitut.com
EOLAS	Généraliste	Gestion image (notoriété & e-réputation	https://www.eolas.fr/
E REPUTATION PARIS	Pure player	-	https://www.ereputation-paris.com
E REPUTATION.PRO	Ad hoc	Gestion de crise en e-réputation	https://e-reputation.pro
ESKIMOZ	Généraliste (Webmarketing)	Gestion de crise en e-réputation	https://www.eskimoz.fr
E-VIIU	Généraliste	Gestion de crise en e-réputation	https://www.e-viiu.agency
EXPERT IS ME	Généraliste (Communication digitale)	Audit & gestion e-réputation	https://www.expertisme.com
EXTERN MARKET	Généraliste (Communication digitale)	Audit & gestion e-réputation	https://www.extern-market.com
FIRSTLINK	Généraliste (Communication digitale)	Veille e-réputation & cartographie d'acteurs	https://first-link.fr
FABIORAMA	Généraliste	Audit, veille e-réputation	https://www.fabiorama.fr/
FB ASSOCIÉS	Ad hoc	Veille & gestion e-réputation personnelle (personal brand)	https://www.fb-associes.com
GO FISH DIGITAL	Editeur	Social listening & protection de la réputation	https://gofishdigital.com
GROUPE 361	Editeur	-	www.groupe361.com/

GUESS SUITE	Tiers de confiance	Collecte & gestion avis clients	https://www.guest-suite.com
HAAS AVOCAT	Ad hoc	Droit de l'e-réputation	https://www.haas-avocats.com
HAPPIDEV	Généraliste	Veille & gestion e-réputation	https://www.happidev.fr
HEALTHCIE	Ad hoc	Gestion e-réputation (santé)	https://www.healthcie.fr
HILL AND KNOWLTON	Généraliste	Veille & gestion e-réputation	https://www2.hkstrategies.com
HOOTSUITE	Editeur	Social Media management	https://hootsuite.com/fr/
HOPSCOTCH (Public système Pr)	Généraliste	Veille & gestion e-réputation	http://www.hopscotch.fr
HUMAN TO HUMAN	Pure player	Audit & stratégie e-réputation	www.humantohuman.fr
IDAOS	Généraliste	Veille & gestion e-réputation	https://www.idaos.com
IDENTITÉ WEB	Généraliste	Gestion de crise en e-réputation	https://identite-web.com
INDICIA	Généraliste	Gestion de crise en e-réputation	https://www.groupe-indicia.com
INDIGO	Généraliste (Communication Digitale)	Gestion e-réputation	https://www.agence-indigo.com
INFLUA	Généraliste (Webmarketing)	Audit e-réputation	https://influa.com/
INMEDIATIC	Généraliste	Veille & cartographie d'acteurs	https://www.inmediatic.net
IPROTEGO	Pure player	-	https://www.iprotego.com
JIN	Généraliste	Veille & gestion e-réputation	https://jin.fr
JNOV	Généraliste (Communication digitale)	Gestion e-réputation	https://www.jnov.fr
KB CRAWL	Editeur	-	https://www.kbcrawl.com/fr/
KEL REFERENCEMENT	Généraliste (Webmarketing)	Veille & gestion e-réputation	https://www.kelreferencement.net
KEYHOLE	Editeur	-	https://keyhole.co

KOMERES	Généraliste (Communication digitale)	Audit & gestion e-réputation	https://www.komeres.fr
KLYDE	Généraliste	Audit & gestion e-réputation	https://www.klyde.fr
KZN	Généraliste (Communication digitale)	Gestion e-réputation	https://agence-kzn.com
LA FRENCH COM	Ad hoc	Gestion de crise en e-réputation	www.lafrenchcom.fr/
LA NANTAISE DU WEB	Généraliste (Webmarketing)	Audit & gestion e-réputation	https://melwynn-rodriguez.fr
LA NETSCOUADE	Généraliste (Communication digitale)	Veille & gestion e-réputation	www.lanetscouade.com/
LAUDANS	Pure player	Audit & gestion e-réputation	https://laudans.com
LES INFOSTRATÈGES	Ad hoc	Gestion de crise en e-réputation	https://www.les-infostrateges.com
LES VEILLEURS DU NET	Généraliste (Webmarketing)	Audit & gestion e-réputation	www.lesveilleursdunet.com
LEXING ALAIN BENSOUSSAN	Ad hoc	Droit de l'e-réputation	https://www.alain-bensoussan.com
LEXIS NEXIS	Editeur	-	https://bis.lexisnexis.fr
LINKFLUENCE	Editeur	-	www.linkfluence.com/fr/
LUSSO	Généraliste (Communication digitale)	Gestion e-réputation	https://lusso.fr
MALIBOO	Généraliste (Webmarketing)	Gestion de crise en e-réputation	https://www.maliboo-referencement.fr
MA REPUTATION EN LIGNE	Ad hoc	Veille & gestion e-réputation (santé)	www.mareputationenligne.fr
MARIE ANTOINETTE	Généraliste	Gestion e-réputation	https://www.marie-antoinette.fr
MAZARINE	Généraliste	Veille & gestion e-réputation	www.mazarine.com/fr/
MEDIATOOLKIT	Editeur	Social listening & alertes sur les mentions	https://www.mediatoolkit.com

MEDIATRIUM	Généraliste (Communication digitale)	Audit & stratégie e-réputation	www.mediatrium.com
MELTWATER	Editeur	-	www.meltwater.com/fr/
MENTION	Editeur	-	https://mention.com/fr/
MENTIONLYTICS	Editeur	-	https://www.mentionlytics.com
METADOSI	Généraliste (Webmarketing)	Audit & gestion e-réputation	https://www.metadosi.fr
MINDFRUITS	Généraliste (Webmarketing)	Gestion e-réputation	https://www.mindfruits.biz
MY HOTEL REPUTATION	Ad hoc	Gestion e-réputation	https://www.my-hotel-reputation.com
MY REPUTATION AGENCY	Pure player	-	www.my-reputation-agency.fr
NANTES WEB	Généraliste	Veille & gestion e-réputation	https://www.nantesweb.bzh
NATIVE COMMUNICATIONS	Généraliste (Communication digitale)	Gestion e-réputation & modération de contenus	www.nativecommunications.com
MP6	Généraliste	Gestion e-réputation	https://mp6.fr
NEODIA	Généraliste (Webmarketing)	Gestion & netoyage pages web	http://www.neodia.fr
NEEDEO	Généraliste	Gestion e-réputation	http://www.needeo.com
NETINO	Pure player	Veille, gestion e-réputation & modération	https://netino.fr
NET OFFENSIVE	Ad hoc	Gestion de crise en e-réputation	https://www.netoffensive.com
NET REPUTATION	Pure player	-	http://www.netreputation.fr
NETVIBES	App gratuite	Veille en e-réputation	https://www.netvibes.com/fr
NET'WASH	Ad hoc	Gestion de crise	https://www.net-wash.fr
NODAL MEDIA	Généraliste (Webmarketing)	Audit, Business intelligence	https://www.nodalmedia.fr
NOTIFY	App gratuite	Social listening & alertes sur les mentions	https://notify.ly
ODW	Généraliste	Gestion e-réputation	https://www.odw.fr
ODIENS	Généraliste (Webmarketing)	Gestion e-réputation	https://www.odiens.com

