

UNIVERSITÉ DE FRANCHE-COMTÉ
**ÉCOLE DOCTORALE « LANGAGES, ESPACES, TEMPS,
SOCIÉTÉS »**

Thèse en vue de l'obtention du titre de docteur en
SCIENCES DE L'INFORMATION ET DE LA COMMUNICATION

**TECHNOLOGIES SÉMANTIQUES POUR UN SYSTÈME ACTIF
D'APPRENTISSAGE**

Présentée et soutenue publiquement par

Ioan SZILAGYI

Le 26 mars 2014

Sous la direction de M. le Professeur Ioan ROXIN

Membres du jury :

Bruno BACHIMONT, Professeur à l'université de Technologie de Compiègne

Sylvie CALABRETTO, Professeur à l'INSA de Lyon, Rapporteur

Sébastien GEORGE, Professeur à l'université du Maine, Rapporteur

Christophe NICOLLE, Professeur à l'université de Bourgogne

Ioan ROXIN, Professeur à l'université de Franche-Comté

Philippe VIDAL, Professeur à l'université de Toulouse III



Technologies sémantiques pour un système actif d'apprentissage

THÈSE

présentée et soutenue publiquement le *26 mars 2014*

pour l'obtention du

Doctorat de l'Université de Franche-Comté

(spécialité : Sciences de l'Information et de la Communication)

par

Ioan SZILAGYI

Membres du jury :

M. Bruno BACHIMONT, Professeur, Université de Technologie de Compiègne, Examineur

Mme. Sylvie CALABRETTO, Professeur, INSA Lyon, Rapporteur

M. Sébastien GEORGE, Professeur, Université du Maine, Rapporteur

M. Christophe NICOLLE, Professeur, Université de Bourgogne, Examineur

M. Ioan ROXIN, Professeur, Université de Franche-Comté, Directeur de thèse

M. Philippe VIDAL, Professeur, Université Toulouse III, Président du Jury

Remerciements

Tout au long de mon travail de recherche, j'ai eu l'énorme chance de pouvoir bénéficier des conseils, du soutien et de l'amitié de nombreuses personnes qui ont enrichi mon expérience de vie à la fois au niveau professionnel et personnel. C'est pour moi un réel plaisir de pouvoir les remercier à l'aboutissement de cette thèse.

Je souhaite exprimer ma reconnaissance à M. Ioan ROXIN, mon directeur de thèse, qui a encadré mes travaux de recherche avec énergie, rigueur et un enthousiasme communicatif. Ses conseils et les nombreuses discussions que nous avons eues m'ont aidé à développer ma réflexion et mon sens critique. Je le remercie de son investissement tout au long de ces années, de ses encouragements, de la confiance qu'il m'a accordée et de m'avoir fait découvrir le monde de la recherche scientifique.

Mes remerciements s'adressent particulièrement à Mme. Sylvie CALABRETTO et à M. Sébastien GEORGE qui ont accepté le rôle de rapporteurs. Leurs appréciations et retours très pertinents m'ont encouragé dans l'aboutissement de ce travail. Je remercie également M. Bruno BACHIMONT et M. Christophe NICOLLE d'avoir accepté d'évaluer ce travail et d'être membres du jury. Je les remercie aussi pour les échanges que nous avons pu avoir à cette occasion. Je tiens à remercier M. Philippe VIDAL pour l'honneur qu'il m'a fait en acceptant d'évaluer ce travail de thèse. Je lui suis aussi reconnaissant d'avoir présidé le jury de thèse et de ses mots d'encouragement.

Je remercie Mme. Ana ROXIN et M. Daniel MERCIER pour leur travail de relecture. J'ai beaucoup apprécié leurs retours, et leurs remarques ont été d'un grand secours pour la finalisation de mon manuscrit.

Pour leurs retours, les corrections et les idées échangées je remercie mes collègues de l'équipe Objets et Usages Numériques (OUN – ELLIADD) : Federico TAJARIOL, Radu BALOG-CRIȘAN, Delphine TIROLE, Diane DUFORT, Antonin SEGALT, Clémence ANDRÉYS, Françoise GREFFIER et Christophe REFFAY. Je remercie également mes collègues pour les suggestions, discussions et retours pendant nos réunions d'équipe : Alexandre COUTANT, Jean-Claude DOMENGET, Antoine MOREAU, Clément BOREL et Pierre-Antoine GERVAIS.

Je remercie le Pays de Montbéliard Agglomération (PMA) d'avoir financé cette thèse et de son implication dans mon accueil. Ils ont été très attentifs aux étudiants étrangers, dont je faisais partie, et ce dès nos premiers jours passés en France dans le cadre du programme d'échange LEONARDO entre l'Université d'Oradea (Roumanie), mon université d'origine, et l'Université de Franche-Comté, mon université d'accueil. Tout particulièrement je remercie M. Jean-François KLOPFENSTEIN d'avoir facilité mon intégration dans le Pays de Montbéliard.

Je tiens également à remercier ma famille qui m'a soutenu malgré la distance, ainsi que mes amis de Montbéliard pour leur convivialité. Leurs encouragements m'ont été d'un grand réconfort.

Je ne pourrais pas terminer sans remercier Lidia, ma chère compagne de vie, d'avoir toujours été à mes côtés et d'avoir partagé tous les moments, les bons comme les moins bons, que j'ai connus pendant cette période.

Table des matières

Introduction.....	1
1.1. Problématique et hypothèses de recherche.....	2
1.2. Contributions de la thèse	3
1.3. Contexte de la thèse.....	5
1.4. Plan du mémoire.....	6
Chapitre 1. Apprentissage. Évolution et conceptions	9
1.1. Théories d'apprentissage – qu'est-ce qu'il faut apprendre ?.....	10
1.1.1. <i>Béhaviorisme – observation externe de l'individu</i>	11
1.1.2. <i>Cognitivisme et constructivisme – facteurs internes à l'individu</i>	13
1.1.3. <i>Socioconstructivisme – collaboration et contexte social</i>	16
1.1.4. <i>Connectivisme – sens et connaissance dans la construction des liens</i>	18
1.2. Innovation dans l'apprentissage ?.....	21
1.2.1. <i>Environnement et expérience d'apprentissage</i>	23
1.2.2. <i>Apprentissage actif</i>	29
1.2.3. <i>Apprentissage personnalisé</i>	35
1.2.4. <i>Apprentissage à travers les technologies de l'information et de la communication</i>	40
1.3. Savoirs et connaissances à l'ère du numérique.....	45
1.3.1. <i>Données, informations et connaissances</i>	49
1.3.2. <i>Représentation et ingénierie des connaissances</i>	54
1.3.3. <i>Objet pédagogique</i>	59
Chapitre 2. Systèmes et technologies de l'information et de la communication pour l'enseignement	67
2.1. Systèmes et systémique.....	67
2.1.1. <i>Système : boîte noire ou ensemble d'éléments en relations</i>	69
2.1.2. <i>Approche systémique : système et modélisation</i>	74
2.1.3. <i>Typologie des systèmes. Système actif et expert</i>	77
2.2. Systèmes d'information, de communication et d'apprentissage	83
2.2.1. <i>Information et communication</i>	83
2.2.2. <i>Systèmes et technologies d'information et de communication : l'Internet et le Web</i>	88
2.2.3. <i>Métadonnée : description des ressources d'apprentissage</i>	94
2.3. Exemples des systèmes d'apprentissage basés sur le Web.....	104
2.3.1. <i>Contenu et ressources d'apprentissage dans les SGA</i>	107
2.3.2. <i>Caractéristiques des systèmes d'apprentissage actuels</i>	109

Chapitre 3. Le Web sémantique et la représentation de l'information	115
3.1. L'utopie d'un savoir universel.....	115
3.1.1. Mundaneum : « savoir au bout des doigts ».....	117
3.1.2. Memex : mémoire associative.....	119
3.1.3. Xanadu : implémentation de l'hypertexte.....	121
3.2. Le Web et la sémantique	123
3.2.1. Comment en-est on arrivé là ?.....	125
3.2.2. Sens et signification dans le Web.....	135
3.2.3. Le Web sémantique.....	140
3.3. Représentation de l'information dans le Web sémantique.....	145
3.3.1. Identification des ressources dans le Web sémantique.....	147
3.3.2. Modèle RDF (Resource Description Framework).....	153
3.3.3. Le schéma RDF. L'inférence dans RDF(S).....	157
3.4. Langage d'interrogation et stockage du RDF.....	162
3.4.1. Sérialisation du RDF : RDF/XML, Turtle.....	162
3.4.2. Le langage d'interrogation SPARQL.....	165
3.4.3. Le stockage des données RDF : le Triplestore.....	167
Chapitre 4. Outils sémantiques pour la modélisation des connaissances	169
4.1. Ontologies et représentation des connaissances.....	173
4.1.1. Ontologie : de la philosophie à l'informatique.....	174
4.1.2. Le langage d'ontologie pour le Web (OWL).....	176
4.1.3. Logiques de description. Profils d'OWL.....	183
4.1.4. La sémantique du modèle théorique OWL.....	190
4.1.5. Règles et inférences dans le Web sémantique.....	193
4.2. Ingénierie ontologique	196
4.2.1. Conception et construction d'ontologies	197
4.2.2. Types d'ontologies.....	199
4.2.3. Approche modulaire.....	200
4.2.4. Exemples d'ontologies et vocabulaires utilisés sur le Web.....	201
4.3. Le Web sémantique pour les systèmes d'apprentissage	203
Chapitre 5. Modélisation du Système Actif et Sémantique d'Apprentissage (SASA)	207
5.1. Identification des objets dans le cadre du système d'apprentissage	208
5.2. Modélisation des objets participant à l'apprentissage	211
5.2.1. Modélisation de l'apprenant.....	211
5.2.2. Modélisation de l'objet pédagogique	216
5.2.3. Modélisation de l'objectif d'apprentissage	221

5.2.4.	<i>Modélisation de l'objet d'évaluation</i>	225
5.2.5.	<i>Modélisation de l'objet d'annotation</i>	228
5.2.6.	<i>Modélisation du cadre d'enseignement</i>	229
5.3.	Ontologies des domaines dans le système d'apprentissage	230
5.3.1.	<i>Ontologie de l'apprenant</i>	231
5.3.2.	<i>Ontologie de l'objet pédagogique</i>	232
5.3.3.	<i>Ontologie de l'objectif d'apprentissage</i>	234
5.3.4.	<i>Ontologie de l'objet d'évaluation</i>	235
5.3.5.	<i>Ontologie de l'objet d'annotation</i>	236
5.3.6.	<i>Ontologie du cadre d'enseignement</i>	236
5.4.	Ontologies au niveau d'application du SASA	237
5.4.1.	<i>Ontologie de liaison</i>	237
5.4.2.	<i>Ontologie d'application</i>	238
Chapitre 6. Implémentation du SASA		241
6.1.	Organisation du système	243
6.1.1.	<i>L'architecture générale du SASA</i>	243
6.1.2.	<i>Les modules du SASA</i>	245
6.1.3.	<i>Le noyau sémantique : un agent intelligent pour l'apprentissage</i>	246
6.2.	SASA : personnalisation du parcours d'apprentissage	247
6.2.1.	<i>Scénarios pour la personnalisation du parcours d'apprentissage</i>	249
6.2.2.	<i>Système réactif : adaptation basée sur les résultats de l'apprenant</i>	251
6.2.3.	<i>Système proactif : adaptation basée sur des préférences prédéfinies</i>	253
6.2.4.	<i>Règles et inférences dans le SASA</i>	258
6.3.	Agrégation des connaissances dans le SASA	261
6.3.1.	<i>Acquisition d'informations sur les apprenants</i>	261
6.3.2.	<i>Agrégation des ressources d'apprentissage externes</i>	262
6.4.	Développement du SASA	265
6.4.1.	<i>Développement du modèle et de la base de connaissance de SASA</i>	265
6.4.2.	<i>Développement de la logique du SASA : les contrôleurs et les interfaces</i>	267
Chapitre 7. Évaluation du SASA		269
7.1.	Évaluation du modèle du SASA	269
7.2.	Tests et simulations du noyau sémantique	272
Conclusions et perspectives		277
Liste des publications		281
Bibliographie		283

Annexes	311
Liste des figures	341
Liste des tableaux	345

Introduction

*« ...depuis le commencement, tout notre discours a été
une recherche de la connaissance »
Platon Théétète 196d*

La connaissance, cette « opinion droite pourvue de raison »¹ est l'élément principal du développement des sociétés humaines. Aujourd'hui, comme l'affirme Edgar Morin, nous vivons « l'époque sans doute la plus exaltante pour le progrès de la connaissance, la fécondité des découvertes, l'élucidation des problèmes, nous nous rendons difficilement compte que nos gains inouïs de connaissance se paient en gains inouïs d'ignorance »² L'apprentissage humain prend un rôle central dans la préservation, la transmission et l'enrichissement de la connaissance au fil des années. Le progrès scientifique apporte ses facilités dans ce processus d'apprentissage et de transfert de ce trésor de l'humanité. Comme l'invention de l'imprimerie a fait exploser la quantité de l'information disponible à travers des livres et du texte imprimé, de nos jours ce sont les technologies numériques de l'information et de la communication qui repoussent les limites d'accès à l'information. Les barrières de l'espace et du temps sont ainsi retracées. L'insertion ubiquitaire de ces systèmes dans la vie courante transforme la société dans ses fonctions essentielles. Le gouvernement est devenu *e-gouvernement*, le commerce *e-commerce*, l'apprentissage est devenu *e-learning*. De plus, la dynamique du changement de la société oblige l'évolution de l'apprentissage, en devenant un processus continu et nécessaire tout au long de la vie. *L'Internet* définit les nouvelles règles d'accès à l'information et fournit un nouvel espace d'existence. La société dans son ensemble (e.g. gouvernement, commerce) et plus particulièrement l'apprentissage sont soumis à une réorganisation afin de pouvoir s'inscrire dans ce nouvel espace.

¹ Platon, *Théétète*, 201d

² (Morin, 1992, p. 1180)

1.1. Problématique et hypothèses de recherche

Si après la révolution de l'imprimerie, les systèmes de classification de livres (e.g. Dewey) ont intégré des aspects organisationnels du savoir, dans les systèmes basés sur Internet, notamment *le Web*, l'organisation du savoir reste un défi. La grande quantité d'informations, le *Big Data*, conséquence de l'évolution des technologies de l'information et de la communication, impose des nouvelles méthodes et approches pour le traitement et l'organisation de ces données. Morin rappelle qu'« aujourd'hui, l'édifice du savoir contemporain s'élève comme une Tour de Babel, qui nous domine plus que nous ne la dominons » (Morin, 1992, p. 1185). Dans le domaine de l'apprentissage, plus sensible à l'organisation du savoir, cette problématique prend une dimension d'autant plus importante.

La vaste quantité de contenus d'apprentissage, sa diversification et son accessibilité font évoluer les approches d'apprentissage. Nous participons à la mondialisation de l'espace d'apprentissage et à l'enrichissement des contenus. L'apprentissage est accessible à travers les *systèmes d'information et de communication* qui permettent la diffusion massive³ de cours quels que soient la localisation géographique, le moment de la journée ou encore le dispositif d'accès utilisé. Les méthodes d'apprentissage et d'enseignement évoluent, et aux modèles classiques viennent s'ajouter de nouveaux paradigmes. Ainsi, au behaviorisme, au cognitivisme ou au constructivisme s'ajoute *le connectivisme*. Visant la facilitation de l'apprentissage humain, les systèmes d'information et de communication se sont spécialisés en *systèmes d'apprentissage*. Nous identifions dans *ces systèmes d'apprentissage* l'unité centrale assurant la convergence des plusieurs domaines qui permettent l'exercice de l'apprentissage.

Pour répondre aux défis d'organisation de l'information et afin d'améliorer la collaboration entre l'homme et la machine⁴, *le Web sémantique* propose un ensemble de technologies permettant d'enregistrer des descriptions pour chaque ressource dans un format interprétable par la machine. Plus concrètement, le Web sémantique propose des

³ *Les cours en ligne ouverts et massifs* (MOOC – Massive Open Online Cours).

⁴ « The Semantic Web is not a separate Web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation. » (Berners-Lee, Hendler, et Lassila, 2001).

outils et langages pour l'enregistrement de métadonnées accompagnant les données effectives. Puisque ces métadonnées peuvent être manipulées par les machines, cela a pour effet d'augmenter le niveau d'efficacité atteint par celles-ci dans le traitement et la « compréhension » des données qu'elles manipulent. L'approche du Web sémantique définit la *trame de description de ressources (Resource Description Framework - RDF)*, le *langage RDF*, les *vocabulaires RDF et schéma RDF (RDFS)*, ainsi que le *langage d'ontologie pour le Web (Web Ontology Language - OWL)*. Ces éléments permettent d'enregistrer d'une manière formelle la sémantique associée aux entités ou ressources manipulées par le système considéré.

En tant que sous-catégorie spécifique des systèmes d'information et de communication, les systèmes d'apprentissages héritent de leurs avantages et inconvénients. Ainsi les systèmes d'apprentissage basés sur le Web bénéficient de la structure et des évolutions technologies du Web. Notre hypothèse de recherche est que les technologies du Web sémantique peuvent améliorer l'apprentissage. Plus particulièrement, ces technologies peuvent améliorer les capacités des systèmes d'apprentissage dans leurs tâches de soutien de l'apprentissage humain. Notre attention est donc centrée sur l'apport et les possibilités d'intégration des technologies du Web sémantique dans la conception d'un système d'apprentissage. Un tel système d'apprentissage serait un système expert, capable de prendre en charge des tâches spécifiques, comme par exemple la recommandation de contenu adapté au besoin de l'apprenant.

1.2. Contributions de la thèse

En suivant une approche systémique, nous avons utilisé les technologies du Web sémantique pour la réalisation d'une plateforme d'apprentissage capable d'enrichir et de personnaliser l'expérience de l'apprenant. Les résultats de nos travaux sont concrétisés dans la proposition d'un prototype pour un *Système Actif et Sémantique d'Apprentissage (SASA)*. SASA désigne à la fois un *système conceptuel* dans la modélisation des entités participant à l'apprentissage et un *système informatique* dans l'implémentation qui permet l'interprétation et la gestion du système conceptuel.

La modélisation à la base de notre prototype SASA est structurée en six ontologies distinctes. Ces ontologies sont organisées autour d'objets centraux et permettent la construction de modules spécifiques correspondant à chaque ontologie.

1. L'ontologie de l'apprenant contient les caractéristiques de l'apprenant déclinées sur quatre dimensions principales. Ainsi, le profil d'apprenant est composé de : son historique d'apprentissage, ses objectifs d'apprentissage, son style d'apprentissage et ses préférences d'apprentissage.
2. L'ontologie de l'objet pédagogique est une représentation des ressources d'apprentissage, et concerne la description du savoir contenu dans l'objet pédagogique. Le modèle représenté dans cette ontologie est basé sur les caractéristiques des objets pédagogiques proposées dans le standard *Learning Object Metadata (LOM)*.
3. L'ontologie de l'objectif d'apprentissage contient les descriptions des objectifs d'apprentissage, dans le but de quantifier le niveau d'apprentissage atteint par l'apprenant dans le cadre du système. Ainsi, le portfolio des objectifs d'apprentissage acquis reflète le niveau d'évolution de l'apprenant dans son parcours d'apprentissage.
4. L'ontologie de l'objet d'évaluation comprend la description des objets utilisés dans l'évaluation de l'apprenant. Cette ontologie spécifie les principaux éléments permettant l'évaluation de l'apprentissage de l'apprenant.
5. L'ontologie de l'objet d'annotation contient les représentations des différents types d'annotation effectuées par les apprenants. Ces annotations constituent des indices clés pour l'évaluation et la classification des ressources d'apprentissage.
6. L'ontologie du cadre d'enseignement permet de décrire les éléments contextuels de l'apprentissage. Elle fournit la modélisation du contexte de l'apprentissage.

Pour l'exploitation de ces modélisations, dans le cadre de notre prototype SASA, nous proposons le *noyau sémantique*. Ce noyau sémantique regroupe les ontologies déclarées, tout en facilitant l'intégration des ressources d'apprentissage ou d'informations existantes sur le Web.

Construit sur la base de la structure offerte par les technologies sémantiques et intégrant une modélisation riche des entités, le système d'apprentissage bénéficie des

possibilités d'inférence sur l'ensemble des modèles représentées. La définition des règles d'inférence spécifiques aux modèles représentés permet l'enrichissement des relations entre les entités et leur adaptation pour une exploitation spécifique au contexte d'apprentissage.

Par rapport à l'apprenant, notre système présente deux scénarios d'interaction :

1. Le système est réactif lorsqu'il répond aux différentes actions de l'apprenant (e.g. avancement de l'apprenant dans le processus d'apprentissage) ;
2. Le système est proactif lorsqu'il anticipe les prochaines actions et, avant de recevoir des stimuli extérieurs de la part d'apprenant, il modifie son état, son organisation, etc.