OPEN2EUROPE	Généraliste (Communication digitale)	Veille & gestion e-réputation	https://open2europe.com
OPINION ACT	Pure player	-	https://www.opinionact.com
OPINION TRACKER	Editeur	Veille & analyse e-réputation	http://www.opinion-tracker.fr/
OPEN LINKING	Généraliste	Gestion de crise en e-réputation	https://www.open-linking.com
OPTIMIZE 360	Généraliste (Webmarketing)	Gestion e-réputation	https://www.optimize360.fr
OTTA	Généraliste (Communication digitale)	Veille & gestion e-réputation	https://otta.fr
OXOLEO	Généraliste (Webmarketing)	Gestion de crise en e-réputation	https://www.oxoleo.fr
PANDASEO	Généraliste (Webmarketing)	Gestion de crise en e-réputation	https://pandaseo.fr
PAPPLEWEB	Généraliste	Gestion e-réputation	https://www.pappleweb.com
PERFORMICS	Généraliste	Audit & gestion e-réputation	https://www.performics.com
PORLON AVOCATS	Ad hoc	Droit de l'e-réputation	https://www.porlon-avocats.com
POWER ON	Généraliste (Webmarketing)	Stratégie e-réputation & cartographie d'acteurs	http://poweron.fr
POWERTRAFIC	Généraliste (Webmarketing)	Veille & gestion e-réputation	https://www.powertrafic.fr
PRAXIS	Généraliste	Gestion e-réputation (noyage pages web)	http://www.praxis-eng.fr
PR FORMANCE	Généraliste	Audit & gestion e-réputation	www.prformance.com
PROOXI	Généraliste	Gestion e-réputation	https://www.prooxi.fr/fr/
PUBLICIS (NET INTELLIGENZ)	Généraliste	Social listening & gestion e-réputation	http://www.publicis-consultants.fr
PULLSEO	Généraliste (Webmarketing)	Audit & gestion e-réputation	http://www.pullseo.com
PULSAR	Editeur	-	https://www.pulsarplatform.com
QSN DIGITAL	Pure player	-	https://blog-ereputation.com

RANKUR	Editeur / App gratuite	-	https://rankur.com
RANKWELL	Généraliste (Webmarketing)	Audit & gestion e-réputation	https://www.rankwell.fr
REFERENCEUR WEB PERFORMANCE	Généraliste (Webmarketing)	Gestion e-réputation (Visibilité)	https://www.referenceur.fr
RÉPUTATION.NET	Pure player	Audit & gestion e-réputation	https://réputation.net
REPUTATION SQUAD	Pure player	-	www.reputationsquad.com
RESONEO	Généraliste (Webmarketing)	Audit & stratégie e-réputation	https://www.resoneo.com
RISE UP	Généraliste (Communication digitale)	Gestion e-réputation	http://agence-riseup.com
RIPOSTE	Ad hoc	Gestion de crise en e-réputation	https://www.ereputation-riposte.com
ROXANE	Généraliste (Communication digitale)	Veille & gestion e-réputation	https://roxane.digital
ROLLINGBOX	Généraliste	Veille & gestion e-réputation	https://rollingbox.com
RPCA	Généraliste	Audit & stratégie e-réputation	https://www.rpca.fr
SEMJI	Généraliste (Webmarketing)	Audit & gestion e-réputation	https://semji.com/fr/
SENSEEGO	Généraliste (Webmarketing)	Gestion e-réputation	www.senseego.com
SESAME (NEVOWEB)	Généraliste (Communication digitale)	Audit & stratégie e-réputation	https://www.agence-sesame.fr
SHAKEBIZ	Généraliste	Social listening	https://www.shakebiz.fr/nos-services
SHARING AGENCY	Généraliste (Communication digitale)	Gestion e-réputation	https://www.sharing.agency
SINDUP	Editeur	-	https://fr.sindup.com
SITBON JAMES	Ad hoc	Gestion de crise en e-réputation	https://sitbonjames-expertgoogle.blogspot.com

SIVVA	Généraliste	Veille & gestion e-réputation	http://sivva.fr
SKA	Généraliste (Webmarketing)	Gestion de crise en e-réputation	https://www.agence-ska.com
SMART WEB GROUP	Généraliste (Webmarketing)	Gestion de crise en e-réputation	https://www.smartweb-group.com
SOCIAL BEARING	App gratuite	Veille en e-réputation sur twitter	https://www.socialbearing.com
SOCIAL SEARCHER	App gratuite	Veille en e-réputation	https://www.social-searcher.com
SOLIKE	Ad hoc	Modération & réponse aux avis	https://www.solike.eu
SORTVOICES	Editeur	-	https://sortvoices.fr
SPITAK	Généraliste	Audit & gestion e-réputation	https://www.spitak.fr
SPRINKLR	Editeur	Social listening	https://www.sprinklr.com
SPROUTSOCIAL	Editeur	Social Media management	https://sproutsocial.com/fr/
STATUM	Généraliste (Webmarketing)	Audit & gestion e-réputation	https://www.statum.fr
STORM COMMUNICATION	Généraliste (Communication digitale)	Veille & gestion e-réputation	https://www.storm-communication.com
STUDIO GAZOLINE	Généraliste (Communication digitale)	Veille & gestion e-réputation	https://www.studiogazoline.com
SOS E REPUTATION	Pure player	-	https://www.sos-e-reputation.com
SYNTHESIO	Editeur	-	https://www.synthesio.com/fr/
TALKWALKER	Editeur	-	https://www.talkwalker.com/fr
TASTE	Ad hoc	Gestion e-réputation (agro alimentaire)	https://agencetaste.fr
TBWA	Généraliste	Audit & gestion e-réputation de crise	https://www.tbwa-corporate.com
TRADITION SITE	Généraliste	Gestion de crise en e-réputation	https://www.traditionsite.fr
TRUST FOLIO	Tiers de confiance	Collecte & gestion avis clients	https://trustfolio.co
UP BY WEB	Généraliste	Audit & gestion e-réputation	https://upbyweb.com

VANKSEN	Généraliste (Communication digitale)	Social listening & gestion e-réputation	https://www.vanksen.com/fr
VISIBRAIN	Editeur	-	https://www.visibrain.com/fr/
VISIPLUS	Généraliste	Gestion de crise en e-réputation	https://www.visiplus.com/fr/prestations/e-reputation.php?gclid=CIL0-PuvtqECFQ6Y2Aod7EKb9Q
WACAN	Généraliste (Communication digitale)	Audit & gestion e-réputation	https://www.wacan.com
WEBMII	App gratuite	Veille en e-réputation personnelle (personal brand)	https://webmii.com
WEB REPORT	Pure player	-	https://www.webreport.fr
WEDIG	Généraliste (Webmarketing)	Veille & gestion e-réputation	https://www.wedig.fr
WE REPUTATION.COM	Pure player	Audit & gestion e-réputation	https://wereputation.com
WKDO	Ad hoc	Gestion e-réputation (santé)	https://wkdo.fr
WINELLO	Editeur	Veille en e-réputation	http://www.winello.com
XIKO	Editeur	Social listening	https://www.xiko.fr
YDCA	Généraliste (Communication digitale)	Audit & gestion e-réputation	https://www.ydca.fr
YUMENS	Généraliste	Veille & gestion e-réputation	https://www.yumens.fr
50A	Généraliste	Veille & gestion e-réputation	https://www.50a.fr

Annexe 3 : La grille de lecture

Identité de l'interviewé

Nom : Prénom : Entreprise :

Objectif : vérifier si les modèles de mesure de la réputation sont utilisés de manière implicite ou objective dans les méthodes d'audit proposées par les professionnels à travers les quatre (4) principales catégories étayées par plusieurs critères.

Cocher la case correspondante : les méthodes prennent-elles en compte tous les éléments structurels des modèles types de mesure de la réputation (variable multiple)?			
Structure générale de l'offre/ solution		Oui	Non
	Input [*Données extraites*]		
	Stockage [*Sources de données*]		
	Traitement [*cfr méthode de calcul*]		
	Output [*Objets des TDB*]		
	Rendus et analyses		
	Autre		

Cocher la case correspondante : la méthode est-elle adaptable à toute stratégie d'audit ou utilisable pour un élément de l'audit? l'audit concerne une personne ou une ressource (produit ou service)?				
Contexte			Oui	Non
	Type	Multi (*Variable quali multiple ouverte)*		
		Contexte unique (*Variable quali mono ouverte)*		
			Oui	Non
	Cible	Personne		
		Ressource		

	Cocher la case correspondante, selon la technique de mesure utilisée dans la solution d'audit d'e-réputation *(question à réponse multiples)*				
Calcul			Oui	Non	
	Technique	Probabilités			
		Fuzzy logic			
		Flow base			
		Subjective logic			
		Somme / moyenne			
		Autre			

Cocher la case correspondante, selon la représentation de la valeur de la réputation *(question à réponse UNIQUE)*				
		Oui	Non	
Valeur	Quanti discrète			
	Échelle de polarité			
	Multiple			
	Floue			
	Pourcentage			
	Autre			
Classer, selon leur importance (3 étant le plus important) le critère le plus pris en compte dans la mesure de la fiabilité des *inputs*				
		(1)	(2)	(3)
Fiabilité	Autorité			
	Récense / Ancienneté			
	Nbre d'avis			
	Autre			

	Classer, selon leur importance et/ ou récurrence, le type de données collectées dans la mesure de la réputation				
Dataset			(1)	(2)	(3)
	Type	Texte			
		Visuel			
		Emoji			
		Vidéo			
		Autre			
	Classer, selon leur importance et/ ou récurrence, les sources les plus utilisées dans la collecte des données				
			(1)	(2)	(3)
	Sources	MS			
		Forums			
		MRG			
		MRS			
		Portails			
		Sites d'avis			
		Autre			

	Cocher la case correspondante : le stockage des données collectées pour la mesure de la réputation est-il aléatoire ou permanent? s'effectue t-il en local ou en externe?				
			Oui	Non	
	Stockage	Local			
		Externe			
		Aléatoire			
		Permanent			
	Classer, selon leur importance et / ou récurrence, l'objet des TDB de la réputation				
	Output	Score de réputation	(1)	(2)	(3)
		Données démographiques			
		Environnement concurrentiel			
		Détection profils influents			
		Sentiment analysis			
		Autre			

Annexe 4 : Le guide d'entretien

Le guide d'entretien comprend 3 principales parties corrélées à nos hypothèses de travail.

Partie 1 : de la perception générale du concept d'e-réputation

Hypothèse de départ : L'objet e-réputation semble, à travers nos recherches dans la littérature, être une notion "polymorphe" dont la définition et les champs d'application ne font pas encore totalement l'unanimité auprès de ses praticiens et auteurs. On pourrait également supposer que la vision *in fine* de la mise en place d'une stratégie d'audit en e-réputation est l'évaluation, la mesure d'un «objet», sur la base des informations collectées sur ce dernier.

Objectif : Appréhender l'objet e-réputation tel que perçu par les interviewés afin d'effectuer par la suite, une analyse comparative de ses différentes déclinaisons. Questions :

- Pouvez-vous me donner votre définition de la notion d'e-réputation et que suppose t-elle?

- Quel regard portez-vous sur les notions d'e-réputation et de réputation traditionnelle? Quel rapport existe t-il entre ces dernières?

- Qu'est-ce qui échappe à la mesure de l'e-réputation aujourd'hui ?

Partie 2 : des outils et/ou solutions utilisés

Hypothèse de départ : Dans la littérature scientifique (surtout anglo-saxonne), il existe une pléthore de modèles théoriques de mesure de la réputation, et on peut supposer que ces derniers ont évolué à travers le temps, et pourraient intrinsèquement être liés à plusieurs facteurs clés dont l'introduction de nouveaux concepts : de « confiance » et « risque », «influence», «identité / image», concepts de facto difficile à appréhender et maîtriser sur la toile.

Objectif : Thème essentiel pour notre étude, il nous permettra d'obtenir quelques éléments de comparaison entre la structure ou l'architecture des solutions proposées sur le marché de la mesure de la réputation, et les modèles issus de notre analyse exploratoire de la littérature scientifique. Questions :

● Quel(s) outil(s) et/ou solutions utilisez-vous dans vos stratégies d'audit (de mesure)? Sont-elles proposées sur le marché par les pure players (exemple Osculteo pour Iprotego, Easy CI de Digimind, Radarly de Linkfluence, etc.) ou les avez-vous développées en interne?

Pour ceux qui utilisent les solutions du marché :

- Comment avez-vous procédé au choix de cette/ ces solution(s)? Et ces éléments sont-ils des facteurs de succès ou limite de votre offre?
- Si vous deviez mettre en place une solution, quelles fonctionnalités privilégieriez-vous ou ajouteriez-vous à l'offre que vous utilisez?