Les comportements du SASA (e.g. actions, décisions, modifications de son organisation) impliquent la coordination de plusieurs éléments le composant : modélisation, générateurs et sources d'informations, règles de décisions, etc. L'intégration de ces éléments, technologies et comportements dans le SASA nous permet de fournir un outil capable d'influer le parcours d'apprentissage de l'apprenant. Les règles au niveau du noyau sémantique permettent d'établir les conditions d'adaptation du contenu d'apprentissage. Le contexte de l'apprenant, ses caractéristiques, préférences ou besoins peuvent influencer la construction de son parcours d'apprentissage.

1.3. Contexte de la thèse

Le sujet de cette thèse se situe dans l'axe de recherche *plateformes sémantiques d'apprentissage* de l'équipe *Objets et Usages Numériques* (OUN) du laboratoire *Édition, Langues, Littératures, Informatique, Arts, Didactiques, Discours* (ELLIADD). Il constitue une continuation des travaux dans l'application des technologies sémantiques aux systèmes d'apprentissage. À la suite de la thèse « *La modélisation d'objet pédagogiques pour une plateforme sémantique d'apprentissage* » soutenue par Radu BALOG-CRIȘAN en décembre 2011, cette thèse propose un approfondissement en intégrant des modélisations supplémentaires des objets participant à l'apprentissage, ainsi que le développement d'un *Système Actif et Sémantique d'Apprentissage*. Ce travail a été financé par la communauté d'agglomération du pays de Montbéliard (*Pays de Montbéliard Agglomération - PMA*).

1.4. Plan du mémoire

Ce mémoire est structuré en sept chapitres.

Chapitre 1. Apprentissage. Évolution et conceptions

Dans ce premier chapitre, nous rappelons quelques définitions de l'apprentissage, puis nous évoquons les principales théories et concepts fondamentaux utilisés dans ce domaine; nous nous interrogeons ensuite sur les changements imposés par l'arrivée des technologies de l'information et de la communication (TIC).

Chapitre 2. Systèmes et technologies de l'information et de la communication pour l'enseignement

Ce chapitre rappelle ce qu'est un système et les notions de base de la pensée systémique. Considérant le Web comme un système de base d'information et de communication, nous étudions le développement des systèmes reposant sur le Web pour la création et la gestion de contenus, ainsi que pour l'apprentissage.

Chapitre 3. Le Web sémantique et la représentation de l'information

Nous présentons, au travers de ce chapitre, le Web sémantique et les principaux outils pour la représentation de l'information.

Chapitre 4. Outils sémantiques pour la modélisation de la connaissance

Nous étudions ici les outils de modélisation proposés dans le cadre du Web sémantique. Nous présentons les langages de modélisation proposés par le W3C et nous donnons des exemples d'utilisation.

Chapitre 5. Modélisation du Système Actif et Sémantique d'Apprentissage (SASA)

Dans ce chapitre nous présentons la modélisation du SASA. Nous nous concentrons sur l'aspect conceptuel du SASA et présentons les modèles et les ontologies associés aux entités identifiées dans le processus d'apprentissage.

Chapitre 6. Implémentation du SASA

Ce chapitre traite de l'implémentation et du développement du système SASA. Nous présentons l'architecture du SASA, son noyau sémantique, ainsi que les fonctionnalités du noyau permettant d'adapter le parcours d'apprentissage à l'apprenant.

Chapitre 7. Évaluation du SASA

Dans ce chapitre nous présentons l'évaluation du SASA dans quelques scénarios d'usage, ainsi que les points d'amélioration identifiés. Enfin nous concluons en rappelant les principaux points de contribution de ces travaux.

Chapitre 1. Apprentissage. Évolution et conceptions

*« L'apprentissage n'est pas un produit de la
scolarisation, mais la tentative permanente de
l'acquérir. »
A. Einstein*

Vu l'infinie complexité du cerveau humain et de ses modes opératoires, comprendre et définir le processus d'apprentissage chez l'être humain n'est pas trivial. Même si plusieurs auteurs se sont efforcés de répondre à la question « qu'est-ce qu'apprendre ? », les évolutions sociales et les innovations technologiques impliquent un renouvellement constat de la problématique de l'apprentissage.

Dans ce premier chapitre, nous rappelons d'abord quelques définitions de l'apprentissage ; nous évoquons ensuite les principales théories et les concepts fondamentaux utilisés dans ce domaine ; nous nous interrogeons enfin sur les changements imposés par les technologies de l'information et de la communication (TIC) à l'apprentissage.

Le dictionnaire *Le Petit Robert* définit *l'apprentissage* comme suit : « 1. Le fait d'apprendre un métier manuel ou technique ; l'ensemble des activités de l'apprenti. 2. Les premières leçons, les premiers essais. 3. Modifications durables du comportement d'un sujet (humain ou animal) grâce à des expériences répétées. Processus d'acquisition des automatismes sensorimoteurs et psychiques. » (Le Robert, Rey-Debove, et Rey, 2010, p. 122)

Différentes conceptions de l'apprentissage enrichissent et éclairent la compréhension du terme. Dans le *Dictionnaire des concepts clés en pédagogie*, une première définition affirme que, apprendre c'est « modifier son comportement et ses représentations [...] traiter l'information de manière à pouvoir la mémoriser et la réutiliser ultérieurement [...] » (Raynal et Rieunier, 2010, p. 34).

D'autres définitions privilégient le transfert de connaissances : « L'apprentissage est un processus systématique et intentionnellement orienté vers l'acquisition de certains savoirs, savoir-faire, savoir-être et savoir-devenir » (J.-M. D. Ketele, Chastrette, et Cros, 1989, p. 216). Et le transfert de connaissances est « la pierre philosophale de l'enseignant » (Mendelsohn, 1994). Pour E. Morin, apprendre « ce n'est pas seulement reconnaître ce qui, d'une façon virtuelle, était déjà connu. Ce n'est pas seulement transformer de l'inconnu en connaissance. C'est la conjonction de la reconnaissance et de la découverte. Apprendre comporte l'union du connu et de l'inconnu » (Morin, 1992, p. 61).

1.1. Théories d'apprentissage – qu'est-ce qu'il faut apprendre ?

Afin d'obtenir une image plus claire et complète de ce qu'est l'apprentissage, nous allons évoquer les théories les plus pertinentes dans ce domaine. Des théories de l'apprentissage ont été déterminées, largement à partir de courants philosophiques – idéalistes ou empiristes – fondant leur analyse des comportements humains sur la séparation de la matière et de l'esprit, mais aussi à partir des sciences de la nature et les sciences de l'esprit. À partir de ces diverses influences, M. Develay (1992, p. 106-109), propose une classification des théories et conceptions de l'apprentissage en trois grandes familles.

Tableau 1-1. Classification de M. Develay (ibid.), adaptée après (Rézeau, 2001, p. 118)

L'idéalisme (Rationalisme)	L'empirisme
La connaissance nous vient d'idées innées dont nous sommes porteurs	La connaissance nous vient du monde extérieur par l'expérience
Platon et Socrate	Aristote
Saint Augustin	La scholastique
Descartes, Grammaire de Port-Royal	Thomas D'Aquin
Les Idéalistes Allemands	Les empiristes Anglais
La gestalt (psychologie de la forme)	Le Behaviorisme
Le constructivisme	
La connaissance nous vient du monde extérieur par l'expérience qui modifie les schèmes innés (Piaget).	

Les deux premières familles de théories, l'idéalisme centré sur le sujet et l'empirisme centré sur l'objet, renvoient à un type de pédagogie et à une conception du rôle de l'enseignant opposés⁵. La troisième théorie – le constructivisme – est présentée comme « la théorie actuelle *faisant la synthèse* des deux autres conceptions » (ibid.). Le constructivisme se rattache en effet autant à l'idéalisme, par les schémas innés dont il nous estime porteurs, qu'à l'empirisme qui voit dans l'expérience du monde extérieur la source de nos connaissances (v. Tableau 1-1).

J. Lecomte (1993, p. 11-16) propose une classification des théories relatives à l'apprentissage qui isole l'individu de son environnement. Il présente une classification en trois catégories. Les facteurs externes à l'apprenant conduisent aux théories behavioristes, les facteurs internes mènent aux théories cognitivistes et l'interaction entre ces deux types de facteurs fait prévaloir les théories socio-constructivistes.

1.1.1. Behaviorisme – observation externe de l'individu

Le behaviorisme est apparu comme une nécessité pour l'étude de la « vie mentale ». Les psychologues ont cherché à créer une psychologie objective fondée sur des faits observables. Selon les behavioristes, la psychologie doit être une science du comportement observé. Le fonctionnement humain doit être expliqué à partir de données observables, de l'étude du comportement, sans recours à l'introspection. J.B. Watson (1878-1958) publie en 1913 un article considéré comme le manifeste du behaviorisme⁶. Dans cet article, il affirme que la psychologie ne doit pas être la science de la vie mentale, mais la science du comportement (Watson, 1913). Un autre psychologue – J.B. Thorndike (1874 - 1949) – s'était prononcé, avant les travaux de J.B. Watson, en faveur d'une psychologie élaborée à partir de l'étude du comportement observable. Il formule ainsi la loi de l'effet : « Tout comportement qui conduit à un état satisfaisant de l'organisme a tendance à se reproduire dans la même situation » (Raynal et Rieunier, 2010, p. 447).

⁵ De l'idéalisme résulte une théorie pédagogique de l'endogène, de l'empirisme une théorie pédagogique de l'exogène. Le constructivisme renvoie à une théorie de « l'exogène » (Rézeau, 2001, p. 118).

⁶ Behavior – le terme anglais pour « comportement ».

Le behaviorisme s'intéresse au comportement de l'individu selon le modèle *stimulus – réponse*. Cette approche analyse les phénomènes psychologiques d'un individu à partir des éléments observables : les stimuli à l'entrée et les comportements en sortie. Dans sa tentative de mise au point d'une méthode d'exploration objective de l'activité psychique, I. Pavlov (1849 - 1936) introduit en 1889, au terme de ses expériences, la notion de *conditionnement répondant*⁷.

Lors de son étude du comportement animal dans le cadre d'expériences pratiques⁸, B.F. Skinner (1904 - 1990) introduit la procédure du *conditionnement opérant*⁹. Il pose ensuite les bases de *l'enseignement programmé*, qui met l'accent sur la nécessité de fixer des objectifs mesurables et autorise le développement des objectifs d'apprentissage, des compétences et des aptitudes spécifiques. (Roxin, 2003, p. 17)

Même si B.F. Skinner n'utilise pas le terme d'interactivité, le concept est présent dans « l'échange continu entre le programme et l'élève » (Skinner et Richelle, 1969, p. 48). En affirmant que les apprenants ne sont pas les bénéficiaires passifs de l'information distribuée par les enseignants, B.F. Skinner favorise l'apprentissage actif. Pour lui, enseigner c'est « inviter l'élève à s'engager dans des nouvelles formes de comportement » (ibid., p. 43). Dans les principes élaborés par B.F. Skinner, nous découvrons un environnement qui pousse l'apprenant dans une nouvelle expérience d'apprentissage. Cette expérience est caractérisée par le niveau élevé d'interactivité, le renforcement positif et la vérification immédiate par chacun de ses résultats. Skinner est unanimement considéré comme le grand théoricien du béhaviorisme.

D'autres points de vue, toujours dans la perspective behavioriste, mettent en avant *l'expérience d'apprentissage* (les échanges *stimulus - réponse* entre l'apprenant et son environnement d'apprentissage) : « L'apprentissage est un processus ou un ensemble de processus qui sous-tend des modifications de comportements, à la suite de l'expérience ou du contact avec l'environnement » (Malcuit, Maurice, et Pomerleau, 1995).

⁷ Conditionnement classique, déterminant un réflexe.

⁸ Il enseigne à des pigeons à jouer au ping-pong (Raynal et Rieunier, 2010, p. 418).

⁹ Conditionnement instrumental, qui génère un comportement.

Dans la perspective béhavioriste, l'apprentissage se produit lorsqu'un apprenant acquiert, en réponse à un stimulus, un comportement nouveau ou modifié. Deux éléments distincts sont observables lors du parcours d'apprentissage. Le premier élément rassemble les causes qui provoquent l'apprentissage : « ensemble des activités », « expériences répétées ». Le deuxième élément constitue le résultat (l'effet) de l'apprentissage : « modification durable du comportement ».

En considérant exclusivement le comportement observable de l'individu (e.g. réponse) et comment l'environnement (e.g. stimulus) façonne ce comportement, le béhaviorisme sous-estime ce qui se passe dans la « boîte noire » du cerveau de l'apprenant. Le résultat est un apprentissage mécanique, par tâtonnements, d'un *savoir procédural*, ne s'appuyant pas sur la réflexion de l'individu. Les insuffisances majeures de ce type d'apprentissage se retrouvent dans les limites de l'adaptabilité face aux changements de l'environnement. Habitué à des caractéristiques et procédures spécifiques, l'individu sera rapidement désorienté à la moindre modification de son environnement.

1.1.2. Cognitivisme et constructivisme – facteurs internes à l'individu

Dans son opposition au béhaviorisme, le *cognitivisme* s'intéresse à ce qu'il se passe à l'intérieur du cerveau de l'individu (« boîte noire »), entre les stimuli d'entrée et les réponses en sortie. Ce courant s'affirme dans les années 1960 et propose d'expliquer le comportement humain à partir de l'hypothèse des variables internes : motivation, attentes cognitives, schémas, structures (Raynal et Rieunier, 2010, p. 106).

Dans le *Dictionnaire actuel de l'éducation*, R. Legendre (1993, p. 205), définit le cognitivisme comme étant « l'ensemble des activités et des processus internes inséparables à l'acquisition des connaissances, à l'information, à la mémoire, à la pensée, à la créativité, à la perception, ainsi qu'à la compréhension et à la résolution de problèmes ».

Le courant cognitivisme s'associe à la psychologie cognitive, qui, selon A. Bertrand et P.H. Garnier (2005, p. 58), « s'intéresse à la manière dont l'esprit humain manipule et élabore mentalement des connaissances (images, représentation, symboles, schémas, etc.) pour produire des pensées et des actions sur le monde, sur les autres et sur soi-

même ». H. Gardner (1985, p. 6) définit la science cognitive comme « un effort contemporain sur une base empirique pour répondre à des questions épistémologiques de longue date en particulier celles qui s'intéressent à la nature de la connaissance, ses composantes, ses sources, son développement et son déploiement ».

Si, dans le béhaviorisme, l'attention était dirigée vers le comportement observable de l'individu, le cognitivisme met en évidence et essaie d'expliquer la *cognition*. *Le Petit Robert* (2010, p. 462) définit la *cognition* comme « connaissance, processus par lequel un organisme acquiert la conscience des événements et objets de son environnement » et le *cognitif* comme ce « qui concerne la connaissance, les processus mentaux en œuvre dans son acquisition ». D'un point de vue plus pragmatique, la cognition peut se définir comme l'ensemble des fonctions mentales qui nous permettent de traiter l'information : mémoire, sens, perception, langage, etc. Les termes *systèmes cognitifs* et *processus cognitifs* vont désigner des systèmes de connaissances et processus de traitement de l'information utilisés par l'esprit humain pour s'adapter à son environnement.

La formule : « cognitivisme = calcul + représentation » (Raynal et Rieunier, 2010, p. 107) met en évidence les deux composantes de la théorie, qui nous indiquent ainsi les deux directions de recherche du cognitivisme : *cognitivisme computationnel*¹⁰ et le *cognitivisme structural*¹¹. « J. Piaget et les *psychologues de la forme*¹² peuvent être considérés comme les représentants d'un *cognitivisme structural* antérieur au *cognitivisme computationnel* des années 1970 » (ibid., p. 350).

J. Piaget (1896-1980) s'intéresse à la construction de connaissance et à l'épistémologie génétique. Il affirme que « les connaissances ne sont pas transmises par quelqu'un qui "sait" vers quelqu'un qui "ne sait pas", elles ne viennent pas des sensations comme le prétendent les associationnistes, elles sont construites (constructivisme) par l'individu par l'intermédiaire des actions qu'il accomplit sur les "objets" » (ibid., p. 349). Le

¹⁰ Met en avant l'aspect computationnel de la cognition humaine, analogie *cerveau – ordinateur*.

¹¹ Met en avant la façon dont on *organise* l'information dans le cerveau.

¹² « Gestalt », signifie, en allemand, forme globale ou « forme organisée ». Khöler, Koffka (gestaltistes allemands) - gestaltisme (psychologie de la forme) – théorie selon laquelle les processus de la perception et de la représentation mentale traitent spontanément les phénomènes comme des ensembles structurés (des formes) et non comme une simple addition ou juxtaposition d'éléments.

constructivisme affirme que le développement d'un individu, et donc aussi l'apprentissage, est un processus permanent de construction et d'organisation des connaissances.

Dans le cognitivisme et le constructivisme, apprendre, c'est assimiler¹³ (« transformer, convertir en sa propre substance »), c'est comprendre¹⁴ (« savoir pourquoi ») mais c'est aussi adapter (« assimilation et accommodation ») : il existe donc deux types d'apprentissage. Le premier est un apprentissage par l'action, décrit par Piaget avec le cognitivisme structurel, le constructivisme et l'épistémologie génétique. Le deuxième type d'apprentissage est un apprentissage par la résolution de problèmes, tel que le décrivent le cognitivisme symbolique et la psychologie du traitement d'information¹⁵. (Roxin, 2003, p. 18) Le cognitivisme et le constructivisme affirment que la connaissance n'est pas passivement reçue de l'environnement, mais elle est *activement* construite par l'apprenant. On apprend en pensant ou en réfléchissant (*learning by thinking*) et l'apprentissage se produit quand, lors d'un processus cognitif, la connaissance est transférée de la mémoire à court terme à la mémoire à long terme. Cet apprentissage s'ajoute à l'*apprentissage en pratiquant (learning by doing)*, cher à J. Dewey (1859-1952), l'un des premiers chercheurs en enseignement. Sa doctrine a influencé É. Claparède et C. Freinet, qui affirmait : « C'est vraiment en forgeant qu'on devient forgeron, c'est en écrivant qu'on apprend à écrire » (Freinet, 1954; cité par Giordan, 1998, p. 113). Davantage de détails sur l'apprentissage par l'action seront présentés dans le paragraphe 1.2.2. Apprentissage actif.

¹³ « [...] pour accommoder son activité aux propriétés des choses, l'enfant a besoin de les assimiler à elle et de se les incorporer véritablement » (Piaget, 1969, p. 209).

¹⁴ « Réussir, c'est comprendre en action une situation donnée à un degré suffisant pour atteindre les buts proposés, et comprendre c'est réussir à dominer en pensée les mêmes situations jusqu'à pouvoir résoudre les problèmes qu'elles posent quant au pourquoi et au comment des liaisons constatées et par ailleurs utilisées dans l'action [...] En un mot, comprendre consiste à dégager la raison des choses, tandis que réussir ne revient qu'à les utiliser avec succès, ce qui est certes une condition préalable puisqu'elle en arrive à un savoir qui précède l'action et peut se passer d'elle [...]. Le monde des "raisons" s'élargit sur les possibles et déborde ainsi le réel. » (Piaget, 1974; cité par Raynal et Rieunier, 2010, p. 114).

¹⁵ Simon Newell

Comme psychologue de l'éducation, J. Bruner s'oriente dans un premier temps vers les approches cognitivistes et constructivistes. « Il s'inspire du modèle de l'équilibration¹⁶ de Piaget pour proposer un modèle d'acquisition de connaissances en spirale¹⁷ » (Raynal et Rieunier, 2010, p. 88). Il souligne l'importance de la médiation et de l'interaction dans l'apprentissage de l'enfant et, avec la notion d'interaction de l'apprenant, il rejoint la thèse de Vygotski (v. §1.1.3). Selon Bruner (1966, 1986), l'individu construit du sens en apprenant.

Au changement du comportement s'ajoute alors un changement de la pensée, de la structure cognitive et de la compréhension des schèmes cognitifs. Les définitions de l'apprentissage s'orientent vers l'acquisition de l'information, le traitement et la compréhension de cette information : « L'apprentissage est le processus continu de transformation de l'information et de l'expérience dans la connaissance, les compétences, les comportements et les attitudes » (Cobb, 2009).

Dans l'approche cognitiviste, l'apprentissage c'est comprendre, acquérir de nouvelles informations et intégrer des nouveaux schémas à sa structure cognitive. La modification de ses représentations ou de ses comportements est l'effet de l'apprentissage (Raynal et Rieunier, 2010, p. 34).

1.1.3. Socioconstructivisme – collaboration et contexte social

En affirmant que le développement cognitif est directement issu de l'interaction sociale et de rapports sociaux, Vygotski (1896-1934) propose une théorie *historico-culturelle du cognitif* (Vygotski, 1997). Considéré comme le représentant le plus célèbre du *socioconstructivisme*, Vygotski affirme que « l'enfant est d'abord un être social et le développement de sa pensée, de son langage et ses fonctions psychiques, est le fruit

¹⁶ La structure cognitive connaît des états dits « d'équilibre », c'est-à-dire des états de schèmes permettant de répondre à toutes les situations déjà rencontrées. Les nouvelles situations provoquent le déséquilibre. Ce déséquilibre conduit à une étape d'assimilation et d'accommodation. Cette étape entraîne la modification du schème ancien ou la création d'un nouveau schème. Un nouvel état d'équilibre en résulte. Les états de déséquilibre rappellent le *conflit cognitif*. [JE NE COMPRENDS PAS ; Difficile à interpréter]

¹⁷ Dès la petite enfance, les notions enseignées doivent être vraies, verbalisées correctement et adaptées à la structure cognitive de l'enfant.

d'une interaction permanente avec le monde des adultes » (Raynal et Rieunier, 2010, p. 469).

Le socioconstructivisme s'inscrit dans le domaine de la psychologie sociale du développement de l'individu. Les théories de la construction sociale de l'intelligence considèrent l'interaction sociale comme l'un des éléments clés du développement intellectuel de l'individu. Avec la théorie du conflit sociocognitif, le socioconstructivisme formule l'hypothèse de développement psychosociale de l'individu. L'hypothèse affirme que l'apprentissage, en tant que *reconstruction cognitive*, vient de la médiation et de la réconciliation d'une déstabilisation favorable (e.g. conflit sociocognitif) provoquée par l'interaction sociale ou par le travail en équipe.

W. Doise et G. Mugny (1979), A. Brossard et P. Clermont (1985) « ont découvert que nombre d'enfants qui ne parvenaient pas au départ à réussir les tâches de conservation ont significativement amélioré leurs performances en discutant avec un autre enfant, même si ce partenaire n'en savait guère plus long qu'eux » (Tomasello, 2004, p. 176).

Les activités d'apprentissage doivent diriger et entraîner l'apprenant vers les interactions sociales. L'apprenant est encouragé à s'impliquer dans les interactions avec ses collègues, dans les discussions, à échanger des idées et à travailler ensemble.

Selon le socioconstructivisme, la construction de connaissances s'opère dans le dialogue avec les autres, dans la confrontation de ses idées, de ses représentations, à celles d'autrui. L'apprentissage se produit à l'aide des conflits sociocognitifs, dans la zone proximale de développement¹⁸, en collaboration et en interaction avec les autres participants à l'apprentissage.

¹⁸ La théorie de la zone proximale de développement suggère que les enfants sont aptes à mieux apprendre les problèmes et à s'améliorer davantage autour d'un enfant plus expérimenté, d'un parent ou d'un enseignant, plutôt que d'un enfant ayant leur niveau cognitif. Cela encourage donc l'apprentissage en milieu scolaire à ce stade de la vie. La zone proximale de développement augmente nettement le potentiel d'un enfant à apprendre plus efficacement.

1.1.4. Connectivisme – sens et connaissance dans la construction des liens

Le béhaviorisme, le cognitivisme et le constructivisme se sont développés dans un contexte où la technologie n'avait pas atteint le niveau d'imprégnation de la vie quotidienne que l'on connaît aujourd'hui. Le changement social, l'évolution technologique, l'augmentation et la disponibilité de la quantité d'information imposent des changements dans les pratiques d'apprentissage.

Le connectivisme, développé par G. Siemens et S. Downes, se définit comme « une théorie d'apprentissage pour l'ère numérique » (« About — Connectivism », 2005; Siemens, 2005).

L'augmentation de la quantité d'information implique d'autres méthodes d'apprentissage. Il ne s'agit plus d'apprendre des informations « par cœur », mais de déterminer les liens entre ces informations. Cité par J. Tardif, M. Develay¹⁹ affirme que « le sens vient des liens construits entre les savoirs et non pas de leur empilement [...] apprendre, ce n'est pas amasser, mais c'est relier des notions pour en construire d'autres plus abstraites. (Tardif et Presseau, 1998, p. 47).

Pour P. Lévy²⁰, la compréhension est « l'association d'un item d'information avec un schéma préétabli » (Lévy, 1990, p. 92). Une telle association « est également un moyen de stocker l'information en mémoire à long terme, en vue d'une réactivation ultérieure. L'établissement de ces interconnexions est non seulement un moyen de stockage efficace, mais il permet également un meilleur accès aux connaissances ainsi constituées comme l'écrit R.J. Mislevy²¹ en 1990, rapporté par Landsheere: "[...] comparés aux novices, les experts dominent plus de faits et établissent plus d'interconnexions ou de relations entre eux. Ces interconnexions permettent de surmonter les limitations de la mémoire à court terme. Alors que le novice ne peut travailler qu'avec au maximum sept éléments simples, l'expert travaille avec sept constellations incarnant une multitude de

¹⁹ Michel Develay, actuellement professeur en Sciences de l'éducation à l'université Lumière-Lyon 2 : <http://www.ac-grenoble.fr/occe26/printemps/develay/>

²⁰ Pierre A. Lévy, Professeur en SIC à l'université d'Ottawa : <http://www.communication.uottawa.ca/eng/faculty/levy.html>

²¹ Robjert J. Mislevy : <http://www.education.umd.edu/EDMS/mislevy/>

relations entre de nombreux éléments” » (De Landsheere, 1992, p. 57; cité par Rézeau, 2001, p. 35).

Constitué de nœuds (contenu) et de liens (connexions), un réseau d'apprentissage ne peut fonctionner que si l'ensemble de ces composants est fonctionnel. Dans un tel système, ce sont les liens qui sont les éléments les plus importants car, à travers les connexions, nous continuons à apprendre et à assimiler du contenu.

L'apprentissage (défini comme connaissance applicable) « peut résider à l'extérieur de nous-mêmes (dans des organisations et des bases de données) et se concentrer sur la connexion de l'ensemble d'information spécialisée, et ces liens qui nous permettent d'apprendre d'avantage sont plus importants que notre état actuel de connaissance [...] la connexion est plus importante que le contenu qui existe dans la connexion » (Siemens, 2005).

Le connectivisme se rapproche du socioconstructivisme puisque l'apprentissage est un processus de construction de connexions entre des nœuds spécialisés (sources d'information). Le *savoir-faire* et le *savoir-quoi* doivent se compléter par le *savoir-où*. En même temps, les connexions entre les nœuds ne sont pas faites au hasard. L'exercice de *la raison* sur les liens construits entre les nœuds rattache le connectivisme au cognitivisme.

Dans la Figure 1-1 nous présentons une image d'ensemble des principales théories et courants de pensée de l'apprentissage.

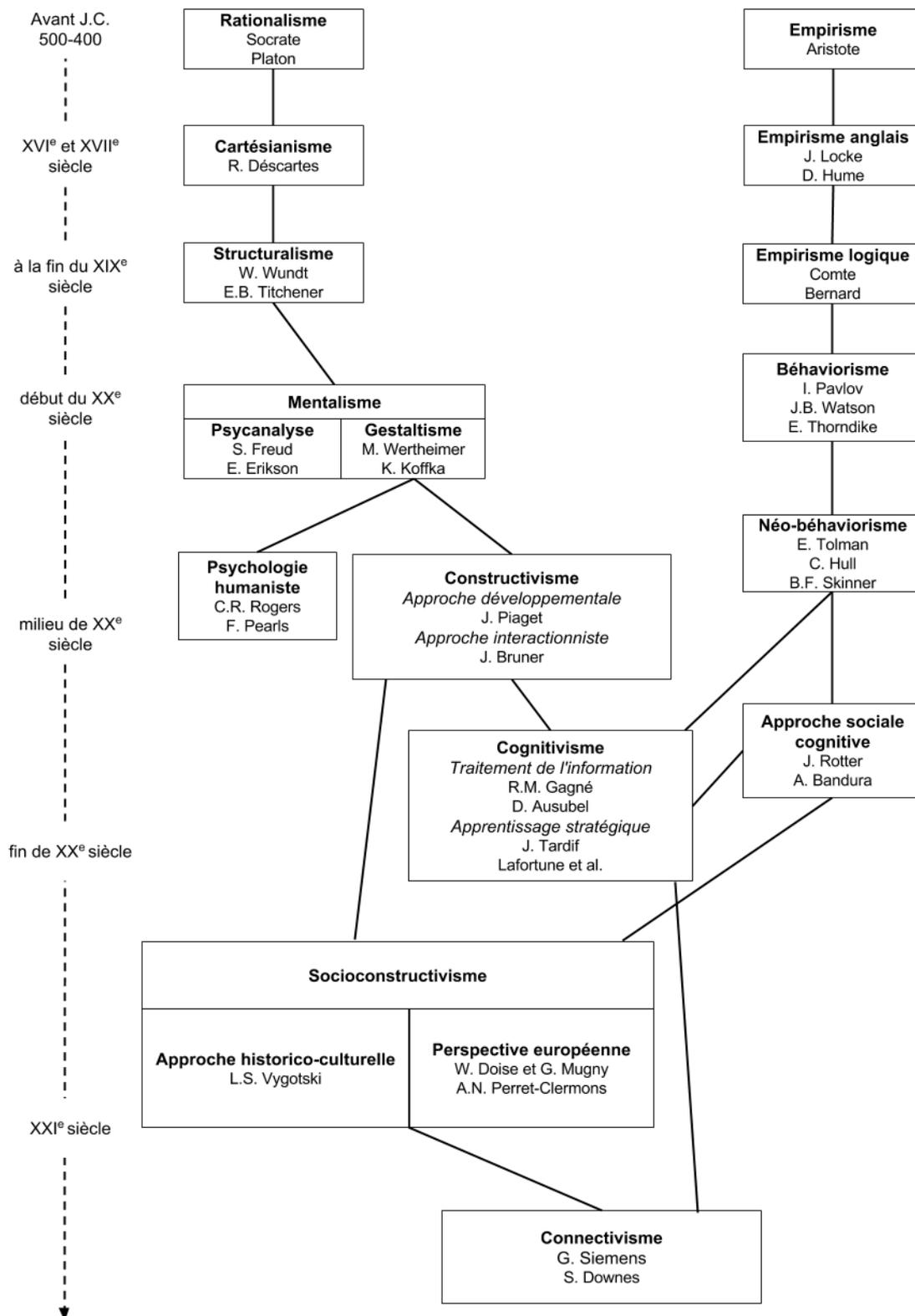


Figure 1-1. Ancrage historique et développement des courants de pensée de l'apprentissage (adapté selon Minier, 2003)

Après cette incursion dans les principaux courants et théories d'apprentissage, la question qui se pose est : quelle théorie, quel courant ou quel modèle devons-nous

privilegier pour l'apprentissage actuel ? Quels sont les éléments de ces théories qui doivent être exploités davantage pour un apprentissage capable de répondre aux besoins actuels d'une société de la connaissance, dans *l'époque post-PC*²² ? La réponse à ces questions est complexe. Les possibilités actuelles des technologies de l'information et de la communication sont, pour le système actuel d'éducation, une provocation à l'innovation.

1.2. Innovation dans l'apprentissage ?

Les avancées de la science et les évolutions technologiques ont, dit-on, aplati la Terre²³. Le développement des technologies de l'information et de la communication a changé notre rapport à l'information et à la connaissance. C. Gonzalez (2004) affirme que la technologie impose de nouvelles exigences aux personnes sur leur lieu de travail en mettant l'accent sur la formation et l'éducation. L'un des facteurs les plus persuasifs est le rétrécissement de la *demi-vie de la connaissance*. La « demi-vie de la connaissance » est le laps de temps situé entre le moment où la connaissance est acquise et le moment où elle devient obsolète. Selon *l'American Society of Formation and Documentation (ASTD)*, cité par G. Siemens (2005), « la moitié de ce qui est connu aujourd'hui n'était pas connu il y a 10 ans. La quantité de connaissance dans le monde a doublé au cours des 10 dernières années et double tous les 18 mois ». G. Siemens continue et affirme que « notre habilité à apprendre ce dont on aura besoin demain est plus importante que ce qu'on connaît aujourd'hui ».

Dans cette perspective le système éducatif doit s'adapter et être prêt à répondre aux défis posés par la société. L'apprentissage formel (e.g. école, lycée, université) doit être en phase avec les autres types d'apprentissage : non-formel (e.g. stages en entreprises, formations, échanges d'expériences) et informel (e.g. apprentissage personnel).

Plusieurs acteurs politiques, parmi lesquels l'Union Européenne, ont mis en place des programmes pour permettre aux individus et aux organisations de s'adapter à cette

²² *Époque post-PC* : terme qui souligne le déclin des *Personal Computers* face aux nouveaux dispositifs d'accès à Internet, notamment les téléphones intelligents et les tablettes numériques.

²³ Thomas L. Friedman, dans le livre « *La terre est plate. Une brève histoire du XXIe siècle* » (Edition Saint-Simon, Paris, 2006) constate que la globalisation (due à l'évolution technologique) a changé de nombreux concepts au cœur de l'économie.

dynamique. Un rapport de l'Union Européenne, sur *les indicateurs de qualité de l'éducation et la formation tout au long de la vie*²⁴, note que « la vitesse considérable à laquelle les connaissances évoluent et deviennent obsolètes souligne au moins deux points:

1. l'apprentissage doit être conceptualisé comme un processus permanent qui commence à la naissance et se poursuit tout au long de la vie ;
2. pour refléter sa diversité et sa complexité, cet apprentissage doit être interprété au sens large, sur un plan tant individuel que social, et ne doit pas être circonscrit à une formation en rapport avec l'emploi. » (Commission Européenne, 2002, p. 84).

L'expérience d'apprentissage doit non seulement être présente tout au long de la vie, mais elle doit encore être efficace et de qualité. L'apprentissage doit implémenter les méthodes d'enseignement les mieux adaptées pour répondre aux besoins actuels de l'homme et de la société. Dans la société d'information²⁵ actuelle, cela revient à intégrer les nouvelles technologies de l'information et de la communication (NTIC)²⁶ aux méthodes d'enseignement.

L'utilisation des TIC doit s'intégrer dans les démarches de construction des systèmes d'apprentissage. En fait, G. Paquette définit l'« ingénierie pédagogique » comme « toute méthode de conception et de construction des systèmes permettant d'échanger, de partager et d'acquérir des informations dans le but de les transformer en connaissances, donc d'apprendre » (Paquette, 2002a, p. 4).

Les systèmes d'apprentissage manipulent l'information dans le but de la transformer en connaissance. L'ingénierie pédagogique doit innover et intégrer les TIC afin d'offrir un

²⁴ Pour définir le terme d'« apprentissage tout au long de la vie », dans une communication de la Commission des Communautés Européennes intitulée *Réaliser un espace européen de l'éducation et de la formation tout au long de la vie (2001b, p. 11)*, la Commission a proposé : « Toute activité d'apprentissage entreprise à tout moment de la vie, dans le but d'améliorer les connaissances, les qualifications et les compétences, dans une perspective personnelle, civique, sociale et/ou liée à l'emploi. »

²⁵ Suite de la société industrielle

http://ec.europa.eu/information_society/eeurope/2002/news_library/pdf_files/french.pdf

²⁶ Les fabuleux développements technologiques font parler aujourd'hui de NBIC : *Nanotechnologies, Biologie synthétiques, Informatique et Cognition*.

cadre de conception de dispositifs et méthodes d'enseignement modernes, efficaces et adaptés au contexte actuel. Le but principal de ces dispositifs d'apprentissage est de générer des expériences d'apprentissage efficaces. Le support du savoir évolue et, dès lors, la relation entre l'apprenant et le savoir est modifiée. Une nouvelle expérience d'apprentissage s'impose.

La conception et la mise en pratique de ces nouvelles expériences d'apprentissage entraînent des observations, des études et des modélisations de l'apprenant et de son environnement d'apprentissage. Dans cette section, nous nous intéresserons à quelques enrichissements de l'expérience d'apprentissage.

1.2.1. Environnement et expérience d'apprentissage

Durant le XX^{ème} siècle, des courants tels que L'École Active et l'Éducation Nouvelle proposent des pédagogies actives, centrées sur l'apprenant et ses besoins (v. Figure 1-2). Ce dernier est encouragé à être actif et à interagir avec son environnement d'apprentissage.

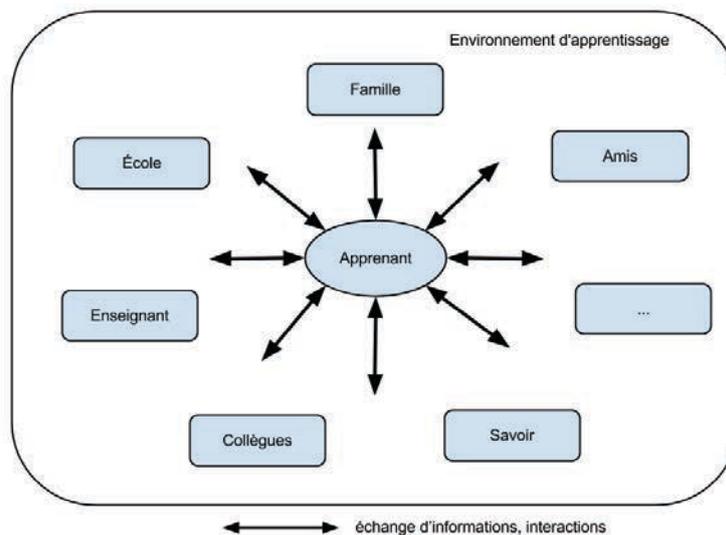


Figure 1-2. L'apprenant dans son environnement d'apprentissage

Le terme *environnement d'apprentissage* est large, et une utilisation négligente peut créer une confusion. Des termes comme *plateforme d'apprentissage numérique*, *école*, *salle de classe*, *meuble*, ou *bureau* (Gayeski, 1995), peuvent être entendus comme faisant partie de l'environnement d'apprentissage.

Un environnement d'apprentissage peut être défini comme l'environnement physique ou virtuel dans lequel se déroule l'apprentissage (« E-Learning Glossary », s. d.-a). Ainsi, un tel environnement fait référence à l'ensemble des composantes et des activités au sein duquel se déroule l'apprentissage (« Learning environment - EduTech Wiki », s. d.).

Dans le cadre de cette thèse, nous utiliserons le terme d'*environnement d'apprentissage* pour désigner l'ensemble des éléments avec lesquels l'apprenant interagit pour apprendre *dans le cadre du système d'apprentissage proposé*. La caractéristique principale de ces éléments est leur organisation dans l'environnement pour favoriser et stimuler l'apprentissage chez l'apprenant.

Conformément à Hannafin, Land et Oliver (1999, p. 123), les environnements d'apprentissage présentent quatre caractéristiques de base :

1. ils favorisent les contextes et ont la capacité de diriger l'apprenant dans des situations d'apprentissage spécifiques (e.g. imposées, induites) ;
2. ils disposent de ressources d'information et d'apprentissage (e.g. numériques, imprimées, humaines) ; le système peut offrir l'information et la ressource demandées ;
3. ils proposent des outils, c'est-à-dire des moyens permettant de travailler avec les ressources et les idées (e.g. chercher, identifier, organiser, communiquer). Le choix d'outils est influencé par le contexte ;
4. ils construisent des « échafaudages » ou « supports »²⁷ c'est-à-dire des processus destinés à soutenir l'effort individuel d'apprentissage : support conceptuel (aide l'apprenant à déterminer ce qu' il doit considérer pour résoudre un problème), support métacognitif (aide au cheminement de la pensée pendant l'apprentissage), support procédural (comment l'apprenant peut utiliser les possibilités offertes par l'environnement), support stratégique enfin (guide l'apprenant dans l'analyse et l'approche d'une situation d'apprentissage ou d'un problème).

²⁷ *scaffolds* en anglais

J. Sandberg (1993, cité par « Learning environment - EduTech Wiki », s. d.), identifie les fonctions de l'environnement d'apprentissage (v. Figure 1-3). Ces fonctions sont assurées par les entités le composant :

1. guidage, instructions : cette fonction peut être assurée par un agent humain (en présentiel ou à distance) ou un agent intelligent ;
2. suivi d'apprentissage : agent humain, qui peut être l'apprenant lui-même ou un agent intelligent ;
3. collaboration : agents humains (collègues) ;
4. fourniture de ressources d'apprentissage, objets pédagogiques, cours, logiciels de simulation ;
5. fourniture de ressources supplémentaires d'information, ou matériels additionnels aux ressources d'apprentissage, livres ;
6. outils : logiciels d'aide à l'apprentissage ;
7. administration : fournit le contexte administratif de l'apprenant, université, école.