Pour ceux ayant développé leur propre solution (les éditeurs) :

- Pourquoi ce choix (raisons marketing, d'expertise, les deux)? Et vous êtes-vous inspiré d'offres existantes en les étoffant ou est-ce une création ex-nihilo?
- Quelles sont les sources des données collectées dans la stratégie d'audit? Classer de 1 à 3 (cfr grille de lecture) selon leur importance (relances : portails, médias sociaux, MRS, MRG, forums, etc.) et typologie (texte, visuel, emoji, etc.)

- Comment mesurez-vous la pertinence/ qualité des sources? Classer de 1 à 3 selon l'importance (cf grille de lecture) : relances (nombre d'avis, ancienneté, autorité de la source, etc.)
- Votre offre tient-elle ses promesses de votre point de vue pour développer votre chiffre d'affaires?

Partie 3 : de l'offre d'audit en e-réputation dans sa globalité

Hypothèse : Avec l'essor du web social, la production et la diffusion exponentielle de "données" semble avoir favorisé le développement de nouvelles méthodes et pratiques de la part des professionnels de la réputation, tant dans la collecte que dans l'analyse des informations collectées.

Objectif : Appréhender la ou les méthode(s) d'audit que proposent les professionnels de la réputation dans sa/leur globalité, afin d'identifier, dans l'optique où elles existent, l'émergence de pratiques qui ont mutées. Questions :

- Comment est structurée votre offre d'audit d'e-réputation? Et quelles en sont les principales étapes? En cas de silence, relances (sourcing, requêtage, crawling, analyse, rapport, dataviz, etc.)
- Pouvez-vous nous citer 2-3 cas concrets de problématiques nécessitant un audit et sur lesquelles vous avez ou pouvez travailler?
- Vos clients sont-ils des entreprises (grands comptes, PME, etc.), des particuliers (dirigeants, hommes politiques, etc.?), les deux? Et votre offre diffère-t-elle selon les publics?
- Quel positionnement priorisez-vous? (offre complète, facilité de prise en main, le sur mesure, ad hoc, la notoriété, etc.)
- Sous quelle forme sont présentés les résultats au client (Cartographies, TDB, etc.) & que contiennent les rapports?

- Quels sont les développements à venir ? Quelles fonctionnalités nouvelles vous paraissent essentielles dans un avenir proche ?
- Sur quelle périodicité sont collectées les données analysées (1 an, pas de temporalité, etc.)?

Pour conclure, pouvez-vous nous parler de vous & de votre entreprise:

- Pourriez vous me parler de la formation que vous avez suivie ? (Lettres et sciences humaines, Communication, technologique, scientifique)?
- Quelles compétences sont selon vous, indissociables de ce profil ? (relances : SMM, CM, ORM, DA, SEO/SEA, etc.)
- Estimez-vous que votre expertise dans le domaine de la réputation vous a demandé une grande part d'auto-formation ?
- Quel est le statut de votre entreprise? (agence de communication, marketing digital, publicité, RP, agence e-réputation, etc.) et quel en est le cœur de métier?
- Quelle part représente, dans la mesure où elle n'est pas le cœur de métier, la mesure de la réputation?
- Depuis quand la mesure de la réputation a-t-elle été ajoutée à l'offre et quelles en sont les raisons?
- Pouvez-nous vous dire quel avenir vous voyez aux études de réputation?

Annexe 5 : Les 20 catégories d'unités de sens

Catégories d'unités de sens issues de l'AIG qui feront par la suite l'objet d'une réduction.

1. E-réputation et cadrage conceptuel : proposition de définitions
2. E-réputation et réputation : transposition de notions, nouvel objet ?
3. E-réputation personnelle versus de marque versus produit et/ service
4. La gestion de crise en e-réputation
5. E-réputation et approches en communication digitale
6. E-réputation et approches en webmarketing
7. E-réputation et pratiques en relations publique
8. De l'e-réputation à la confiance : processus itératif entre input et output
9. La mesure de l'influence en e-réputation
10. Les limites des solutions de *listening* proposées par les éditeurs
11. E-réputation : une approche combinée homme - machine (algorithme et humain)
12. Les méthodes de mesure proposées sur le marché
13. E-réputation et *sourcing* : de l'importance de la catégorisation des bouquets de sources
14. Les phases de cadrage et paramétrage comme clé d'entrée de la qualité des output
15. Les tableaux de bord de l'e-réputation : de la granularité des widgets
16. Des rendus : entre alertes, newsletters et rapports d'audit
17. E-réputation et nouvelles restrictions légales
18. Les nouvelles problématiques de réputation
19. Des compétences indispensables / indissociables de la gestion de l'e-réputation
20. De la position des entités dans le marché de l'e-réputation

LISTE DES FIGURES

Figure n°1 : Nuage de mots clés créé à partir des suggestions sur la plateforme « nuage de mots » 15
Figure n°2 : Ce que suppose la réputation 56
Figure n°3 : La relation entre le truster et le trustee, source Grandison & Sloman (2000)....73
Figure n°4 : Les cinq typologies de confiance, source Josang al. (2017) à partir des travaux de Grandison et Sloman 79
Figure n°5 : Les dimensions de mesure de la confiance, source : classification of trust measure, josang et al. (2007)....84
Figure n°6 : Panorama 2020 des médias sociaux, source Fred Cavazza 99
Figure n°7 : Catégories de classification des modèles, source Koutrouli et Tsalgatidou....129
Figure n°8 : Distinction entre les logiques booléennes et floues, source TechTarget 174
Figure n°9 : Processus de collecte des avis vérifiés, source Avis vérifiés 190
Figure n°10 : The e reputation market map Aref JDEY, source : site demain la veille....196
Figure n°11 : Cartographie des acteurs du marché de la réputation, Source : Blueboat 197
Figure n°12 : Proposition d'une cartographie des acteurs du marché de la réputation....200
Figure n°13 : Exemple de pure players en e-réputation 201
Figure n°14 : Exemple d'éditeurs de solutions en e-réputation 202
Figure n°15 : Exemple d'agences généralistes en e-réputation 204
Figure n°16 : Exemple d'agences ad hoc en e-réputation 205
Figure n°17 : Exemples d'opérateurs booléens utilisés pour le requêtage, source site web corporate Sproutsocial....209
Figure n°18 : Classification des requêtes, interface de démonstration du tableau de bord de la plateforme Digimind 212
Figure n°19 : Analyse de sentiment, source interface de démonstration du tableau de bord de la plateforme Digimind 214
Figure n°20 : Notification de contenus non pris en compte dans l'analyse de sentiment, source interface de démonstration du tableau de bord de la plateforme Digimind....214
Figure n°21 : Exemple de mention sur une entité sur Twitter, source : interface de démonstration du tableau de bord de la plateforme Digimind 215