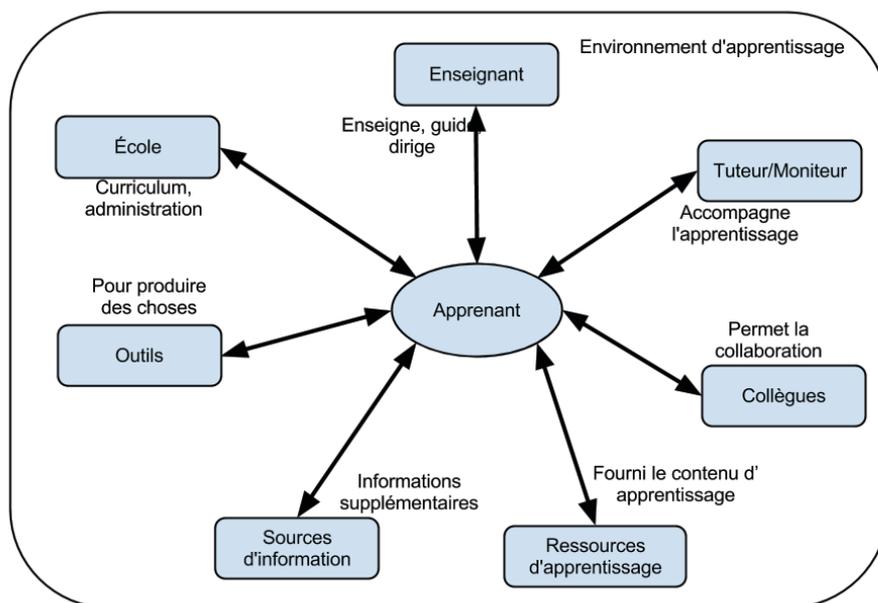


Figure 1-3. Fonctions de l'environnement d'apprentissage

Dans cette thèse, l'environnement d'apprentissage présenté est le même que celui géré par le système d'apprentissage qui représente notre contribution. Cet environnement est adapté aux possibilités et fonctionnalités offertes par le système d'apprentissage. Le système d'apprentissage offre un environnement qui intègre et améliore l'accès à des ressources d'apprentissage (notamment extérieures au système). Alors que certaines de ses caractéristiques sont accessibles au niveau du système sous forme de fonctions et

fonctionnalités, d'autres caractéristiques sont accessibles au niveau de la ressource d'apprentissage. L'apprenant interagit avec les éléments définis dans son environnement d'apprentissage. Nous nous concentrons sur les interactions de l'apprenant qui ont lieu dans l'environnement d'apprentissage généré par notre système. Le Chapitre 5. Modélisation du Système Actif et Sémantique d'Apprentissage (SASA) contient une description plus détaillée du système que nous avons conçu.

Dans le triangle pédagogique de Houssaye (2000) , illustré par la Figure 1-4, la situation d'apprentissage comporte trois acteurs : l'apprenant, le savoir et l'enseignant. « Houssaye estime que, dans cette relation triangulaire, il y a toujours un élément qui, comme au bridge, fait le mort, donc qui participe tout en étant hors-jeu. » (Raynal et Rieunier, 2010, p. 454) Le processus « enseigner » privilégie l'interaction enseignant-savoir, le processus « former » privilégie l'interaction enseignant-apprenant et le processus « apprendre » privilégie l'interaction apprenant-savoir.

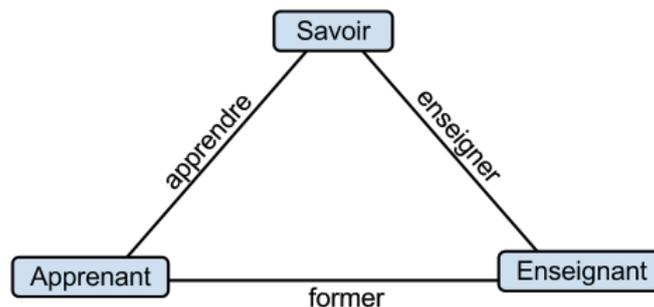


Figure 1-4. Le triangle pédagogique de J. Houssaye

Plaçant l'apprenant au centre de l'apprentissage, Moore (1989) définit trois types d'interaction dans l'apprentissage : apprenant – enseignant, apprenant – apprenant et apprenant – contenu. L'interaction apprenant – enseignant est un facteur majeur de l'apprentissage cognitif. L'interaction apprenant – apprenant favorise l'apprentissage par collaboration. L'interaction apprenant – contenu favorise l'interactivité entre l'apprenant et le contenu d'apprentissage, c'est-à-dire le savoir. Zhang (2005, p. 149) rappelle qu'il est essentiel d'offrir un accès simple, intuitif et rapide au contenu d'apprentissage, qui doit être bien traité et organisé.

Dans le système d'apprentissage que nous proposons, nous privilégions l'interaction de l'apprenant avec le savoir. L'enseignant peut être présent dans l'environnement

d'apprentissage et dans le déroulement de l'expérience d'apprentissage, mais son intervention n'est pas obligatoire. Il est davantage présent dans la phase préparatoire de l'expérience d'apprentissage – préparation du contenu et du parcours d'apprentissage –, mais il se contente d'un rôle d'observateur, de correcteur et d'accompagnateur dans la phase effective d'apprentissage. Il peut intervenir à tout moment afin de corriger et d'améliorer les interactions de l'apprenant avec l'environnement d'apprentissage, mais il part d'une position d'observateur.

Conformément à la typologie de (Bjørnåvold et European Centre for the Development of Vocational Training, 2000), nous distinguons trois types d'apprentissage en fonction du cadre de l'environnement :

1. l'apprentissage formel, à l'œuvre dans un cadre organisé et structuré (e.g. universités, écoles, formations en entreprises); spécialement conçu pour l'apprentissage, il peut conduire à une reconnaissance officielle (e.g. diplôme, certificat) ;
2. l'apprentissage non-formel ou apprentissage intégré dans des activités planifiées qui ne sont pas explicitement conçues pour l'apprentissage ;
3. l'apprentissage informel tel qu'il découle des activités de la vie quotidienne liées au travail, à la famille ou aux loisirs ; il n'est pas structuré en termes d'objectifs d'apprentissage, de temps ou de soutien à l'apprentissage ; généralement, ce type d'apprentissage ne conduit pas à une certification.

Le système d'apprentissage proposé dans cette thèse a pour but d'offrir un environnement d'apprentissage qui soutient l'apprentissage formel, avec des objectifs et des parcours d'apprentissage spécialement conçus à cette fin.

Pour définir comment l'apprenant se comporte et réagit à son environnement d'apprentissage, nous utiliserons le terme d'*expérience d'apprentissage*. D'après F. Marton et S. Booth (1998) l'expérience d'apprentissage est composée de deux constituants déjà présents dans la définition de « apprendre » : apprendre comme *processus* et apprendre comme *produit*. L'expérience d'apprentissage révèle comment l'apprenant a procédé dans des situations spécifiques d'apprentissage et comment il a ressenti l'apprentissage. D. Laurillard (1995) évoque une nouvelle expérience de

l'apprenant rendue possible par le multimédia. L'expérience d'apprentissage est enrichie grâce à l'ouverture des nouveaux supports à différents modes d'apprentissage : par acquisition (e.g. média narratifs), par discussion, par découverte (e.g. média interactif) ou encore par découverte guidée (e.g. média adaptatif).

Plaçant l'apprenant au centre d'apprentissage, nous utilisons le terme *expérience d'apprentissage* pour désigner l'ensemble des échanges qu'il a avec son environnement d'apprentissage. Cet ensemble est constitué d'échanges d'informations, d'interactions (avec l'enseignant et/ou le savoir) et d'activités (mentales et/ou physiques) que l'apprenant entretient avec et dans son *environnement d'apprentissage*, afin que cet ensemble produise le résultat final désiré et défini à l'avance.

Dans cette définition de l'expérience d'apprentissage nous identifions trois éléments : a) l'apprenant ; b) l'ensemble d'activités et d'interactions ; c) l'environnement d'apprentissage. L'enseignant fait lui aussi partie de l'environnement d'apprentissage. Les interactions de l'apprenant s'opèrent avec des entités existantes dans l'environnement d'apprentissage : l'enseignant, l'objet pédagogique, les collègues, les exercices, etc.

Pour l'amélioration de la qualité et de l'efficacité de l'expérience d'apprentissage, nous devons intervenir sur l'environnement et sur l'ensemble des activités et des interactions. L'intervention sur l'environnement présente deux dimensions. La première concerne la composition de l'environnement et les caractéristiques des entités qui le composent. La seconde dimension porte sur la régulation de l'interaction de l'apprenant avec son environnement. L'ensemble des activités et des interactions avec chaque entité de l'environnement devrait être régulé et normalisé pour stimuler une expérience d'apprentissage qui se veut efficace. L'expérience d'apprentissage doit s'adapter au fur à mesure de l'évolution d'acquisition des compétences de l'apprenant.

Dans un rapport sur les pratiques utilisées dans les systèmes d'apprentissage en ligne, citons une étude effectuée par le département de l'éducation des États-Unis en 2010, dans laquelle l'activité ou l'expérience d'apprentissage est analysée selon trois aspects :

1. l'objectif: *remplacement* d'un cours (e.g. cours virtuel) ou *renforcement* des éléments d'un cours donné précédemment en face-à-face ;
2. le type d'expérience d'apprentissage: *expositoire* (la technologie délivre le contenu); apprentissage *actif* (interactions avec le contenu à travers les simulations) ; apprentissage *interactif* (interaction avec les autres apprenants) ;
3. la dimension temporelle du type d'expérience d'apprentissage: *synchrone* (l'enseignement se produit dans le temps réel) ; *asynchrone* (décalage entre le temps de l'enseignement et le temps de réponse de l'apprenant).

(« Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies », 2010)

Dans cette perspective, le système d'apprentissage que nous proposons s'inscrit dans la catégorie des systèmes qui offrent un cours virtuel en utilisant les trois types d'expériences d'apprentissage en asynchrone. Dans un tel système d'apprentissage, la pédagogie progresse et occupe un rôle beaucoup plus important. À partir de « toute activité déployée *par une personne* pour développer des apprentissages chez autrui » (Raynal et Rieunier, 2010, p. 332), elle s'enrichit de toute l'activité déployée *par le système d'apprentissage* pour le développement d'une expérience qui provoque un apprentissage. Plusieurs types de pédagogies existent – active, de groupe, de projet, différenciée, individualisée – chacune ayant des caractéristiques susceptibles de convenir aux différents apprenants ou aux diverses situations d'apprentissage.

1.2.2. Apprentissage actif

Comme nous l'avons évoqué précédemment, l'apprentissage n'est pas un simple transfert de savoirs mais un processus dans lequel l'apprenant est activement engagé à construire ses connaissances. Il est l'acteur principal de son propre apprentissage. Cette idée fut promue par J. Dewey, le principal fondateur du mouvement *éducation nouvelle*²⁸. Initiateur du *learning by doing* et de la *pédagogie active*, il défend le principe que l'apprenant doit plutôt agir qu'écouter. A. Ferrière, atteint d'une maladie qui l'a rendu

²⁸ Éducation nouvelle: apparu vers les années 1900, ce courant pédagogique défend que l'apprenant doit *activement* participer à sa propre formation. L'école active, les méthodes actives et la pédagogie active sont des idées inhérentes à l'éducation nouvelle. Adolphe Ferrière (1879-1960) est l'un de principaux promoteurs de ce courant. Il est parmi les premiers à utiliser l'appellation « école active ».

sourd dès l'âge de 20 ans, renonce à enseigner et s'intéresse à des méthodes actives d'apprentissage, rejoignant ainsi le courant de l'éducation nouvelle. Auteur de *L'École Active* (Ferrière, 1922), il rédige aussi 30 *invariants pédagogiques*²⁹ ainsi que la charte sur laquelle se fonde la *Ligue Internationale pour l'Éducation Nouvelle (LIEN)*³⁰. La pédagogie active et l'éducation nouvelle défendent l'idée que l'apprenant est au centre de l'apprentissage et qu'il doit *activement participer* à sa propre formation.

En introduisant une imprimerie dans sa classe rurale en 1924, C. Freinet propose une pédagogie fondée sur des méthodes novatrices : « plan de travail, production de textes libres, imprimerie, individualisation du travail, enquêtes et conférences, ateliers d'expression-crédation, correspondance scolaire, éducation corporelle, réunion de coopérative » (« Education nouvelle et pédagogie active », s. d.). Il « cherche par tous les moyens de rendre *l'apprenant* actif » (Raynal et Rieunier, 2010, p. 207)

E. Claparède, l'un des psychologues et pédagogues qui s'intéressent à l'éducation nouvelle, soutient que l'éducation doit se retrouver dans l'action, cette dernière étant provoquée par un besoin, (Claparède, 1931) et que l'intelligence est une fonction active d'adaptation aux situations nouvelles.

Thorndike et Skinner mettent l'accent sur l'importance de l'activité³¹ de l'apprenant dans la construction de ses comportements : « C'est par l'action que des réponses sont effectuées et ensuite sélectionnées (par le biais d'agents de renforcement et de feedback) en fonction de la situation pour une association adéquate. » (Rebetez, 2003).

²⁹ Les invariants pédagogiques sont présentés comme un code pédagogique ou une nouvelle gamme de valeurs scolaires. Commenant par affirmer que « l'enfant est de la même nature que nous », il renforce les principes de la pédagogie active.

³⁰ La *Ligue Internationale pour l'Éducation Nouvelle (LIEN)* créée en 5 août 1921 par Adolphe Ferrière, Jean Piaget, Maria Montessori, considère que « le but essentiel de toute éducation est de préparer l'enfant à vouloir et à réaliser dans sa vie la suprématie de l'esprit [...] ». Ils rappellent l'importance de *l'individualité de l'enfant* et de *l'apprentissage de la vie*, et que toutes les expériences « doivent donner libre cours aux intérêts innés de l'enfant [...] *activités variées* d'ordre manuel intellectuel, esthétique, social et autres » (« Ligue Internationale pour l'Éducation Nouvelle », 1922).

³¹ D. Zhang (2005, p. 149) affirme que « l'apprentissage est un processus actif dans lequel un apprenant utilise l'entrée sensorielle et construit le sens à partir des informations reçues, et non pas une acceptation passive des connaissances qui existent déjà ».

C.C. Bonwell affirme que l'apprentissage actif peut être défini par tout ce qui amène les étudiants à faire des choses et à penser aux choses qu'ils font (Bonwell, 1991). Il note aussi quelques caractéristiques de l'apprentissage actif :

- les apprenants sont plus impliqués et engagés dans le processus d'apprentissage ;
- l'accent est mis sur le développement des aptitudes de l'apprenant ;
- la motivation des apprenants est augmentée ;
- le niveau de pensée est renforcé (analyse, synthèse, évaluation).

Les facettes de l'apprentissage actif se retrouvent dans chaque théorie d'apprentissage évoquée. Dans l'approche béhavioriste, les actions et les activités faites par les apprenants (e.g. les réponses de l'apprenant aux stimuli) sont au centre de l'apprentissage. Dans les approches cognitivistes et constructivistes, l'apprentissage actif est soutenu par les processus cognitifs des apprenants, qui doivent penser et comprendre leurs actions, et par la construction mentale de leurs actes. L'apprentissage actif est présent aussi dans les discussions et interactions sociales qui forment la base du socioconstructivisme. Dans le connectivisme, l'apprentissage actif consiste dans la construction des liens entre les nœuds (contenu d'apprentissage). Rendre la connaissance implicite (interne) en connaissance explicite (externe) et construire une correspondance entre l'abstrait et le concret est considéré comme un des points forts du connectivisme.

D'après un article paru sur le site de *Stanford – Center for Teaching and Learning*, l'apprentissage actif, caractérisé comme but ultime, favorise la pensée critique, la créativité, la collaboration, l'investissement de l'apprenant, la motivation et la performance (« The Ultimate Goal: Active Learning | Center for Teaching and Learning », s. d.). J. Grunert (1997) affirme que les apprenants apprennent davantage quand on leur demande de « participer d'une façon active à l'expérience d'apprentissage, que ce soit par discussions, exercices pratiques, révisions ou applications ».

Dans « Creating significant learning experiences », Fink (2003, p. 107) présente une image d'ensemble de l'apprentissage actif (v. Figure 1-5).

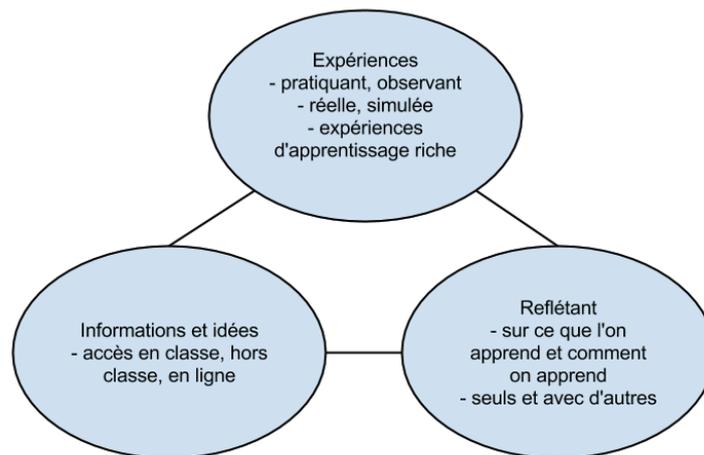


Figure 1-5. Image d'ensemble sur l'apprentissage actif (adapté de Dee Fink)

Les expériences d'apprentissage (e.g. la pratique, l'exercice, l'observation), doivent se tisser en relation avec les savoirs (e.g. contenu d'apprentissage) et la cognition (e.g. qu'est-ce qu'on apprend, comment apprend-on) (ibid.).

Jacques Henry et Jocelyne Cormier³² présentent leur résumé sur la démarche générale dans l'apprentissage et la construction de l'activité d'apprentissage (v. Figure 1-6). Dans leur schéma, les activités mentales (réflexion) sont en relation avec les actes de l'apprenant. Comme le processus cognitif est étroitement lié aux actions et à l'activité d'apprentissage, la modélisation des actions et des activités de l'apprenant dans son environnement d'apprentissage permettent de mesurer et d'interpréter l'activité cognitive de l'apprenant.

³² Consultants-pédagogues et fondateurs de DISCAS
<http://csrdsn.qc.ca/discas/infoDiscas/infoDISCAS.html>

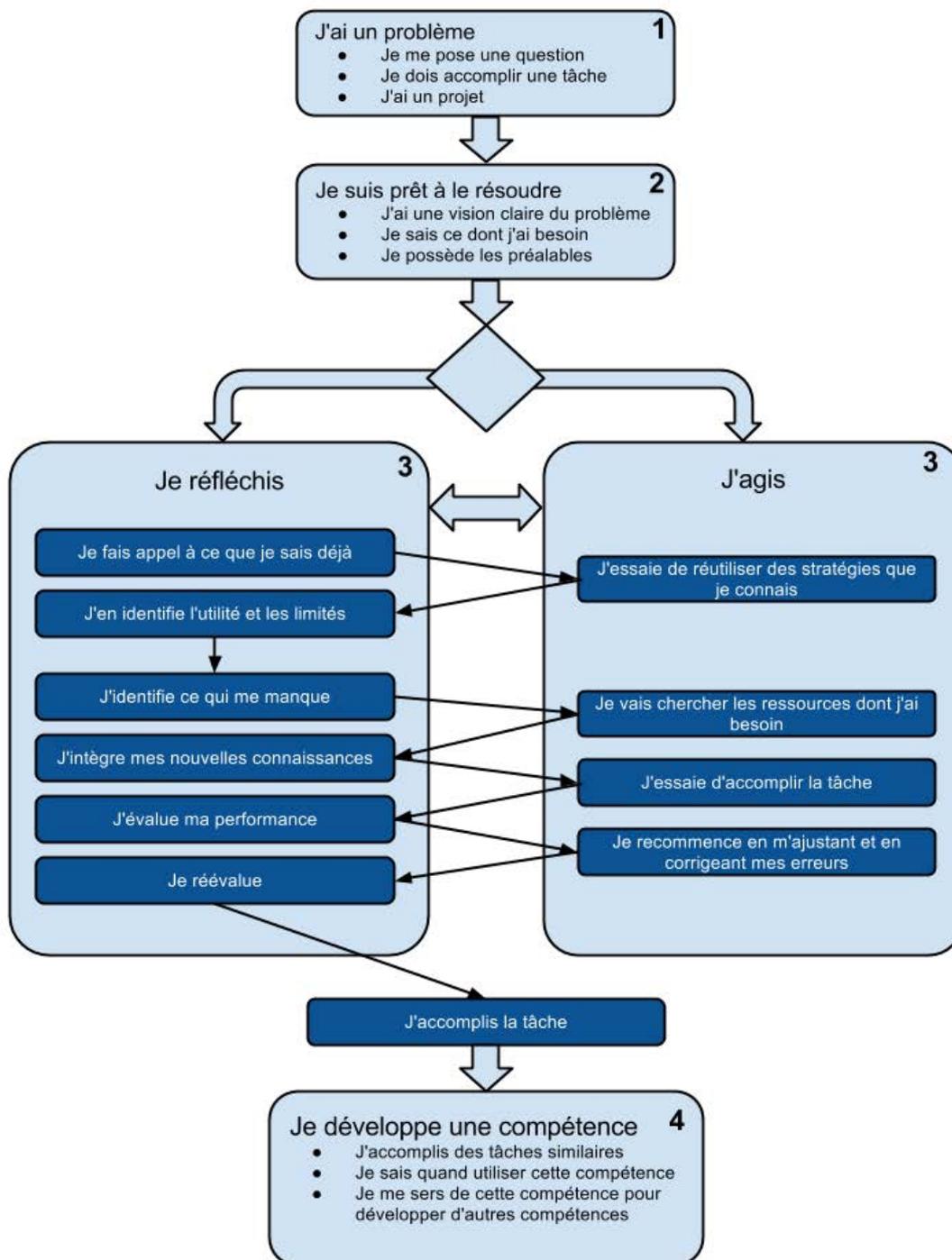


Figure 1-6. Démarche générale d'apprentissage (source : « L'activité d'apprentissage », s. d.)

La réponse et l'implication de l'apprenant dans l'expérience d'apprentissage sont mesurées à l'aide des actions et des activités définies dans le système d'apprentissage.