Figure n°22 : Typologie d'emojis pris en compte dans l'analyse du sentiment, Source site web corporate de Linkfluence....216
Figure n°23 : Critères de réputation du tableau de bord, source : Interface de démonstration (Source Digimind)....217
Figure n°24 : Interface de démonstration du tableau de bord relative aux tendances sur une requête (Source : Digimind)....218
Figure n°25 : Exemple de (re) présentation de données démographiques, source : site web corporate de Linkfluence....219
Figure n°26 : Exemple de présentation des mentions par type de média, source : Digimind220
Figure n°27 : Exemple de classification des services d'une entité par tonalité, source Digimind....221
Figure n°28 : Analyse concurrentielle de la réputation d'une entité et de son marché , source : site web corporate Digimind....222
Figure n°29 : Contexte de mesure de la réputation....244
Figure n°30 : Structure générale des méthodes de mesure de la réputation....245
Figure n°31 : La dimension de calcul dans la mesure de la réputation....248
Figure n°32 : Exemple de Mail / Inmail envoyé....254
Figure n°33: Catégorisation des unités de sens, réalisée à partir des entrevues....297
Figure n°34 : Les méthodes de mesure et les approches théoriques qui les ont inspirées....302

LISTE DES TABLEAUX

Tableau n°1 : Approche comparative par discipline de la perception et de l'objet pris en compte dans la mesure de la réputation 53
Tableau n°2 : Synthèse des dimensions de classification des modèles théoriques 150
Tableau n°3 : Concepts utilisés dans notre tableau de classification des modèles.....151
Tableau n°4 : Nos dimensions de classification des modèles théoriques de la réputation.....183
Tableau n°5 : Caractéristiques des typologies d'entretiens par Imbert (2010), inspiré des travaux de De Ketele et Roegiers (1996, p.172) 251
Tableau n°6 : Comparaison des approches d'analyse qualitative de Blais et Martineau (2006) à partir des travaux de Thomas (2006, p.241).....256
Tableau n°7 : Processus de codage des données brutes de Blais et Martineau (2006), adapté de Thomas (2006, p.242), d'après Creswell (2002, p.266).....259

LEXIQUE

Ce lexique vise à faciliter la compréhension de certains des concepts utilisés dans nos travaux. Nous tenons à souligner que les terminologies proposées dans cet index sont indépendantes des définitions génériques associées aux notions, et sont plutôt relatives au contexte au sein duquel les notions ont été employées dans la littérature scientifique sur la mesure de la réputation.

Agent : un agent, dans une approche informatique, se dit d'un processus, d'un robot, d'un humain, etc., toute forme qui interagit avec son environnement. Les agents sont, dans la mesure de la réputation, utilisés pour représenter les entités au sein d'un système multi agents (Maes, 1994[453], Maamar & Moulin, 1997[454], Jennings & Wooldridge, 1998[455]).

Authoritative node detection : il s'agit d'une méthode permettant de détecter les noeuds (ou utilisateurs) les plus influents en termes d'autorité au sein d'un réseau ou d'une communauté (Kumar & al., 2008[456]; Sebastian & al., 2017[457], Oro & al., 2017[458]).

---

[453] P. Maes. Agents that reduce work and information overload. Communications of the ACM, 37(7) :31-40, July 1994;

P. Maes. Social interface agents : Acquiring competence by learning from users and other agents. In O. Etzioni, editor, Software Agents - Papers from the 1994 Spring Symposium (Technical Report SS-94-03), pages 71-78. AAAI Press, March 1994 .

[454] Z. Maamar and B. Moulin. An agent based approach for intelligent and cooperative systems. In Proc. of IEEE Knowledge and Data Engineering Exchange Workshop'97 (KDEX-97), pages 41-47, Newport Beach, California, 1997.

[455] N. R. Jennings, M. Wooldridge, and K. Sycara. A roadmap of agent research and development. Int Journal of Autonomous Agents and Multi-Agent Systems, 1(1) :7- 38, 1998.

456 Ravi Kumar, Kevin Lang, Cameron Marlow, Andrew Tomkins (2008). Efficient Discovery of AuthoritativeResources, IEEE.

[457] Sebastian, R.S.A., Aguilera, F., Nun~ez-Gonzalez, J.D. and Gran~a, M. (2017), "Semantically enhanced network analysis for influencer identification in online social networks", Neurocomputing, Vol. 326, pp. 71-81

[458] Oro, Ermelinda & Pizzuti, Clara & Procopio, Nicola & Ruffolo, Massimo. (2017). Detecting Topic Authoritative Social Media Users: A Multilayer Network Approach. IEEE Transactions on Multimedia. PP. 1-1. 10.1109/TMM.2017.2763324. Accessible sur http://staff.icar.cnr.it/staff/pizzuti/public_html/pubblicazioni/IEEETM2017.pdf

Authority score : l'authority score de SEMrush est comme ils le soulignent sur leur site « *une métrique utilisée pour mesurer la qualité globale et les performances SEO d'un domaine ou d'une page web* »[459]. Il vise à accroître le positionnement d'une page dans les moteurs de recherche.

Black hat SEO : qui rappelons-le ne respecte pas la charte éthique et les recommandations des moteurs de recherche, fait référence à un ensemble de pratiques de référencement naturel visant à accroître la visibilité d'une page (à la suroptimiser) dans les résultats de recherche. Il contribue à la construction de l'e-réputation même si la pratique reste déconseillée (1re position[460], Content marketing[461], Semrush[462], Smart keyword[463]).

Cloaking : se dit d'une technique de SEO black hat (utilisée par certaines agences généralistes de notre cartographie), qui vise à optimiser l'indexation d'une page pour les robots. Le principe du cloaking est simple : afficher un contenu différent pour chacun des visiteurs d'une page : contenus suroptimisés pour les robots d'indexation, et contenus génériques pour les internautes.

Crawl : technique d'exploration du web, le crawl vise à extraire et indexer un corpus relatif à un objet de recherche. C'est la méthode utilisée dans la collecte des données (veille, *social listening*).

Domain Authority : ou l'autorité de domaine se dit d'un score de *ranking* des moteurs de recherche (compris entre 1 et 100) développé par l'entreprise MOZ[464]. Il permet de prédire, selon sa pertinence dans un domaine spécifique, le classement d'une page web dans les résultats de recherche.

---

[459] Définition accessible en ligne sur le blog de SEMrush http://urlr.me/BQTXh

[460] Accessible en ligne sur https://www.1ere-position.fr/blog/black-hat-seo-la-fausse-bonne-idee/

[461] Accessible en ligne sur https://www.redacteur.com/blog/seo-techniques-black-hat/

[462] Accessible en ligne sur https://fr.semrush.com/blog/seo-black-hat/

[463] Accessible en ligne sur https://smartkeyword.io/referencement-naturel-black-hat-seo/

[464] Définition proposée par l'entreprise accessible en ligne sur https://moz.com/learn/seo/domain-authority

Flow based : ou approche basée sur les flux, est une technique de calcul du score de réputation utilisée par plusieurs modèles (le Pagerank par exemple).