Par rapport à la pédagogie traditionnelle, qui se concentre plus sur la transmission des connaissances, la pédagogie active rend l'apprenant acteur de ses apprentissages. L'environnement d'apprentissage doit prendre en charge les fonctions d'interaction et

assurer l'entraînement de l'apprenant dans des expériences d'apprentissage actif. L'élément clé responsable de la réalisation d'une expérience d'apprentissage actif est la ressource d'apprentissage. Elle doit être conçue conformément aux principes de la pédagogie active. En offrant des possibilités élevées d'interaction, la ressource d'apprentissage doit rendre l'apprenant actif. L'activité de l'apprenant est donc conditionnée par la ressource d'apprentissage. Dans le cas le plus simple, un type de ressource d'apprentissage est constitué d'un texte simple. Ce type de ressource impose une activité de lecture. Des questions ou un questionnaire enrichissant ce type de ressource d'apprentissage vont permettre un retour sur l'implication de l'apprenant dans sa lecture, mais aussi sur la compréhension de sa lecture.

La complexité des types d'interaction possibles à travers l'ordinateur, rend possible un plus grand nombre d'activités d'apprentissage (e.g. manipulation des objets virtuels, simulation des circuits électroniques). Le même sujet d'apprentissage peut être conçu selon différentes ressources d'apprentissage en utilisant différentes formes de présentation (e.g. texte, son, vidéo) et types d'interaction (e.g. réponses écrites aux questions – textes, réponses audio pour l'apprentissage des langues). Cette diversité permet de répondre à la fois à l'adaptation de la forme de présentation du contenu à apprendre, et à la diversité de types d'apprentissages humains (e.g. styles d'apprentissage).

Pour la conception du système actif d'apprentissage, nous considérons l'existence de plusieurs ressources d'apprentissage adaptées aux différentes formes et types d'interaction. Le *système actif* permet à l'apprenant d'intervenir sur son parcours d'apprentissage. Même si l'interaction essentielle de l'apprenant avec l'environnement d'apprentissage se situe au niveau de la ressource d'apprentissage, la possibilité d'influer sur le choix des ressources d'apprentissage reste un facteur important dans l'induction d'une expérience d'apprentissage active et adaptée.

Le but de ce travail de recherche est d'étudier la modélisation des composantes de l'environnement et de l'expérience d'apprentissage. D'une part, les caractéristiques des ressources d'apprentissage et leur enchaînement, d'autre part l'apprenant avec ses particularités d'apprentissage. La correspondance entre les caractéristiques de ces deux

parties permet l'adaptation du parcours d'apprentissage et la génération d'un apprentissage actif tout en ouvrant la voie vers la personnalisation de l'expérience d'apprentissage. Nous utilisons la modélisation de ces ressources pour faciliter leur choix dans le but de stimuler une expérience d'apprentissage active.

Un système d'apprentissage doit sélectionner les ressources d'apprentissage qui stimulent l'activité pédagogique de l'apprenant. Le système influe donc le parcours d'apprentissage, mais aussi l'activité pédagogique de l'apprenant. Comme le souligne S. George et A. Derycke : « actuellement les activités pédagogiques sont davantage mises en avant [...] une plate-forme peut être vue comme un système qui permet de gérer et de donner accès à un ensemble d'activités et de ressources pédagogiques » (George et Derycke, 2005).

Le fait d'être un apprenant actif implique un retour de la part de l'environnement d'apprentissage (e.g. objet pédagogique, tuteur). Ce retour de l'environnement d'apprentissage doit être intelligent, adapté et en mesure de stimuler et de valoriser le potentiel d'apprentissage de l'apprenant.

Une partie importante de l'implication de l'apprenant dans l'activité d'apprentissage est donnée par son interaction avec les autres apprenants. La technologie doit soutenir la participation et les interactions entre les apprenants pendant leur apprentissage (Manson, 2007). L'utilisation des outils de communication, comme le chat, le forum, les liste de discussions constituent un facteur important dans l'activité de l'apprenant. Pour visualiser ces interactions May et al. (May, George, et Prévôt, 2011) présente TrAVis (Tracking Data Analysis and Visualisation).

1.2.3. Apprentissage personnalisé

Etant données la complexité et la diversité des êtres humains et de leur façon d'apprendre, l'apprentissage ne peut pas être abordé avec une approche du type *one size fits all*, mais dans une démarche où chacun est partie prenante d'une expérience d'apprentissage adaptée à ses particularités.

Dans « Les sept savoirs nécessaires à l'éducation du futur », E. Morin affirme que les êtres humains « doivent se reconnaître dans leur humanité commune, en même temps que reconnaître leur diversité tant individuelle que culturelle » (Morin, 1999, p. 23). Notre diversité se reflète dans notre façon d'être, de penser, d'agir, d'apprendre. C'est cette diversité individuelle et culturelle qui nous pousse vers des démarches et stratégies d'enseignement diversifiées dans la perspective d'une expérience d'apprentissage personnalisée. En répondant à cette diversité, la personnalisation de l'expérience d'apprentissage a pour but d'améliorer l'apprentissage en regardant au plus proche les particularités de l'apprenant. Celles-ci sont constituées par les caractéristiques de l'apprenant considérées en relation avec son apprentissage. Ce sont elles qui font de l'apprenant un être distinct d'autres apprenants en rapport avec l'apprentissage.

En elle-même, la personnalisation de l'expérience d'apprentissage consiste en l'utilisation de méthodes d'adaptation de la pédagogie, de l'enseignement, du contenu d'apprentissage et de la forme d'apprentissage pour répondre aux particularités de l'apprenant. L'expérience d'apprentissage personnalisée est l'effet d'une pédagogie adaptée et d'un enseignement ajusté qui ont pour but la maximisation des effets de l'apprentissage. Présents sous différentes facettes, les aspects spécifiques mettent en évidence les différences des approches et termes existants : formation individualisée, pédagogie différenciée, formation personnalisée, etc.

La *formation individualisée* permet à l'apprenant d'apprendre et d'atteindre à son propre rythme, seul ou à plusieurs, selon un parcours diversifié, des objectifs d'apprentissage définis. Elle s'appuie sur des « dossiers de formation individualisée et/ou d'autoformation » (Raynal et Rieunier, 2010, p. 203,231).

Influencé par les écrits de J. Dewey, Helene Parkhurst, pédagogue américaine, fonde une méthode de travail individualisé connue sous le nom de plan Dalton (Helen Parkhurst, 1922). Observant que « les différences individuelles, les capacités et les intérêts sont trop variés pour permettre à la classe de fonctionner comme une entité » (Dewey, 1922, p. 19), le plan Dalton propose d'organiser l'apprentissage en fonction de besoins, intérêts et capacités de l'apprenant. C'est le premier exemple de pédagogie différenciée.

Le terme de *pédagogie différenciée* a été proposé par Louis Legrand au début des années 1970. La pédagogie différenciée admet que chaque apprenant apprend à sa manière selon des « styles d'apprentissage » particuliers, chacun ayant des capacités, compétences et difficultés spécifiques. La notion de « style d'apprentissage » renvoie à la psychologie différentielle, aux domaines cognitif et conatif (Raynal et Rieunier, 2010, p. 337; Frayssinhes, 2012, p. 157).

La psychologie différentielle a pour objectif de « mesurer, décrire et expliquer les différences stables entre les individus » (Lautrey, 2006). Ces différences sont exprimées à travers des caractéristiques individuelles ou « traits d'une personne qui font que ses comportements sont différents de ceux des autres personnes » (Viau, 2005, p. 17). Ces caractéristiques individuelles sont en relation avec l'apprentissage. Viau (ibid., p. 18) reprend la classification des caractéristiques individuelles de (Corno et Snow, 1986) et (R. Snow et Farr, 1987) les groupent dans trois domaines principaux : le domaine cognitif, le domaine conatif et le domaine affectif (v. Figure 1-7). La juxtaposition des cercles symbolise l'interdépendance de ces caractéristiques.

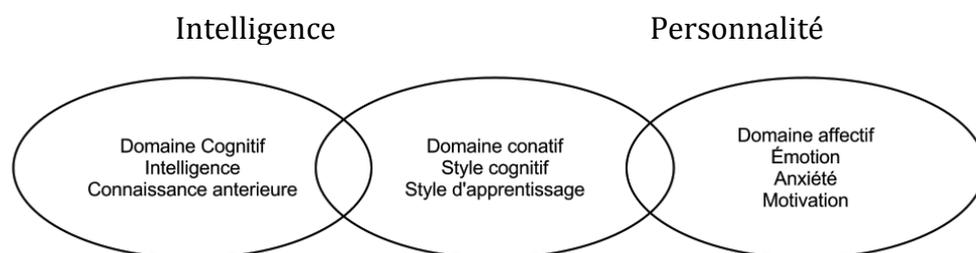


Figure 1-7. Caractéristiques individuelles de l'apprenant, selon (Hilgard, 1980)

Les caractéristiques sont mesurées par « l'intermédiaire des comportements auxquels elles donnent lieu. On mesure, par exemple, la motivation à l'aide de comportements comme la persévérance d'un élève dans l'accomplissement d'une activité » (Viau, 2005, p. 18).

En recouvrant tout ce qui fait appel aux activités cognitives et aux processus mentaux qui permettent l'acquisition des connaissances (e.g. perception, mémorisation, raisonnement, jugement), le domaine cognitif permet la modélisation de l'apprenant en relation avec son niveau de connaissances et sa capacité d'acquérir des connaissances.

Le domaine conatif renvoie à l'orientation, l'autorégulation et le contrôle des conduites (Frayssinhes, 2012, p. 157). « Si le mot *conation* s'applique à la manifestation active d'une tendance, il peut être utilisé pour désigner l'orientation des conduites, c'est-à-dire des activités finalisées et organisées. Le terme dénote aussi un effort exigé par cette manifestation » (Reuchlin, 1990, p. 11). Dans (R. E. Snow et Jackson, 1997, p. 3), la conation est située « entre les domaines cognitif et affectif ». Le domaine conatif est influencé par le domaine affectif. La motivation, l'orientation et le contrôle des conduites sont influencés par les processus affectifs (e.g. l'émotion, le tempérament). Les styles d'apprentissage se trouvent à l'interaction de ces trois domaines.

Deux points de vue spécifiques distinguent les tenants des différents styles d'apprentissage : d'abord le point de vue *innéiste* selon lequel le style d'apprentissage est une « caractéristique immuable, figée » (Dunn et Dunn, 1978, 1993) ; ensuite le point de vue *empiriste* selon lequel le style d'apprentissage est une caractéristique essentiellement acquise (Honey et Mumford, 1992; Kolb, 1973, 1984).

Les styles d'apprentissage peuvent-être considérés comme « des différences généralisées dans les orientations d'apprentissage basées sur le degré relatif d'accent mis par les gens sur les quatre modes du processus d'apprentissage tels que mesurés par un questionnaire "self-report" nommé le "Learning Style Inventory" » (ibid., p. 67). Vus en relation avec les différents types d'activités d'apprentissage, « les styles d'apprentissage sont des tendances générales à préférer traiter l'information de différentes façons » (Jonassen et Grabowski, 1993, p. 233; cité par Frayssinhes, 2012, p. 164)

La relation entre les styles d'apprentissage et les ressources d'apprentissage, en passant par les activités d'apprentissage (correspondantes à chaque ressource d'apprentissage), permet la personnalisation du parcours d'apprentissage.

La personnalisation implique la construction du profil d'apprenant, notamment la spécification explicite de son style d'apprentissage et l'adaptation de l'expérience d'apprentissage à ce profil. Différent du profil d'utilisateur, le profil d'apprenant contient

des caractéristiques plus spécifiques en relation avec l'apprentissage et les préférences personnelles liées aux activités d'apprentissage de l'utilisateur.

Les méthodes d'adaptation de l'apprentissage se retrouvent à différents niveaux de l'expérience d'apprentissage. Une première approche est centrée sur la conception des ressources d'apprentissage. La construction de plusieurs ressources d'apprentissage avec les mêmes objectifs d'apprentissage, mais qui utilisent des formes différentes pour la communication du contenu d'apprentissage (e.g. vidéo, son, textes, exercices interactifs), permet une première possibilité d'adaptation du contenu.

Nous allons regarder comment nous pouvons enregistrer les descriptions du contenu d'apprentissage pour l'adapter au profil de l'apprenant au bon moment du parcours d'apprentissage. Dans ce qui suit, nous utilisons l'expression *apprentissage personnalisé* pour illustrer l'adaptation de l'expérience d'apprentissage aux particularités de l'apprenant. Cette adaptation est assurée en deux temps : tout d'abord, en tenant compte de ses objectifs et niveau d'apprentissage, puis, en tenant compte de son style d'apprentissage.

L'évolution technologique a rendu la personnalisation de l'expérience d'apprentissage accessible : « On affirme que l'accès aux savoirs s'en trouve miraculeusement démocratisé, on considère que les technologies de l'information permettent, enfin, cette *individualisation de la formation* ». (Meirieu, Kambouchner, et Stiegler, 2012)

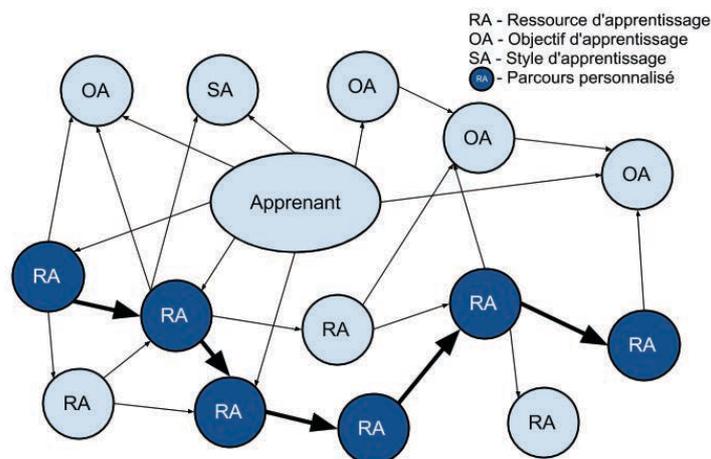


Figure 1-8. Parcours d'apprentissage personnalisé

Le système d'apprentissage que nous avons développé utilise des descripteurs des ressources d'apprentissage et les profils des apprenants pour adapter le parcours d'apprentissage (v. Figure 1-8). La personnalisation de l'expérience d'apprentissage dans le cadre du système est obtenue grâce aux outils exploités pour la description des ressources et aux mécanismes de régulation des interactions entre les entités définies dans le système d'apprentissage. Chaque apprenant peut bénéficier de son agent personnel d'apprentissage qui enregistre son profil d'apprentissage (e.g. style d'apprentissage, objectifs d'apprentissage). Cet agent personnel analyse les résultats obtenus aux questionnaires et peut proposer un contenu d'apprentissage ajusté au niveau, besoins ou préférences de l'apprenant.

1.2.4. Apprentissage à travers les technologies de l'information et de la communication

L'évolution technologique a imposé des changements dans la majorité des domaines d'activité de la société actuelle. Le courrier est devenu électronique (*mail* en anglais, devenu *email*), le livre est maintenant électronique (*book*, *ebook*), le commerce est devenu *e-commerce*, le business est *e-business*, le gouvernement communique avec les citoyens à travers le *e-gouvernement*. Le petit *e* en préfixe vient du mot *electronic* et désigne l'augmentation du service par des moyens électroniques tels que la télévision interactive, le CD-Rom, l'Internet, etc.

Le *e-Learning*³³, traduit par « apprentissage électronique », correspond à l'apprentissage par des moyens électroniques. L'Union Européenne définit l'e-learning comme « l'utilisation des nouvelles technologies multimédias et de l'Internet pour améliorer la qualité de l'apprentissage en facilitant l'accès à des ressources et des services, ainsi que les échanges et la collaboration à distance » (Commission des Communautés Européennes, 2001a, p. 2).

³³ Le terme e-Learning est utilisé dans un contexte professionnel en 1999 (« History of e-Learning », s. d.). Cependant, des articles utilisant le terme sont publiés dès 1997 (« A bright future for distance learning: One Touch/Hughes alliance promotes interactive "e-learning" service », 1997)

Le e-Learning est couramment associé à la formation en ligne et à la formation ouverte à distance (FOAD). En introduisant le qualificatif *ouvert*, FOAD met en évidence une plus grande accessibilité, souplesse et flexibilité dans son mode d'organisation.

Lors du développement et de l'adoption de ces technologies, plusieurs appellations ont été développées :

1. TI – technologies de l'information (IT – Information Technology en anglais) : la composante principale est l'informatique ;
2. TIC – technologies de l'information et de la communication (ICT – Information and Communication Technology) : à la TI s'ajoute la communication, qui implique les réseaux³⁴ ;
3. NTIC – nouvelles technologies de l'information et de la communication (NICT – New Information and Communication Technology)³⁵ ;
4. TICE – technologies de l'information et de la communication pour l'enseignement.

L'ensemble des technologies de l'information et de la communication fournit les ressources nécessaires à la manipulation de l'information : ordinateurs, logiciels, éléments de réseau. Les TICE intègrent les outils numériques conçus et utilisés pour l'administration d'une expérience d'apprentissage dans un environnement numérique (e.g. production des ressources numériques d'apprentissage, création et administration des ressources numériques d'apprentissage, création et administration de l'environnement numérique d'apprentissage).

Proposé par l'Union Européenne, le terme « *Technology-Enhanced Learning* »³⁶, traduit en français par « *Apprentissage amélioré par la Technologie* », met l'accent sur l'importance de l'utilisation des TIC dans l'apprentissage. La construction de l'environnement d'apprentissage avec les outils des TIC permet l'enrichissement de l'expérience d'apprentissage. L'environnement d'apprentissage devient « virtuel » ou

³⁴ Le Web 2.0 est un exemple d'usage des TIC

³⁵ Les TIC renvoient au terme général pour les technologies utilisées dans la communication (qui peuvent être la télévision, le radio); les NTIC renvoient aux technologies plus récentes, comme le web, le chat, l'email, etc.

³⁶ La première conférence EC-TEL« European Conference on Technology Enhanced Learning » a eu lieu en 2006 en Grèce : <http://www.ectel06.org/>

« numérique ». Les moyens de communication sont enrichis et les possibilités d'interaction avec les ressources d'apprentissages sont augmentées.

Les objets informatiques et les outils de l'information et de la communication permettent la construction de logiciels éducatifs (e.g. exercices, tutoriels, pratique). Les environnements virtuels d'apprentissage sont un type spécifique de logiciels éducatifs (v. Figure 1-9). Plusieurs de leurs composantes communiquent et travaillent ensemble pour construire une cohérence dans l'expérience d'apprentissage de l'apprenant. La communication, les relations et la cohérence entre ces composantes renvoient à la notion de système. Dans le cadre de ce travail, nous utilisons le terme de *système d'apprentissage* pour désigner cet ensemble de modules (e.g. logiciels) et les relations correspondantes.

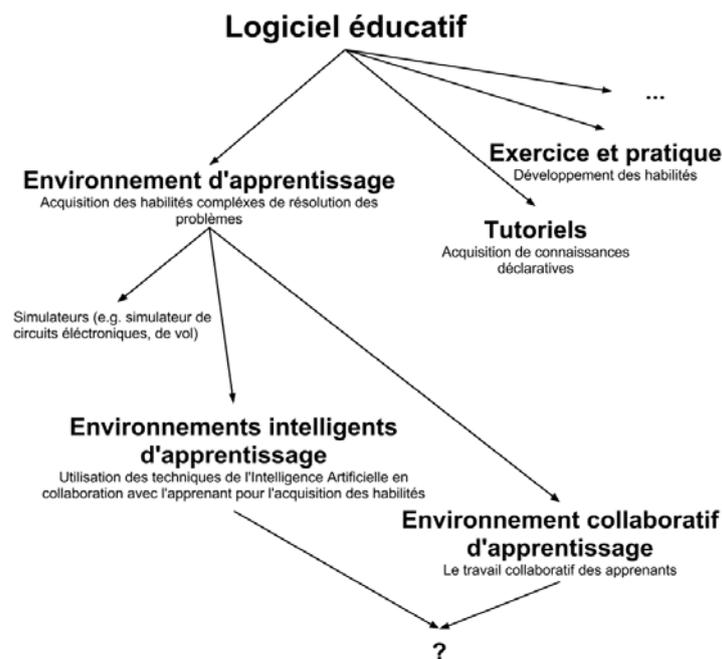


Figure 1-9. Vue générale des logiciels éducatifs (adapté après Schneider, 1997)

L'environnement virtuel d'apprentissage retrace les limitations d'espace et de temps tout en permettant la personnalisation de l'expérience d'apprentissage. Deux conceptions s'opposent sur les fonctions et les capacités des outils (e.g. l'ordinateur) dans l'enseignement. La première affirme qu'en tant qu'outils, leur rôle est neutre : « L'ordinateur est un outil, incapable d'agir de lui-même. Il n'a pas la sagesse innée, sans avoir son propre esprit, aucune initiative, et aucune capacité inhérente d'apprendre ou d'enseigner » (Ahmad *et al.*, 1985, p. 2). Levy (1997, p. 181, 208) affirme que les outils

sont « neutres » et que leur fonction n'est pas d'enseigner ; la puissance de l'outil [l'ordinateur] se trouve dans « sa polyvalence et sa capacité d'augmenter les capacités de l'homme ».