Fuzzy logic : ou logique floue est une méthode de mesure de l'e-réputation utilisée par certains modèles (REGRET, 2002; Yu & al., 2014; Yan & al., 2015; Chandran & al., 2016 par exemple).

Grey hat SEO : technique de référencement naturel qui se situe à la limite des méthodes de white Hat et de black hat. À la différence de ces dernières, les pratiques de Grey Hat ne vont pas, de manière spécifique, à l'encontre des limites imposées par les moteurs de recherche. Elles consistent, à titre d'exemple, à l'achat de liens, les variations de contenu (*content spinning*) ou encore à l'utilisation de manière excessive de mots-clés (SEMrush[465], Boostability[466], seo.fr[467], Webconversion[468], Search engine journal[469], Panda seo[470]).

Information gathering / aggregation : phase de collecte des données dans le processus d'évaluation.

Incentive / punishment feedback : principe de récompense / de punition visant à garantir d'une part l'exhaustivité, et d'autre part la fiabilité des données collectées dans la mesure de la réputation. Ce procédé est utilisé dans la littérature scientifique par plusieurs modèles (TRAVOS de Teacy & al., 2006 et le modèle de Jurca & Faltings, 2003 par Exemple).

Input : ensemble de données d'entrée qui servent à mesurer la réputation d'une entité. Dans la littérature scientifique sur la mesure de la réputation, les inputs peuvent provenir de plusieurs sources directes et indirectes.

---

[465] Accessible sur https://www.semrush.com/blog/white-hat-vs-black-hat-seo-and-gray-hat-too/

[466] Accessible en ligne https://www.boostability.com/content/black-hat-vs-white-hat-vs-gray-hat-seo

[467] Accessible sur https://www.seo.fr/definition/black-hat-seo

[468] Accessible sur https://www.webconversion.fr/white-hat-black-hat-grey-hat/

[469] Accessible sur https://www.searchenginejournal.com/white-hat-vs-black-hat-vs-gray-hat-seo/365142/

[470] Accessible en ligne https://pandaseo.fr/differences-entre-white-hat-grey-hat-seo/

Key opinion leader (KOL): leader d'opinion ou influenceur numérique se dit d'un internaute qui possède un indice ou score d'influence élevé (Sebastian & al., 2017[471]; More & Lingam, 2017[472]; Quan & al., 2020[473]). La littérature scientifique identifie trois (3) principaux types d'influenceurs sur les plateformes socionumériques : les *hub*, les autorités et les *spreaders*.

Mobile Ad hoc Network (MANET) : ou réseau mobile ad hoc, est un système décentralisé construit autour de noeuds (d'utilisateurs) mobiles sans fil et de routeurs. Certains modèles de réputation (que vous retrouvez dans le tableau présenté en annexe 1) sont construits autour de cette structure distribuée (Chlamtac & al., 2003[474]; Belhoul, 2013[475], Eurocom[476], Science Direct[477]).

Model's granularity : il s'agit de la structure à partir de laquelle les modèles conceptuels sont construits et les contextes au sein desquels ils peuvent être utilisés (centralisé, décentralisé ou générique).

Multi Agent Systems (MAS) : un système multi-agents est composé de plusieurs agents qui interagissent dans un environnement partagé afin d'atteindre un objectif commun. Construits autour d'une approche distribuée, les MAS ont fait l'objet de nombreuses applications dans la

---

[471] Sebastian, R.S.A., Aguilera, F., Nun~ez-Gonzalez, J.D. and Gran~a, M. (2017), "Semantically enhanced network analysis for influencer identification in online social networks", Neurocomputing, Vol. 326, pp. 71-81

[472] More, J.S. and Lingam, C. (2017), "A SI model for social media influencer maximization", Applied Computing and Informatics, Vol. 15 No. 2, pp. 102-108.

[473] Quan, T.-T., Mai, D.-T., & Tran, T.-D. (2020). CID: Categorical Influencer Detection on microtext-based social media. Online Information Review, ahead-of-print(ahead-of-print). doi:10.1108/oir-02-2019-0062

[474] Imrich Chlamtac, Marco Conti, Jennifer J.-N. Liu (2003). Mobile ad hoc networking: imperatives and challenges. doi:10.1016/S1570-8705(03)00013-1. Accessible sur http://manet.eurecom.fr/chlamtac.pdf

[475] Yacine Belhoul. Graph-based Ad Hoc Networks Topologies and Business Process Matching. 2013

[476] Accessible en ligne http://manet.eurecom.fr

[477] Accessible en ligne https://www.sciencedirect.com/topics/computer-science/mobile-ad-hoc-network

mesure de la réputation (The Alan Turing Institute[478]; Singh, 1974[479]; Jennings, 1995[480]; Glaser, 1996[481]; Flores-Mendes, 1999[482]; Chaib-draa & al., 2017[483]) .

Natural Language Processing : ou traitement automatique du langage naturel est une branche de l'intelligence artificielle qui fait référence à l'interaction homme-machine, et plus précisément à la programmation informatique dans le traitement et l'analyse de données (à un ou plusieurs niveaux linguistiques) issues du Language naturel (Chowdhury, 2003[484], Liddy, 2001[485], Allen, 2003[486], IBM Cloud Lear hub[487], Machine Learning mastery[488]).

Private Blog Networks (PBN) : fait référence à une technique utilisée par certaines agences généralistes et ad hoc dans le processus de construction de la réputation. Elle vise à créer un ensemble de sites web qui génèrent des *backlinks* vers un autre site afin d'accroître son d*omain authority,* sa position dans les résultats de recherche.

---

[478] Accessible sur https://www.turing.ac.uk/research/interest-groups/multi-agent-systems

[479] M. P. Singh. Multiagent Systems : A Theoretical Framework for Intentions, Know- How,and Communications (LNAI Volume 799). Springer-Verlag : Heidelberg, Germany, 1994.

[480] N. R. Jennings. Controlling cooperative problem solving in industrial multi-agent systems using joint intentions. Artificial Intelligence, 74(2), 1995.

[481] N. Glaser. Contributing to Knowledge Modeling in a Multi-Agent Framework : the CoMoMAS Approach. PhD thesis, Université Henri Poincaré, Nancy 1, France, 1996.