La deuxième conception affirme que « l'outil est au centre de l'histoire de l'homme depuis ses origines. Relation circulaire au cœur de la pédagogie, l'homme fabrique l'outil et en retour l'outil³⁷ façonne l'homme » (Dieuzeide, 1994, p. 18).

L'utilisation et l'intégration des TIC représentent une étape normale dans le développement des outils pour l'apprentissage. J. Rosnay (1974, p. 295) affirme que « l'ordinateur représente, en principe, le prolongement idéal du livre programmé et de la machine à enseigner ». Ces outils rendent plus accessibles les possibilités de personnalisation de l'apprentissage. B.F. Skinner (1968) met en avant les possibilités de l'individualisation de l'enseignement : l'activité soutenue induite par la machine conduit à un échange continu entre le programme et l'apprenant ; la machine impose que chaque point soit parfaitement compris avant d'aller plus loin ; elle présente la matière, la ressource et l'objet que l'élève est préparé à aborder ; elle aide l'élève à produire la réponse correcte ; utilisant un feed-back immédiat, la machine modèle le comportement de l'apprenant et le tient en activité. (Rézeau, 2001, p. 243-244)

Les possibilités de personnalisation sont également évoquées par M. Alberganti (2000, p. 155) : « Devant son ordinateur, l'élève se retrouve [...] dans la position privilégiée du cours particulier. Le logiciel d'apprentissage, lui, joue le rôle du précepteur attaché à un seul et unique élève. [...] L'ordinateur se glisse sans difficulté dans la peau de Socrate ».

L'environnement d'apprentissage devient *personnel*. Il est propre à l'apprenant et reflète ses besoins et ses particularités. Les fonctions offertes par l'environnement répondent aux particularités de l'apprenant. Ses fonctions ne se limitent pas à un simple échange d'information, mais il introduit l'apprenant dans des situations d'apprentissage adaptées à ses besoins et préférences. Tous les modules de l'environnement sont synchronisés

³⁷ « [...] contrairement à l'intuition commune, l'instrument n'est absolument pas neutre par rapport au réel [...] [les instruments] contiennent une « *conception du monde* » qui s'impose peu ou prou à leurs utilisateurs, et [ils] influencent ainsi le développement de leurs compétences » (Rabardel, 1995, p. 213; cité par Rézeau, 2001, p. 271).

pour rendre une expérience d'apprentissage cohérente et adaptée aux besoins de l'apprenant.

Nous reprenons l'illustration de A. Derycke (2009) pour éclairer la métamorphose de l'apprentissage en combinaison avec les TIC (v. Figure 1-10). Les deux modes d'apprentissage en fonction de la variable espace, la formation présentielle et la FOAD convergent pour créer un mode hybride de l'apprentissage : la formation pervasive, le p-Learning.

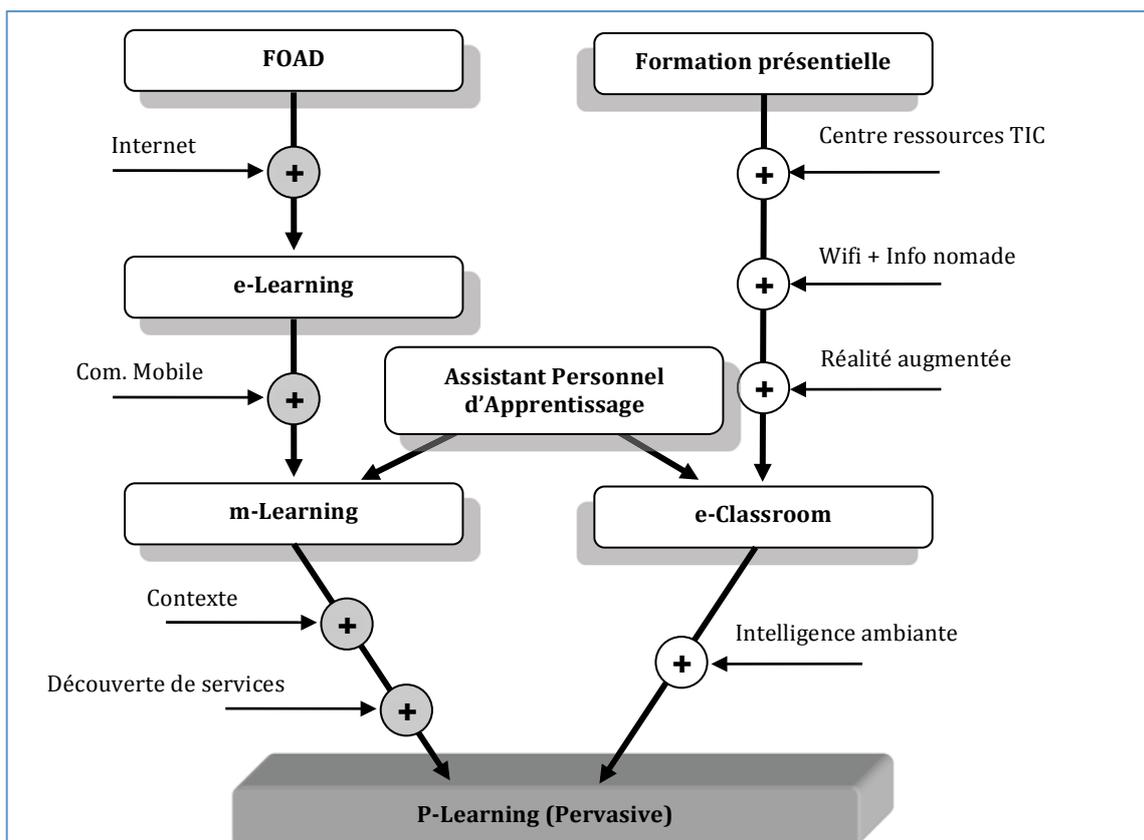


Figure 1-10. Hybridation des modes d'apprentissage (selon *ibid.*)

Le développement des dispositifs mobiles (e.g. tablettes, téléphones intelligents) favorise le passage du « e-Learning » au m-Learning (l'apprentissage mobile). Avec l'ajout des ressources, des outils créés par les TIC et de la réalité augmentée, la salle de classe devient e-Classroom. L'assistant personnel d'apprentissage peut intervenir à ces niveaux pour personnaliser l'expérience d'apprentissage. L'ajout des services de géolocalisation, des services contextuels et de l'intelligence rendent l'apprentissage ubiquitaire ou *pervasif* (p-Learning).

1.3. Savoirs et connaissances à l'ère du numérique

Dans la conférence « Les nouvelles technologies : révolution culturelle et cognitive », à Lille en 2007, l'académicien M. Serres présente l'évolution des systèmes de l'information et de la communication (*Inria - Michel Serres - 2007 - Intro*, 2010). Il voit dans l'évolution technologique une externalisation des fonctions corporelles : c'est ce qu'il appelle l'*exo-darwinisme*.

M. Serres définit le message comme une information sur laquelle on peut appliquer quatre opérations de base : stockage, traitement, émission et réception. Entre le message et le support d'information, il y a toujours un couplage. Dans l'évolution du support de l'information et du couplage message-support, il observe quatre stades d'évolution : (1) oral ³⁸ ; (2) l'écriture manuscrite ; (3) l'imprimerie ; (4) l'ordinateur. Premier constat : la quantité d'information n'a cessé d'augmenter et l'évolution du support a facilité le développement des quatre opérations de base réalisées sur l'information.

Dans le stade oral, le support du message était le corps humain (le cerveau et le corps entier, la mémoire humaine jouait alors un rôle essentiel). La transmission de l'information, du message, était faite par voie orale. Ensuite, l'écriture devient le premier support extérieur au corps humain (vélin, papyrus, papier). L'invention de l'écriture va changer le couplage message-support et va provoquer la première révolution culturelle et cognitive³⁹. M. Serres affirme que la « pédagogie⁴⁰ et l'Etat⁴¹ » sont les fils de l'écriture. Le troisième stade dans l'évolution du support de l'information est déclenché par l'invention de l'imprimerie : « Il suffit de connaître l'endroit où se trouve le livre. C'est une catastrophe pour la mémoire ! » (*Inria - Michel Serres - 2007 - Les Hommes : le cognitif*, 2010).

³⁸ Oral au sens linguistique.

³⁹ « La première catastrophe est l'invention de l'écriture. » (Serres, 2007, p. 132).

⁴⁰ La pédagogie désigne l'art d'éduquer (du grec, « pais, paidos » - enfant et « gogia » - action de transformer, de conduire). Par extension, la pédagogie, science de l'éducation, désigne également la formation intellectuelle des adultes (ou andragogie).

⁴¹ L'organisation des villes devient possible grâce à l'écriture d'un droit écrit stable (le premier texte étant le célèbre code d'Hammurabi) et mène à l'invention de l'Etat.

Chaque changement du support provoque un changement culturel et politique (possibilités de mobilisation de citoyens) et religieux (la bible imprimée). Cette révolution change aussi notre manière de connaître, de savoir et d'apprendre. L'espace cognitif est totalement bouleversé.

Le quatrième stade de l'évolution du support du message est l'invention de l'ordinateur. Le passage au numérique des dispositifs de traitement de l'information ainsi que les évolutions des technologies de communication (e.g. les réseaux, l'Internet) ont permis une extraordinaire ouverture des dimensions spatiale (e.g. capacité de stockage des dispositifs, densité des dispositifs de traitement) et temporelle (e.g. accès à l'information, vitesse de transmission de l'information) dans lesquelles *vit* le *message*.

À ce stade, la mémoire s'est externalisée dans les ordinateurs et leur capacité de stockage dépasse les capacités des objets de stockage du stade précédant, d'où des conséquences immédiates dans l'augmentation de la quantité d'information. Au troisième stade, l'information se trouve sur un support physique (matériel) différent et spécifique pour chaque média (e.g. cassette VHS pour la vidéo ; disque vinyle pour la musique ; livre pour le texte et l'image). En cherchant à accéder à cette information, nous heurtons à des contraintes d'espace et de temps. Dans le « tout numérique », qui définit le quatrième stade de l'évolution, ces contraintes sont remplacées par la condition d'utilisation d'un dispositif d'interfaçage (e.g. ordinateur, téléphone portable, tablette, liseuse). À l'ère du numérique, la *pervasivité* et l'*ubiquité* sont des caractéristiques originelles des systèmes d'information et de communication.

Si « nous avons la mémoire et le savoir devant nous », la question qui se pose est : qu'avons-nous gagné ? Que nous reste-t-il ? Comme l'affirme M. Serres : « les nouvelles technologies nous ont condamné à devenir intelligents, [...] et nous sommes condamnés à devenir inventifs » (*Inria - Michel Serres - 2007 - Les Hommes*, 2010). Aujourd'hui, le travail intellectuel doit être un travail intelligent et non pas un travail répétitif.

En plus de leur capacité de stockage, les ordinateurs facilitent l'indexation et la classification de l'information. Les capacités d'inférence et de raisonnement des ordinateurs sur un grand ensemble de données ont permis la réalisation des bases de

connaissances. En intelligence artificielle, les représentations des connaissances font évoluer les bases de *données* vers les bases de *connaissances*. Ces bases de connaissances sont des « *ensembles de données destinées à être exploitées par un programme informatique, le moteur d'inférences* » (Marie-Christine ROUSSET dans Houdé *et al.*, 1998, p. 69-70).

Les termes *savoir* et *connaissance* sont utilisés pour capter le contexte de leur existence. Si on utilise le terme *savoir* pour désigner ce qui est « stocké dans les bibliothèques, [et] dans les banques de données », et le terme *connaissance* pour ce qui est « emmagasiné dans la tête des individus » (Raynal et Rieunier, 2010, p. 401), alors l'apprentissage est le processus qui permet la transformation des savoirs en connaissances. Cependant l'apprentissage n'est pas seulement le processus de stockage du savoir dans la tête des individus. Car la connaissance est construite par l'apprenant.

Le terme savoir⁴² indique un contenu : « ce que l'on sait » (La Borderie, Paty, et Sembel, 2000, p. 45). Les savoirs sont « des données, des concepts, des procédures ou des méthodes qui existent hors de tout sujet connaissant et qui sont généralement codifiés dans des ouvrages de référence, manuels, cahiers de procédures, encyclopédies, dictionnaires » (Meichel, 2009).

Les connaissances, par contre, sont « indissociables d'un sujet connaissant. Lorsqu'une personne intériorise un savoir en prenant conscience, précisément, elle transforme ce savoir en connaissance » (ibid.). Dans une perspective constructiviste, la connaissance est le savoir approprié, construit, compris, représenté par une personne. Chez une autre personne le même savoir ne sera pas représenté identiquement. « Comme il n'existe aucune connaissance parfaite et absolue, seul un savoir beaucoup plus ample permet de s'orienter raisonnablement en contexte d'incertitude. Un "connaisseur" ne dispose pas ainsi de toute l'expérience de recherche qui caractérise le "savant" » (ibid.).

Plusieurs définitions existent pour caractériser la connaissance en fonction de son domaine d'application. De manière générale, la connaissance est une « croyance vraie,

⁴² Du mot latin « *sapere* » qui signifie « avoir du goût, de l'intelligence , du jugement » (« DicoLatin », s. d.)

justifiée » ou une « représentation adéquate justifiée » (Le Ny, 1989, p. 26-27). Dans la psychologie cognitive, le terme connaissance est défini comme « *structure stabilisée en mémoire à long terme* » (Richard, Bonnet, et Ghiglione, 1990, p. 35). Les différentes nuances des *représentations de connaissances* sont soulignées par plusieurs auteurs dans des expressions telles que « *représentations cognitives, représentations mentales, représentations types* » par J.F. Le Ny (dans Ehrlich, 1985, p. 231-247), ou « *représentations sémantiques* » par S. Ehrlich (1985), ou encore « *représentation mentales des états du monde, connaissances sémantiques* » par G. Tiberghien (dans Monteil et Fayol, 1989).

En 1941, aux États Unis, l'apprentissage s'adaptait pour former les ouvriers dans un contexte de guerre. La société devenait une société de l'industrie et l'apprentissage par objectifs y était largement implémenté. En suivant l'évolution technologique, la société s'est transformée en une « société de l'information »⁴³, une « société du savoir »⁴⁴, et la connaissance est mise au centre de l'économie : l'« industrie de la connaissance ». Lors du Conseil européen de Lisbonne, les 23 et 24 mars 2000, les Chefs d'Etat et de Gouvernement ont constaté que « *l'Union européenne se trouve face à un formidable bouleversement induit par la mondialisation et par les défis inhérents à une nouvelle économie fondée sur la connaissance* ».

Les effets induits dans le monde du travail par les outils des technologies de l'information et de la communication sont soulignés aussi par Michel Serres qui constate la difficulté de s'exprimer sur le métier de quelqu'un qui travaille devant l'ordinateur. Plusieurs articles ont proposé des listes de métiers qui n'existaient pas dans la décennie antérieure (e.g. architecte d'information, manager des communautés en ligne, développeur d'applications mobiles) (« 10 Jobs That Didn't Exist 10 Years Ago - Forbes », 2012, « Great New Jobs That Didn't Exist 10 Years Ago », 2011). Dans notre société, l'information et la connaissance sont considérées comme les ressources premières de son développement (« TAG - Fiche du terme : société de l'information », s. d.).

⁴³ Expression mise en circulation par la Commission européenne en 1993

⁴⁴ Titre d'un rapport de l'Unesco, début des années 2000

1.3.1. Données, informations et connaissances

Amplifiés par le contexte technologique actuel, les concepts de donnée, d'information et de connaissance sont devenus essentiels à notre société. La connaissance est définie en termes d'information, et l'information en termes de donnée. Ces trois concepts associés à celui de sagesse complètent la hiérarchie : Donnée-Information-Connaissance-Sagesse (DICS)⁴⁵. Initialement proposée par R.L. Ackoff (1989), la *pyramide de la sagesse* contient 5 niveaux distincts : donnée, information, connaissance, compréhension et sagesse (v. Figure 1-11).

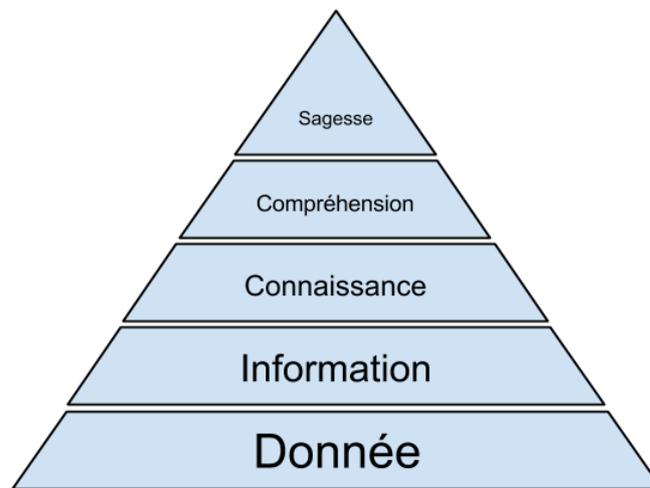


Figure 1-11. Hiérarchie donnée-information-connaissance-compréhension-sagesse d'Ackoff

Même si cette hiérarchie est associée à R.L. Ackoff, ses éléments sont déjà préfigurés dans un article de M. Zeleny (1987), ou dans d'autres formes plus poétiques⁴⁶. Sur l'axe temporel, la donnée, l'information, la connaissance et la compréhension existent dans le passé du sujet (v. Figure 1-12). La compréhension est une étape distincte dans la transition vers la sagesse. Résultant d'une compréhension de la connaissance, la sagesse permet la projection dans l'avenir.

M. Zeleny associe le concept de *donnée* au *connaît-rien* (know-nothing) et au traitement électronique de données. *L'information*, c'est le *connaît-comment* (know-how) et elle est

⁴⁵ En anglais : DIKW hierarchy : Data – Information – Knowledge – Wisdom. Elle est aussi connue sous les noms de « pyramide/hiérarchie de la sagesse » ou « pyramide/hiérarchie de la connaissance ».

⁴⁶ *Information is not knowledge/Knowledge is not wisdom/Wisdom is not truth/Truth is not beauty/Beauty is not love/Love is not music/and Music is THE BEST.* The 1979 song 'Packard's Goose' by Frank Zappa, on the Album 'Joe's Garage Act II and III contains the lines: (Tower Records, 1979) <http://www.youtube.com/watch?v=kQHMKleiN0&t=3m40s>.

associée aux systèmes de management d'information. *La connaissance*, c'est du *connaît-quoi* (know-what) et elle est associée à l'intelligence artificielle et aux systèmes décisionnels. *La sagesse* est le *connaît-pourquoi* (know-why) et elle est associée aux systèmes de management humain et aux systèmes de support de management (ibid., p. 60). Il attire l'attention sur le fait que, dans la société et dans l'interaction humaine, la classification se complique. Prenons l'exemple des données, définies comme *réalités objectives*. Dans la vie réelle, ces *réalités objectives* ne sont pas toujours les mêmes pour tout le monde. Ce qui se qualifie comme *donnée* pour quelques-uns peut se qualifier comme *information* pour d'autres. Dans une vision de modélisation de connaissances, M. Zeleny définit l'expression *objets de connaissance* comme des entités distinctes de connaissance. L'organisation et la structure des *objets de connaissances composées* se réalisent avec des relations de type : « est-un » (is-a), « partie-de » (part-of), « à-pour-part » (has-part)⁴⁷.

Après R.L. Ackoff, différents auteurs (Bellinger, Castro, et Mills, 2004; Burton-Jones, 1999; D. Clark, 2004; Markus, 2005) reprennent ces concepts et en proposent une hiérarchie sur quatre niveaux, la compréhension n'étant plus un niveau distinct, mais elle est toujours présente dans la construction des niveaux (v. Figure 1-12). Pour R.L. Ackoff, la compréhension est en relation avec le « pourquoi ».

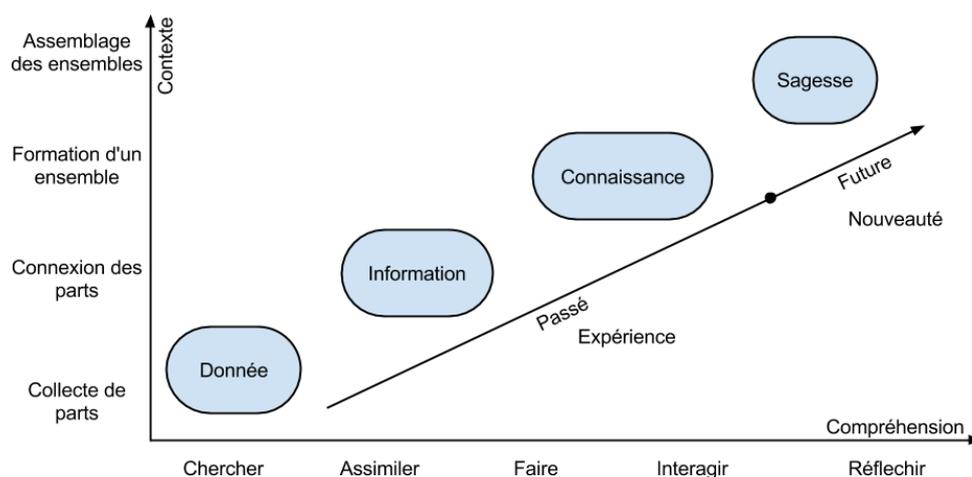


Figure 1-12. Hiérarchie DICS (selon D. Clark, 2004)

⁴⁷ Ces types de relations sont beaucoup utilisés dans la construction des ontologies. La relation *est-un* (is-a) est la relation de base qui permet la classification des instances dans les hiérarchies des classes dans les ontologies (rdfs:subClassOf).