[482] Roberto A. Flores-Mendez (1999). Towards the Standardization of Multi-Agent Systems Architectures: An Overview. ACM Crossroads' special issue on Intelligent Agents. Accessible sur https://pages.cpsc.ucalgary.ca/~robertof/publications/crossroads99/crossroads99.pdf

[483] B. Chaib-draa, I. Jarras et B. Moulin (2017). Systèmes multiagents : Principes généraux et applications, Département d'Informatique, Pavillon Pouliot, Université Laval, Ste-Foy, PQ, Canada. Accessible sur http://www.sietmanagement.fr/wp-content/uploads/2017/12/Chaib-draa2001.pdf

[484] Chowdhury, G. (2003) Natural language processing. Annual Review of Information Science and Technology, 37. pp. 51-89. ISSN 0066-4200. Accessible en ligne https://strathprints.strath.ac.uk/2611/1/strathprints002611.pdf

[485] Liddy, E.D. 2001. Natural Language Processing. In Encyclopedia of Library and Information Science, 2nd Ed. NY. Marcel Decker, Inc. Accessible sur http://urlr.me/k4rsN

[486] James F. Allen. 2003. Natural language processing. Encyclopedia of Computer Science. John Wiley and Sons Ltd., GBR, 1218–1222.

[487] Accessible sur https://www.ibm.com/cloud/learn/natural-language-processing

[488] Accessible sur https://machinelearningmastery.com/natural-language-processing/

Output : par *output*, on entend les résultats issus du processus de mesure de la réputation. Selon les objectifs établis en amont, l'*output* peut prendre plusieurs formes : indice ou score de réputation, analyse de sentiments, détection de co-occurence de termes, cartographie d'acteurs par Exemple.

Reputation dashboard : tableau de bord modulaire proposé par les éditeurs de solutions en e-réputation qui reprend l'ensemble des outputs à l'issue de l'analyse.

Reputation estimation / Assessment / **scoring** : processus de mesure de la réputation. Dans les approches théoriques et professionnelles, la mesure de la réputation suppose trois (3) principales étapes : la phase de collecte (la paramétrage et le sourcing), le traitement des données (nettoyer, classer et catégoriser), et enfin la phase d'analyse et de visualisation des résultats qui vise à les rendre intelligibles.

Reputation representation : représentation de la / des valeur (s) de réputation.

Social Comparison Theory (SCT) : ou théorie de la comparaison sociale développée en 1954 par Leon Festinger, est relative au processus psychologique à partir duquel les individus se comparent aux autres. Cette théorie a notamment contribué, dans les modèles théoriques, au développement de mécanismes visant à comprendre et à mesurer l'influence numérique (Festinger, 1954[489], Goethals & Darley, 1977[490], Suls & & Wheeler, 2000[491], Fridman & Kaminka, 2007[492]).

---

[489] Festinger, L. 1954. A theory of social comparison processes. Human Relations, 7: 117–140.

[490] Goethals, G.R. & Darley, J.M. (1977). Social comparison theory: An attributional approach. In J.M. Suls & R.L.
Miller (Eds.), 'Social comparison processes: Theoretical and empirical perspectives (pp 259–278). Washington, DC: Hemisphere.

[491] Suls, J., & Wheeler, L. (2000). A Selective history of classic and neo-social comparison theory. Handbook of Social Comparison. New York: Kluwer Academic/Plenum Publishers

[492] Fridman, N., & Kaminka, G.A. (2007). Towards a Cognitive Model of Crowd Behavior Based on Social Comparison Theory, In AAAI.

Search Engine Result Page (SERP) : désigne la page de résultats d'un moteur de recherche lorsqu'on exécute une requête (une recherche) sur un phénomène (Techtarget[493], page on power[494], 1re position[495], seo.fr[496]).

Trust Flow / Citation Flow (TF / CF) : sont des indicateurs proposés par la plateforme Majestic SEO. Le *trust flow* se dit du degré de confiance (indicateur qualitatif) attribué aux *backlinks* qui renvoient vers un site site. Le *citation flow,* évalue le nombre (indicateur quantitatif) des *backlinks* vers un site web (Netoffensive[497], Oncrawl[498], josptimisemonsite[499], seo.fr[500]).

Vehicular Ad hoc Network (VANET) : ou réseau ad hoc de véhicules s'appuie sur le principe des MANET, et est composé d'un ensemble de véhicules en mouvement ou à l'arrêt connectés par un réseau sans fil (Science direct[501]). C'est un des domaines d'application (relatif à l'IoT) des modèles de mesure de la réputation (le modèle TRIP de Marmol & Perez, 2012 ou celui de Wang & al., 2016 par exemple).

---

493 Accessible sur https://whatis.techtarget.com/definition/search-engine-results-page-SERP

494 Accessible sur https://www.pageonepower.com/search-glossary/what-is-serp

495 Accessible sur https://www.1ere-position.fr/definitions/serp-search-engine-result-page/

496 Accessible sur https://www.seo.fr/definition/serp-search-engine-results-page

497 Accessible sur http://urlr.me/PxrX8

498 Accessible sur http://urlr.me/4XLzq

499 Accessible en ligne sur https://www.joptimisemonsite.fr/definition-interpreter-trust-flow-citation-flow/

500 Accessible sur https://www.seo.fr/definition/trust-flow

501 Accessible sur https://www.sciencedirect.com/topics/computer-science/vehicular-ad-hoc-network

TABLE DES MATIÈRES

REMERCIEMENTS 3
RÉSUMÉ 5
ABSTRACT 6
SOMMAIRE 7
LISTE DES PRINCIPALES ABRÉVIATIONS 11
INTRODUCTION GÉNÉRALE 13
1. Rappel historique 17
1.1. L'émergence du world wide web ! 18
1.2. L'avènement du web social dit 2.0 19
1.3. Le web des données 21
2. Contexte de la recherche 24
3. Objectif de la recherche 27
4. Problématique 28
5. Hypothèses de travail 30
6. Présentation du plan de la thèse 32
PARTIE I : DE LA RÉPUTATION À L'E-RÉPUTATION 34
CHAPITRE 1 : LA RÉPUTATION, UN OBJET DES SCIENCES HUMAINES ET SOCIALES 36
Introduction du chapitre 1 37
1. La réputation : historique, étymologie et définition 38
2. La réputation comme ressource et capital : approches quantifiantes 41
3. La réputation comme construit social de l'évaluateur ou du récepteur 45
4. Les espaces publics de la réputation : le paradigme info communicationnel 48
5. La réputation : une définition stabilisée à travers les différentes disciplines 50
Conclusion du chapitre 1 57
CHAPITRE 2 : DE L' E-RÉPUTATION À L'INFLUENCE 58
Introduction du chapitre 2 59
1. La trace numérique comme source de l'e-réputation 62
2. La construction des identités numériques 67
3. La construction de la confiance numérique 72
3.1. La confiance : un concept abstrait, complexe et multidimensionnel… 72
3.2. La confiance comme concept composite : entre risque, certitude et délégation 74
3.3. Les interactions au coeur de la construction de la confiance 81
4. La mesure de l'influence numérique 85
4.1. La détection de profils d'influence 89
4.2. La classification des influenceurs selon leur typologie 92
4.3. L'attribution d'un score d'influence 93

Conclusion du chapitre 2 95
CHAPITRE 3 : DES MEDIAS SOCIAUX À LEUR ANALYSE 97
Introduction du chapitre 3 98
1. Le web 2.0 et les plateformes collaboratives : l'avènement du "consom'acteur" 100
2. L'émergence de la fouille d'opinion et du sentiment analysis dans l'analyse des réseaux sociaux (ARS) 104
2.1. Le paramétrage 105
2.2. La collecte et le traitement des données : préparation du corpus et archivage 107
2.3. La phase d'analyse : le dilemme de l'analyse de sentiment 110
3. La visualisation des données : le boom des cartographies dans l'analyse des réseaux sociaux 113
Conclusion du chapitre 3 116
Conclusion de la partie I 117
PARTIE II : DES MODÈLES THÉORIQUES AUX MÉTHODES DE MESURE 120
Introduction de la partie II 121
CHAPITRE 4 : LES MODÈLES CONCEPTUELS DE MESURE DE LA RÉPUTATION EN LIGNE, PROPOSITION D'UNE TAXONOMIE 122
Introduction du chapitre 4 123
1. Revue de la littérature : les approches scientifiques de classification 126
1.1. Les dimensions de classification de Sabater et Sierra, 2005 126
1.2. Les dimensions de classification de Koutrouli et Tsalgatidou 2006 / 2012 129
1.3. Les dimensions de classification de Marti and Garcia-Molina, 2006 130
1.4. Les dimensions de classification de Wang and Vassileva, 2007 132
1.5. Les dimensions de classification de Hoffman et al., 2009 133
1.6. Les dimensions de classification de Yuan et al., 2012 138
1.7. Les dimensions de classification de Vavilis et al., 2014 141
1.8. Les dimensions de classification de Hendrikx et al., 2015 145
1.9. Les dimensions de classification de Granatyr et al. 2015 147
2. La mesure de la réputation en théorie : approche contextuelle et méthodologique 152
2.1. Les contextes d'utilisation des modèles 153
2.2. La phase de collecte : des sources de données à leur catégorisation 156
2.2.a. Le sourcing : des sources et typologies de données collectées 156
2.2.b. La catégorisation : les métriques utilisées dans le classement 160
2.3. La phase d'analyse des données : des métriques de réputation à la méthode de calcul utilisée 162
2.3.a. Des métriques de réputation comme objet de mesure 162
2.3.b. Les méthodes de calcul utilisées 169
3. Proposition d'une nouvelle taxonomie : nos dimensions de classification 179
3.1. Les modèles primaires ou sommaires 179
3.2. Les modèles ad hoc 180

3.3. Les modèles sophistiqués 181
3.4. Les approches modulaires / génériques (flexibles) 182
Conclusion du chapitre 4 184
CHAPITRE 5 : LE MARCHÉ DE LA MESURE DE LA RÉPUTATION EN LIGNE 185
Introduction du chapitre 5 186
1. L'e-réputation : un marché en pleine mutation 187
1.1. Les techniques et méthodes de collecte des données (les données utilisées) : échantillon versus quantité exponentielle de données 187
1.2. Les techniques et méthodes de mesure : approche traditionnelle versus algorithmique.193
1.3. Des problématiques liées à la réputation : l'objet sur lequel porte la mesure (image versus totalité des attributs) 195
2. Les agences et le marché : un marché français en pleine expansion....196
3. Les services proposés : une offre foisonnante et sur mesure 207
3.1. L'analyse de la présence en ligne : entre intelligence économique, veille, audit et social listening 208
3.2. La gestion de la réputation : de la gestion de crise au recours légal 224
3.2.a. La modération de contenus sur les media sociaux 224
3.2.b. La suppression ou le flooding (noyage) des contenus négatifs dans les résultats de recherche 227
3.2.c. Le recours légal en cas de diffamation 229
3.3. La construction de la réputation : l'approche par popularité 232
Conclusion du chapitre 5 236
Conclusion de la partie II 237
PARTIE III : UNE ENQUÊTE AUPRÈS DES PROFESSIONNELS DE LA RÉPUTATION EN LIGNE....239
Introduction de la partie III 240
CHAPITRE 6 : LA MÉTHODOLOGIE DE TRAVAIL 241
Introduction du chapitre 6 242
1. De la nécessité d'une approche quantifiante : utilisation d'une grille de lecture....243
1.1. Le contexte d'utilisation ou l'objectif visé : la phase de cadrage 244
1.2. La structure (de la solution ou de la méthode) 245
1.3. La méthode de mesure utilisée 247
1.4. Les données de mesure....249
2. Une approche inductive par entretiens semi directifs 250
2.1. Déroulé des entretiens 253
2.2. Proposition de catégories d'analyse : du processus de réduction des données 255
Conclusion du chapitre 6 263
CHAPITRE 7 : LE TERRAIN ET LES RÉSULTATS....264
Introduction du chapitre 7 265

1. Échantillonnage : d'un processus d'analyse exploratoire du web à la constitution du corpus 266
1.1. Le cadrage de la recherche 266
1.2. La constitution du corpus 268
1.3. La phase de traitement : de la prise de contact à la collecte 269
2. Présentation des résultats et analyse des données 270
2.1. L'e-réputation, un objet intangible nécessitant une approche pluridisciplinaire 270
2.2. L'e-réputation comme transposition de la réputation via l'interface numérique? 277
2.3. L'e-réputation, une offre sur mesure 280
2.4. La mesure de l'e-réputation, une méthode en construction 284
2.5. Différents types de livrables 289
2.6. La mesure de l'e-réputation, un avenir mitigé? 291
3. Interprétation des résultats et discussion 300
3.1. Les modèles théoriques sur lesquels sont construites les méthodes de mesure : des questions de recherche à la problématique… 300
Conclusion du chapitre 7 307
Conclusion de la partie III 308
CONCLUSION GÉNÉRALE ET PERSPECTIVES 309
1. Quelles réponses avons-nous pu apporter à nos hypothèses ? 311
2. Quelles premières conclusions? 317
3. Quelles sont les limites de nos travaux et les perspectives futures? 321
BIBLIOGRAPHIE 324
ANNEXES 352
Annexe 1 : Tableau de classification des modèles théoriques de mesure de l'e-réputation 353
Annexe 2 : Le marché de l'e-réputation 364
Annexe 3 : La grille de lecture 375
Annexe 4 : Le guide d'entretien 380
Annexe 5 : Les 20 catégories d'unités de sens 384
LISTE DES FIGURES 385
LISTE DES TABLEAUX 387
LEXIQUE 388
TABLE DES MATIÈRES 395