On distingue trois concepts de base essentiels dont la valeur est définie par leur degré d'interconnexion et leur niveau de compréhension. Pour A. Burton-Jones (1999), le niveau de compréhension permet de faire accéder la donnée au niveau de l'information et finalement à celui de la connaissance.

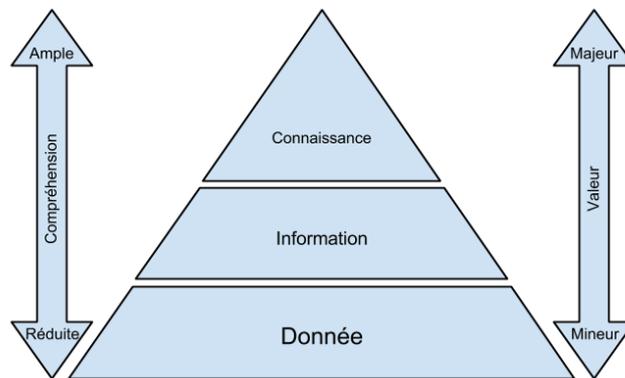


Figure 1-13. Donnée, information et connaissance (selon *ibid.*)

Pour faire une synthèse des définitions (Zeleny, 1987; Ackoff, 1989; Burton-Jones, 1999; Rowley, 2008; Davenport et Prusak, 1998) ces trois concepts sont définis de la façon suivante :

- donnée⁴⁸ : « discrète, brute, réalités objectives, observations » ; « symbole qui représente une propriété d'un objet, d'un événement ou de son environnement » ; « le produit de l'observation » mais elle n'est utilisable que si elle est dans une forme adéquate, pertinente ; les données ne sont pas organisées ou traitées ; elles n'ont ni sens ni valeur car elles sont dépourvues de contexte et d'interprétation ;
- information⁴⁹ : « donnée organisée, structurée et traitée » ; « contenu dans des descriptions » ; « répond aux questions de type *qui, quoi, où...* » ; « traité de façon à être pertinente pour un but spécifique ou contexte », « plus de valeur, utile » ; les systèmes d'information génèrent, stockent, récupèrent ou traitent les données ; la différence entre la donnée et l'information est *fonctionnelle* et non structurelle ;

⁴⁸ Le mot « *donnée* », du mot latin *datum*, signifie « ce qui est donné », « cadeau »⁴⁸; la donnée peut être considérée comme *ce qui est admis comme base de raisonnement ou d'inférence* ou comme *ce qui sert de base* de calcul ou de mesure.

⁴⁹ Le mot latin « *informare* » : instruire

- connaissance⁵⁰ : se définit par référence à l'information; le *connaît-comment* rend possible la transformation de l'information en instruction ; « synthèse de plusieurs sources d'information dans un intervalle de temps » ; obtenue par instruction (de quelqu'un qui la possède), par étude ou par expérience ; présente non seulement dans les documents, mais aussi dans les routines, processus, pratiques et normes.

Quand à la sagesse, elle est une mesure d'évaluation de la compréhension. Elle a la capacité d'augmenter l'efficacité et de créer de la valeur par l'application de fonctions mentales, par exemple de jugement.

Pour revenir sur le concept de la connaissance dans le domaine des sciences humaines, J.F. Le Ny la définit comme « croyance vraie justifiée, ou plus généralement représentation adéquate justifiée » (Le Ny, 1989, p. 26-27). Une première classification de la connaissance distingue deux grandes catégories : connaissances *déclaratives* et connaissances *procédurales* (Raynal et Rieunier, 2010, p. 120-121).

Les connaissances déclaratives correspondent aux connaissances théoriques qui facilitent la compréhension et l'explication du monde. Elles sont relatives aux données, aux informations, aux faits, aux termes. Elles sont en relation avec le *savoir-que*⁵¹.

Les connaissances procédurales permettent à une personne « d'agir dans son environnement ; elles correspondent à ce que plusieurs enseignants appellent les habiletés ou le savoir-faire. » (Viau, 2005, p. 14) Les connaissances procédurales sont relatives aux procédures, elles sont en relation avec le *savoir-comment*⁵² ou *savoir-faire*⁵³ (Raynal et Rieunier, 2010, p. 122; Richard, 1990, p. 18).

⁵⁰ J. Locke affirmait que « La connaissance est la perception de l'accord ou du désaccord de deux idées » (Locke, 1689)

⁵¹ *Savoir-que (knowing-about)* représente les notions de base dans un domaine, les concepts d'initiation dans une discipline. Exemples : *savoir-que* l'un des angles du triangle rectangle a une valeur de 90° ; *savoir-que* la somme des angles d'un triangle est de 180°.

⁵² *Savoir-comment (know-how)* : calculer l'hypoténuse d'un triangle rectangle en connaissant les deux autres côtés du triangle.

⁵³ *Savoir-faire (knowing-to do)* : conduire une voiture, résoudre un problème de mathématique; le *savoir-faire* permet l'accomplissement d'une tâche.

Les nuances de la connaissance se définissent aussi en fonction du type du savoir emmagasiné. Aujourd’hui le *savoir-que*, le *savoir-faire* et le *savoir-être*⁵⁴ doivent être complétées par le *savoir-où*⁵⁵ et le *savoir-transformer* (e.g. ajuster, adapter, recombinaison, innover) (Siemens, 2006, p. 10). Siemens affirme que le *savoir-où* est plus important que le *savoir-que* et le *savoir-comment* (Siemens, 2005).

Une autre classification (v. Figure 1-14) de la connaissance selon la possibilité d’articulation de cette connaissance est réalisée sous trois catégories : explicite, tacite et implicite (Cortada et Woods, 2000, p. 14).

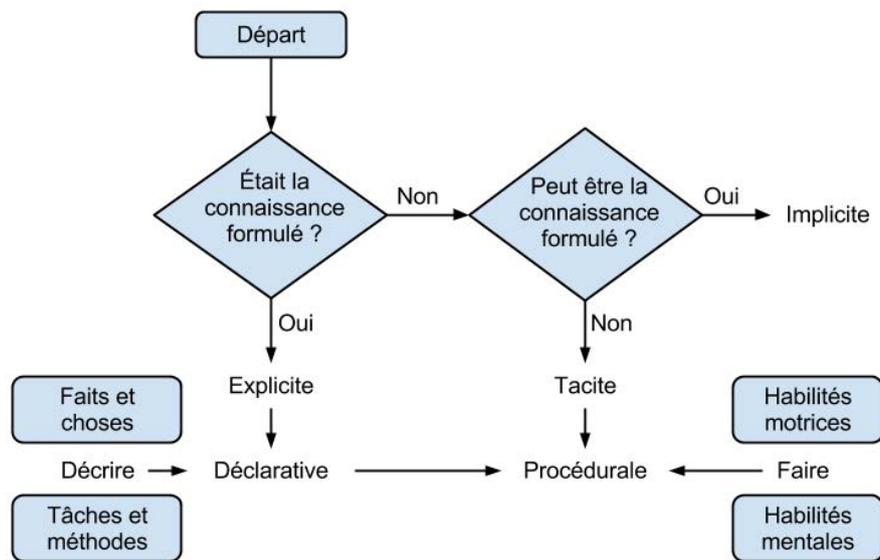


Figure 1-14. Relations entre les types de connaissances (selon *ibid.*, p. 18)

La connaissance explicite est une connaissance qui peut être articulée, captée sous forme de textes, tableaux, diagrammes. Elle est « formelle et systématique » (Nonaka et Takeuchi, 1995). Au contraire, la connaissance tacite, comme le note M. Polanyi (1966) qui propose le terme (et l’illustre par l’expression : « Nous savons plus que nous pouvons exprimer »), ne peut pas être articulée⁵⁶. La connaissance tacite est dans

⁵⁴ *Savoir-être* : définit la capacité d’agir ou de réagir de manière adaptée à l’environnement social ou physique.

⁵⁵ *Savoir-où* : trouver la connaissance quand nécessaire, chercher sur le web, dans les bibliothèques

⁵⁶ « Polanyi utilise l’exemple de la capacité à reconnaître le visage d’une personne, tout en étant seulement vaguement capable de décrire comment cela s’opère. Ceci est un exemple de reconnaissance de formes. Ce que nous reconnaissons est la totalité ou la « gestalt » et on échoue

l'action, dans le « faire ». La connaissance qui peut être articulée, mais qui ne l'a pas été est la connaissance implicite. Elle est impliquée, déduite ou inférée du comportement observable ou de la performance.

L'acquisition des connaissances auprès de l'apprenant est le but principal de l'apprentissage. L'apprenant doit s'approprier un contenu et construire des connaissances, qui sont bien plus que des données ou des informations, afin d'être capable de les utiliser dans les aspects concrets de sa vie professionnelle ou personnelle.

Pour faciliter l'association des connaissances aux objectifs d'apprentissage, D.R. Krathwohl (2002, p. 214) classe la connaissance en quatre catégories : (1) *connaissance factuelle* (e.g. éléments de base, terminologie, détails spécifiques); (2) *connaissance conceptuelle* (e.g. relations entre les éléments de base, classifications, théories); (3) *connaissance procédurale* (e.g. comment faire quelque chose, habiletés, méthodes, procédures) et (4) *connaissance métacognitive* (e.g. connaissance stratégiques, connaissance des tâches cognitives, connaissance de soi). Cette classification de la connaissance est utilisée en relation avec la définition des objectifs d'apprentissage selon la taxonomie de Bloom (voir le paragraphe 5.2.3).

1.3.2. Représentation et ingénierie des connaissances

J.Y. Prax affirme dans « Le guide du knowledge management » que : « À l'échelle du développement des nations, on peut affirmer que le savoir est devenu le principal déterminant de la croissance et de la qualité de la vie, avant la terre, les machines et le travail » (Prax, 2000, p. 3). La société de l'information⁵⁷ ou de la connaissance⁵⁸ a compris l'importance du savoir et de la gestion de la connaissance : elle doit s'adapter aux exigences imposées par ces nouveautés.

à la décomposer en ses éléments constitutifs afin d'être en mesure d'exprimer son essence » (Cortada et Woods, 2000, p. 14)

⁵⁷ http://ec.europa.eu/information_society/eeurope/2002/news_library/pdf_files/french.pdf

⁵⁸ Basée sur la société de l'information, elle met l'accent sur l'importance du savoir, de la connaissance, de la créativité et de l'innovation, en plus de l'information

Après l'invention de l'imprimerie, et jusqu'à l'invention de l'ordinateur, le livre⁵⁹ est le support de base pour le stockage et la transmission des données et des informations. L'imprimerie a fait exploser la quantité des livres existants et a donc imposé des stratégies d'organisation adaptées. La « *bibliologie* »⁶⁰ place le livre au centre de ses études. La bibliothéconomie (e.g. organisation des bibliothèques), la bibliométrie (e.g. méthodes statistiques appliquées aux livres), le catalogage (e.g. indexation, archivage) sont quelques domaines qui se rapportent à l'effort d'organisation et d'amélioration de l'accès au savoir. L'enrichissement du livre dans l'espace virtuel renforce encore les capacités de stockage et de transport du savoir.

Lors du stockage du savoir, les informations sont enregistrées sans qu'elles aient nécessairement des relations entre elles. À ce niveau, l'intelligence humaine est l'agent de raisonnement qui permet d'établir la validité et la vérité des informations. L'extension de cette capacité de raisonnement à des agents et systèmes informatiques conduit au domaine de l'Intelligence Artificielle (IA)⁶¹.

La *représentation des connaissances* est un sous-domaine de l'IA. Elle est concernée par la façon dont la connaissance peut être représentée symboliquement et manipulée d'une façon automatique par les agents de raisonnement (Brachman et Levesque, 2004). La représentation des connaissances couvre les outils, les techniques et les technologies utilisés pour manipuler les connaissances : logique, règles, réseaux sémantiques, etc.

En s'efforçant de définir la *représentation des connaissances*, R. Davis (1993) lui attribue cinq caractéristiques ou rôles essentiels distincts. Elle a ainsi toutes ces caractéristiques à la fois.

1. Un rôle de substitution : la représentation des connaissances opère avec des *représentations*⁶². Ces représentations sont des *substituts*, des « remplaçants » des

⁵⁹ Biblio – du grec *biblion* – « livre » (Le Robert, Rey-Debove, et Rey, 2010, p. 248)

⁶⁰ Bibliologie – Science du livre, du texte et de l'écrit. Mot constitué de deux éléments : biblio- : « livre » et logos : « discours, récit, parole »

⁶¹ Terme créé dans les années 1950 par John McCarthy (1927 – 2011).

⁶² Les représentation ne sont pas complètes (voir aussi le cas de systèmes) : elles ne sont pas parfaites et toute imperfection peut être une source d'erreur (Davis, Shrobe, et Szolovits, 1993, p. 19).

choses qui existent dans le monde (e.g. référant). La dichotomie *substitut – référant* est l'essence de la représentation des connaissances.

2. Un ensemble de déclarations, *d'engagements ontologiques* : il s'agit d'un ensemble des décisions relatives à ce qui va exister et comment cela va exister dans la représentation du monde. Le choix de telle ou telle autre ontologie peut conduire à des points de vue très variés, voire différents. L'essentiel n'est pas dans le langage dans lequel l'ontologie est écrite, mais dans les composantes et les connexions qui les relient.
3. Une *théorie fragmentaire de raisonnement intelligent* : les définitions du *raisonnement intelligent* sont construites sur les différentes conceptions qui caractérisent les différents domaines concernés : logique mathématique (e.g. calcul formel, déduction), psychologie (vue comme caractéristique humaine), biologie (l'architecture de la machine, une caractéristique comportementale de type stimulus-réponse qui émerge de l'interconnexion parallèle d'une grande collection des processeurs), statistique (respecter les axiomes de la théorie de la probabilité, l'incertitude) et économie (valeurs, préférences). La *représentation des connaissances* est une *théorie fragmentaire* parce qu'elle intègre seulement une partie de la connaissance et cette connaissance à son tour représente une partie d'un phénomène complexe de raisonnement intelligent.
4. *Un environnement efficace pour le calcul*: le raisonnement étant un processus computationnel, la représentation de connaissances doit offrir un cadre favorable au raisonnement.
5. *Un environnement pour l'expression humaine* : la représentation de connaissances est un moyen dont nous exprimons des choses sur le monde, un moyen dont nous décrivons le monde aux machines.

Dans cette énumération, nous notons le rôle central des agents intelligents qui ont la capacité d'*inférer*⁶³ sur un ensemble de connaissances. Acteurs clés dans les systèmes à

⁶³ Les moteurs de raisonnement ont la capacité de d'enrichir la base de connaissances à travers l'*inférence*. En partant de propositions tenues pour vraies (prémises, connaissances explicites), l'*inférence* permet la déduction des nouvelles propositions (conclusions, connaissances implicites). Les moteurs de raisonnement permettent de valider la pertinence de la base de connaissances. Si du processus de raisonnement résultent des propositions contradictoires, la base de connaissances est déclarée inconsistante ou incohérente. L'ajout des nouvelles propositions ne doit pas contredire l'ensemble des propositions existant dans la base.

base de connaissances, ils ont pour but de prendre en charge des tâches censées être résolues par l'homme. Ces systèmes optimisent le travail intellectuel et surmontent la complexité des informations prises en compte. « Dans cette optique, l'ingénierie des connaissances mobilise des technologies cognitives pour réaliser des systèmes techniques instrumentant le travail cognitif et intellectuel et outillant la connaissance et la pensée » (Bachimont, 2004a, p. 4). L'ingénierie des connaissances est aussi « une technologie, dont la technique (*techno-*) procède de deux savoirs scientifiques (*-logie*), un savoir nomologique de modélisation formelle, un savoir idiographique critique de l'interprétation sémiotique » (Bachimont, 2004b, p. 56).

L'ingénierie des connaissances consiste à intégrer la connaissance dans les systèmes informatiques afin de résoudre des problèmes complexes nécessitant normalement un haut niveau d'expertise humaine (Feigenbaum et McCorduck, 1983). Les systèmes résultants sont appelés des *systèmes experts* et ils sont spécialisés dans des domaines spécifiques dont les connaissances sont *représentées*.

B. Bachimont affirme que « le caractère souvent réducteur des approches relevant des sciences cognitives et de l'intelligence artificielle a conduit à développer des recherches portant davantage sur les systèmes assistant le travail intellectuel⁶⁴ que sur la nature de l'intelligence elle-même » (Bachimont, 2004a, p. 4).

Dans l'ingénierie des connaissances, la représentation des connaissances est considérée sous les deux sens : *représentation* – la connaissance est représentée par des symboles – et *interprétation* – les symboles représentent des connaissances. Si l'ingénierie⁶⁵ documentaire⁶⁶ a comme objet central *le document*⁶⁷ constitué d'information et d'un

⁶⁴ Un exemple est fourni par les systèmes d'aide à la décision.

⁶⁵ En anglais *engineering*, « celui qui construit des *engins* et fait preuve de son talent ou de son *ingenium* » (Bachimont, 2004a, p. 3). Terme introduit en 1964 : engineering -> génie. Conception, étude globale d'un projet industriel sous tous ses aspects (techniques, économiques, financiers, sociaux) (Le Robert et al., 2010, p. 1330).

⁶⁶ L'ingénierie documentaire consiste dans (1) la mise en place et l'administration de production documentaire (produits documentaires : bulletins, revues de presse, liste de nouveautés etc.) et (2) la mise en place et l'administration des services d'information et documentaire (Guyot, 2011, p. 5).

⁶⁷ Objet documentaire ou document : « supports de données enregistrés à des fins de preuve ou de renseignement. Il possède deux niveaux constitutifs : l'*information* et le *support d'enregistrement* » (ibid., p. 20).

support, l'ingénierie des connaissances est « en fait une *ingénierie des supports de connaissances* » (ibid., p. 5). Toutefois, l'ingénierie documentaire est elle-même une partie essentielle de l'ingénierie des systèmes d'information (Zacklad, 2005). Rappelant le rôle de représentation de la connaissance, B. Bachimont affirme que « la connaissance n'est pas un objet et qu'elle ne s'appréhende qu'à travers des objets dont elle est l'interprétation » (Bachimont, 2004a). Une des illustrations très parlante de l'idée de représentation est le tableau intitulé *La trahison des images* de René Magritte (v. Figure 1-15).



Figure 1-15. *La trahison des images*, de René Magritte

Ces objets, appelés *inscriptions*, sont des objets physiques et matériels intentionnels, dont l'usage « renvoie à une construction interprétative selon des normes liées au contexte de l'interprétation » (ibid., p. 5). Une inscription est donc un objet intentionnel considéré « non pas pour ce qu'il est, sa constitution physique propre, son apparence matérielle, mais pour ce qu'il n'est pas : ce qu'il signifie, ce qu'il représente » (ibid., p. 5).

Le système d'apprentissage intègre donc une représentation des caractéristiques des entités qu'il gère (e.g. l'apprenant, l'objet pédagogique). Les représentations de ces entités permettent au système d'apprentissage d'inférer et de prendre des décisions sur le déroulement de l'apprentissage de l'apprenant. Pour B. Bachimont la connaissance est « la capacité d'exercer une action pour atteindre un but » (ibid., p. 65) et elle possède un objet, car « tout connaissance est connaissance de quelque chose, d'un *objet* ». La connaissance des entités participant à l'apprentissage fournit au système la capacité d'exercer des actions pour atteindre des buts spécifiques (e.g. personnalisation du parcours d'apprentissage, amélioration de l'expérience d'apprentissage).

1.3.3. Objet pédagogique

L'*objet*⁶⁸ *pédagogique* est un concept central de l'apprentissage assisté par l'ordinateur. Cette abstraction permet d'adapter le contenu à apprendre grâce à une forme accessible et convenable pour l'apprentissage. C'est à travers des objets pédagogiques que le savoir prend forme et interagit avec l'apprenant. L'objet pédagogique est l'unité porteuse du savoir, il sert à la construction de la connaissance de l'apprenant. C'est une « brique pédagogique » ayant un objectif d'apprentissage et qui permet l'encapsulation d'un contenu autonome. Le terme d'*objet* renvoie à la notion de *granularité*. L'assemblage de ces granules construit *le cours*. L'objet pédagogique constitue l'unité centrale dans l'organisation du contenu d'apprentissage.

Plusieurs définitions ont été données pour définir le concept d'objet pédagogique. La plus répandue est celle donnée par le *Learning Technology Standards Committee d'IEEE*⁶⁹ (IEEE LTSC). Il définit l'objet pédagogique comme étant « toute entité, numérique ou non-numérique, qui peut être utilisée pour l'apprentissage, l'enseignement ou la formation » (IEEE, 2002, p. 5).

En rappelant l'idée centrale de l'objet pédagogique, à savoir « décomposer le contenu d'apprentissage en petits morceaux qui peuvent être réutilisés dans différents environnements d'apprentissage », D.A. Wiley affirme que l'objet pédagogique est « toute ressource numérique qui peut être utilisée pour soutenir l'apprentissage » (Wiley, 2003)

Cependant, ces définitions étant très générales, toute ressource peut être un objet pédagogique, A.C. Laverde *et al.* (2007, p. 675) résumet et proposent la définition suivante de l'objet pédagogique : « entité numérique, autonome et réutilisable avec un objectif d'apprentissage clair qui contient au moins trois composants internes : le contenu, les activités pédagogiques et des éléments de contexte. En complément, l'objet d'apprentissage devrait comporter une composante externe d'information qui permet son identification, son stockage et sa récupération : les métadonnées ».

⁶⁸ La notion d'*objet* renvoie à la *programmation orientée objet*. Ce paradigme de programmation est fondé sur la modularité, la flexibilité et la réutilisabilité des *objets* (composantes d'une application informatique ou instances des classes).

⁶⁹ Institute of Electrical and Electronics Engineers

L'Agence Universitaire de la Francophonie (AUF) utilise plutôt le terme d'*objet d'apprentissage* qu'elle définit comme « un bloc de contenus qui focalise sur un objectif d'apprentissage spécifique. Ces objets d'apprentissage peuvent être constitués d'un ou de plusieurs composants ou blocs d'informations incluant du texte, des images, de la vidéo, des sons, des animations, etc. » (Simard, 2002).

R. McGreal (2004) propose une synthèse des définitions formulées par les publications en langue anglaise sur le sujet et constate une grande diversité d'appellations.⁷⁰ Il propose une caractérisation en quatre niveaux (v. Figure 1-16).

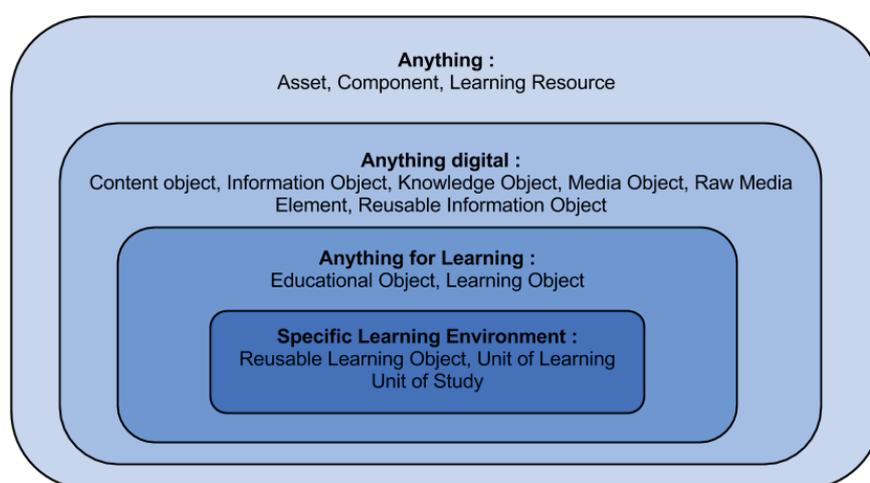


Figure 1-16. Terminologie de l'objet pédagogique, adaptée selon (ibid.)

Les termes utilisés sont distribués sur quatre classes à partir du niveau le plus général, jusqu'au niveau le plus restrictif : (1) tout objet ; (2) tout objet numérique ; (3) tout objet pour l'apprentissage (objet créé autour d'un objectif d'apprentissage spécifié) et (4) les objets d'apprentissage spécifiques (objet pédagogique créé pour des plateformes utilisant des standards spécifiques).

⁷⁰ « **Asset** (Wiley 2000); **Component** (Ip et. al., 1997; Koutlis et. al., 1999; Merrill, 1983; Quinn and Hobbs, 2000; Roschelle et. al., 1999); **Content object** (ADL, 2003; OASIS, 2003; Shabajee, 2002; Slosser, 2001); **Educational object** (Ilich, 1971; Friesen, 2001; EOE, 2003) ; **Information object** (Wiley, 1999; Epsilon Learning Systems 2003; Wieseler, 1999) ; **Knowledge object** (Merrill, 1999; Paquette and Rosca, 2002) ; **Learning object** (generic) term credited to W. Hodgins (Jacobsen, 2001) ; **Learning resource** (IMS Global Learning Consortium, 2000; Papatheodorou et. al., 2002; Koper, 2003; Paquette and Rosca, 2002) ; **Media object** (ADL, 2001; Shabajee, 2002) ; **Raw Media Element** (CanCore, 2003); Duval & Hodgins, in press) ; **RIO** (Reusable Information Object) (CISCO, 1999; Wieseler, 1999) ; **RLO** (Reusable Learning Object) (Cisco Systems, 2001; Barritt and Lewis, 2002; MERLOT, 2002) ; **Unit of Learning** (Sloep, in press) ; **Unit of Study** (Koper, 2001) »(McGreal, 2004).

Dans le compte rendu de nos travaux, nous utiliserons le terme *d'objet pédagogique*⁷¹, pour désigner toute entité numérique qui peut être utilisée dans l'apprentissage, sa caractéristique principale étant l'existence et la relation avec un objectif d'apprentissage.

Sous un certain aspect, ces classes sont liées au niveau de l'agrégation des objets pédagogiques. En fait, une des difficultés de la conception des objets pédagogiques procède de ses avantages mêmes : *la granularité*. Quelle taille (e.g. quantité du contenu d'apprentissage, temps de parcours) devrait avoir un objet d'apprentissage ? Plusieurs modèles formant un cadre permettant de définir la granularité de l'objet pédagogique ont été proposés. Une synthèse de ces modèles est présentée dans le Tableau 1-2.

Le IEEE LTSC propose 4 niveaux de granularité :

1. un niveau minimum d'agrégation : les données médias bruts, fragments (e.g. images, texte, clips audio, clips vidéo) ;
2. une collection d'objets de niveau 1 d'apprentissage (e.g. leçon) ;
3. une collection d'objets de niveau 2 d'apprentissage (e.g. cours) ;
4. un niveau maximum d'agrégation (e.g. ensemble de cours qui mènent à un certificat). Ce niveau peut contenir des objets de niveau trois ou, de manière récursive, des objets de niveau quatre.

⁷¹ Dans cette thèse les termes *objet pédagogique* et *objet d'apprentissage* sont interchangeables et expriment la même chose.

Tableau 1-2. Sommaire de modèles de contenu et niveaux de granularité de l'objet pédagogique (selon Balatsoukas, Morris, et O'Brien, 2008)⁷²

	SCORM	IEEE LTSC LOM	Learnativity	Cisco RIO, RLO	Les LCMS	DHER & LO
Niveau faible d'agrégation	Ressources	Niveau 1 Donnée brute	Ressources	Sous-thème	Ressources	Objets d'information
	SCOs	Niveau 2 d'agrégation (Leçon)	Objets d'information Objets pédagogiques	Thème (RIO) Leçon (RLO)	Objets pédagogiques	Ressources d'information Objets pédagogiques
Niveau élevé d'agrégation	Agrégation du contenu (e.g. Leçon, Cours)	Niveau 3 d'agrégation (Cours)	Composante d'apprentissage	Module	Unité d'apprentissage	Unité d'étude
		Niveau 4 d'agrégation (Cours certifié)	Environnement d'apprentissage	Course		Module Cours Collection

Le niveau de granularité est inversement proportionnel à la capacité de réutilisation de l'objet pédagogique (v. Figure 1-17).

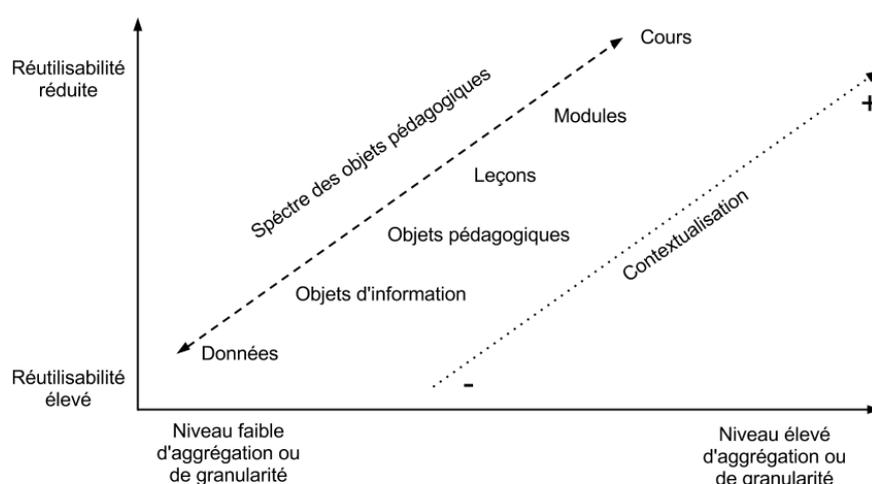


Figure 1-17. Niveau de granularité – réutilisabilité (selon ibid.)

⁷² RIO : Reusable Information Object ; RLO : Reusable Learning Object ; SCO : SCORM Object.

Les niveaux de granularité des objets pédagogiques donnent leurs dimensions en terme d'espace et de temps. L'espace dans le sens du nombre d'objets d'information contenus dans un objet pédagogique⁷³; le temps considéré la durée nécessaire au parcours de cet objet pédagogique. Mettant en évidence la dimension temporelle, Wisconsin Online Resource Center⁷⁴ caractérise l'objet pédagogique par une durée comprise entre 2 et 15 minutes (Beck, 2008).

Dans la Figure 1-18, est présentée la pyramide des contenus d'apprentissage. À la base se trouvent les données brutes et les objets d'information. De la composition de ces types d'objets résulte la création d'objets pédagogiques, qui à leur tour, peuvent se combiner pour devenir des contenus d'apprentissage avec un niveau élevé d'aggregation.

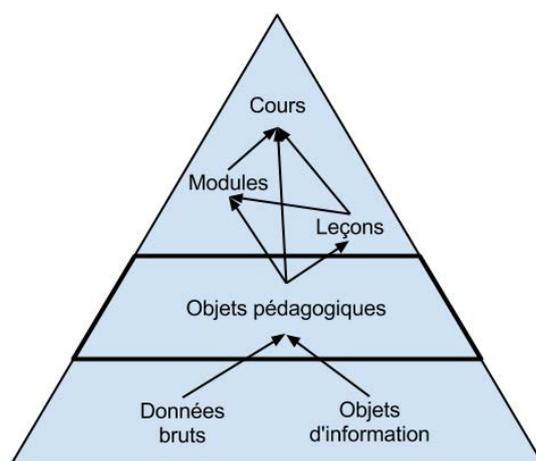


Figure 1-18. La pyramide des contenus d'apprentissage, selon (Balatsoukas *et al.*, 2008)

Afin d'illustrer la granularité et les possibilités de réutilisation des objets pédagogiques, deux métaphores ont été proposées : celle du jeu de LEGO et celle de l'atome (Wiley, 1999). L'analogie avec les pièces de LEGO permet de comprendre que : « (1) chaque brique (granule) est combinable avec n'importe quelle autre granule, de n'importe quelle façon ; (2) la combinaison de granules est une activité intuitive (amusante ?) à la

⁷³ Les spécialistes de Cisco Systems proposent qu'un objet pédagogique soit composé de cinq à neuf objets d'information (Balatsoukas, Morris, et O'Brien, 2008, p. 119).

⁷⁴ **Wisc-Online** est une bibliothèque numérique de ressources d'apprentissage basées sur le Web et appelées « objets pédagogiques ». Cette bibliothèque a été principalement développée par Wisconsin Technical College System (WTCS – <http://www.wisc-online.com/>).

portée de chacun ; (3) les constructions obtenues tiennent parfois debout, parfois non » (Roxin, 2003, p. 33; Wiley, 1999). Le principal défaut de cette analogie est qu'elle ne nous assure pas que le résultat de la combinaison des différents objets pédagogiques est *pédagogiquement utile* (Wiley, 1999). La deuxième métaphore considérant l'objet pédagogique comme un atome, cela génère plus des contraintes quant aux possibilités de combinaison : (1) chaque atome ne peut pas se combiner avec n'importe quel autre atome ; (2) les atomes ne peuvent se combiner que selon les formes prescrites par leur structure ; (3) enfin, un niveau de compréhension est requis pour combiner les atomes (ibid.).

Pour permettre la combinaison des objets pédagogiques, il est nécessaire d'ajouter une couche supplémentaire d'information : les métadonnées. On qualifie comme *métadonnées pédagogiques* les données décrivant les ressources pédagogiques (numériques ou non). Il s'agit de « descripteurs » - informations de nature *sémantique* (auteur, titre) - qui permettent de les rendre plus facilement identifiables (accessibles) et manipulables (interopérables, réutilisables, durables, adaptables). Chaque descripteur représente une caractéristique particulière de la ressource pédagogique et peut être mono ou multi-évalué, obligatoire ou optionnel.

La structure générale d'un objet pédagogique comporte le contenu d'apprentissage (e.g. objets d'information), les exercices (e.g. questions, questionnaires), l'évaluation et les métadonnées. Plus de détails sur les métadonnées sont présentés dans le paragraphe 2.2.3.

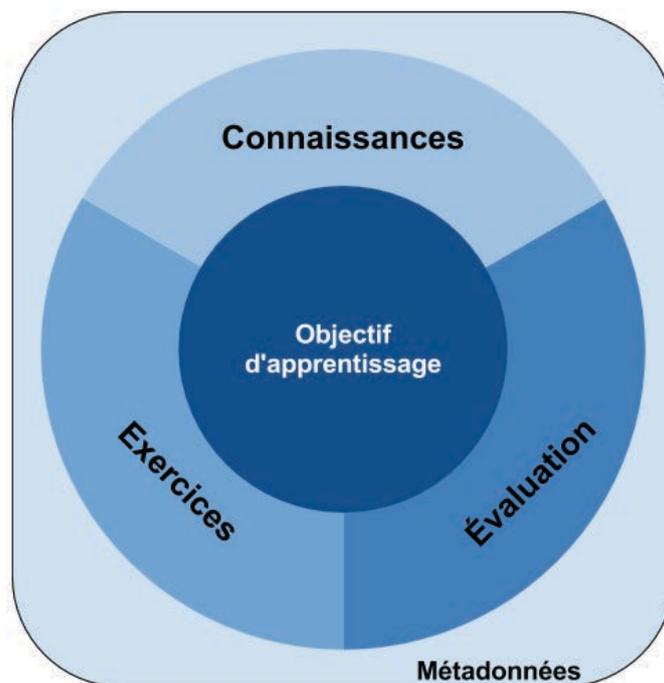


Figure 1-19. La structure générale d'un objet pédagogique (source Balog-Crişan, 2011, p. 24)

Pour S. Metros (Metros, 2005, p. 12), une entité numérique est un objet pédagogique si elle regroupe les éléments suivants : un objectif d'apprentissage, une activité d'exercice pratique et une évaluation. L. Mortimer (2002) ajoute à ces trois éléments le contenu d'apprentissage et les métadonnées. Nous distinguons donc maintenant les éléments que doit nécessairement intégrer une ressource numérique pour prétendre au statut d'objet pédagogique (v. Figure 1-19) :

1. contenu d'apprentissage (e.g. information) ;
2. objectif d'apprentissage (e.g. compétence à acquérir) ;
3. exercices, activités d'apprentissage (e.g. questions, questionnaires) ;
4. moyens d'évaluation (e.g. tests) ;
5. métadonnées (e.g. titre, auteur, niveau ciblé).

Dans cette thèse, nous qualifions *d'objet pédagogique* toute ressource numérique intégrant ces cinq composantes. Les autres types de ressources numériques susceptibles d'être utilisées dans un processus d'apprentissage sont appelés *ressources d'apprentissage* (e.g. articles wiki).

Le contenu d'apprentissage, les exercices et les moyens d'évaluation contiennent les informations que l'apprenant va transformer en connaissances. Ces trois composants, et

plus particulièrement le contenu d'apprentissage, contiennent le savoir à apprendre. Les métadonnées contiennent les informations qui permettent au système d'apprentissage d'opérer avec l'objet pédagogique en cause. Si un objet pédagogique porte une connaissance, les métadonnées sont une *connaissance sur cette connaissance* portée par l'objet pédagogique. Les systèmes d'apprentissage opèrent avec ce deuxième niveau de connaissance.

L'intégration et l'échangeabilité des objets pédagogiques dans les systèmes d'apprentissage nécessite l'existence et l'implémentation de mécanismes d'interopérabilité et de communication (Balatsoukas *et al.*, 2008, p. 119). Les *objets pédagogiques interactifs* (OPI), proposés dans (Guéraud *et al.*, 2004)⁷⁵, visant à stimuler l'apprentissage actif, sont caractérisés par un niveau élevé d'interaction avec l'apprenant. Contrairement à l'objet pédagogique *expositif*, l'OPI autorise la modification de son état interne et de sa représentation en fonction des manipulations de l'apprenant. Il permet la mise en place des *situations actives d'apprentissage* (e.g. simulations pédagogiques, exercices). Dans ces situations l'apprenant « interagit étroitement avec un dispositif l'induisant à une découverte-construction de connaissances » (ibid., p. 2). Les contraintes techniques d'un OPI imposent une intégration profonde avec le dispositif d'apprentissage. Dans une analogie avec l'architecture client-serveur, on peut considérer l'OPI comme le serveur et le dispositif d'apprentissage comme le client capable de recevoir les signaux générés par l'OPI.

Dans cette perspective et définition de l'objet pédagogique nous pensons que les technologies du Web sémantique peuvent apporter des solutions dans deux directions spécifiques. La première consiste à exprimer les métadonnées de l'objet pédagogique en utilisant les technologies du Web sémantique (Balog-Crişan, 2011; Nilsson, 2010). La seconde direction propose d'ouvrir le système d'apprentissage en offrant des points d'accès permettant aux objets pédagogiques de lancer des requêtes en utilisant les technologies du Web sémantique (e.g. SPARQL), sur la base des connaissances du système d'apprentissage.

⁷⁵ Dans cet article les auteurs présentent le projet FORMID (FORMation Interactive à Distance).

Chapitre 2. Systèmes et technologies de l'information et de la communication pour l'enseignement

*« Les systèmes ne sont pas dans la nature mais dans
l'esprit des hommes ».
J.L. Le Moigne*

Pour l'élaboration du prototype du système actif d'apprentissage, nous allons emprunter une méthode de modélisation utilisant des systèmes. La *systemique* nous permet d'aborder la modélisation de la complexité de notre système d'apprentissage sous une forme simplifiée tout en gardant une image holistique à travers la structure et les relations des éléments le composant.

Dans ce chapitre nous rappelons le concept de système et les notions de base de la pensée systemique. Considérant le Web comme système d'information et de communication de base, nous étudions le développement des systèmes basés sur le Web : des systèmes de création et de gestion de contenu ainsi que des systèmes d'apprentissage.

2.1. Systèmes et systemique

L'ingénierie pédagogique « tire son origine de la science des systèmes » (Le Moigne, 1995; Simon, 1974; cité par Paquette, 2002a, p. 107), elle se présente comme une « méthode systemique particulière vouée à la résolution des problèmes de conception des systèmes d'apprentissage » (ibid., p. 107). Les systèmes et la systemique sont à la base de l'ingénierie pédagogique, qui se trouve à l'intersection du design pédagogique, du génie logiciel et de l'ingénierie cognitive (v. Figure 2-1).

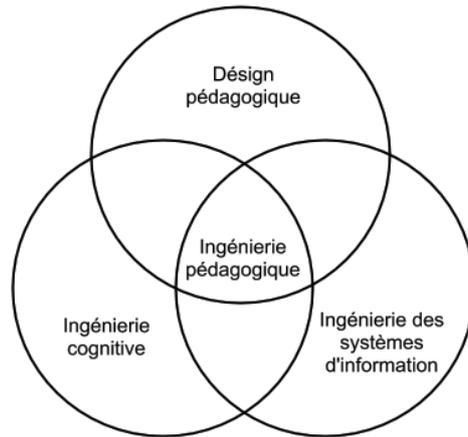


Figure 2-1. Bases de l'ingénierie pédagogique, selon (ibid.)

Les développements théoriques fondés sur le concept de système remontent loin dans le temps et trouvent des disciples retenant le concept de système bien avant l'élaboration spécifique de la théorie des systèmes. Déjà le philosophe grec Aristote formulait l'idée de système dans des termes proches de ceux qu'utilisera le fondateur de la théorie : « la totalité est plus que la somme des parties » (L. Von Bertalanffy, 1972, p. 407). Les théoriciens de la nature physique et les philosophes ont toujours pensé qu'il y avait une unité sous-jacente en-deçà des phénomènes, des manifestations et des objets de la nature.

Signalons le très révélateur *Traité des Systèmes* (1749) de l'abbé Étienne Bonnot de Condillac (1715-1780) qui pose les grands principes qui seront ceux de la future *systemique*. L'idée de système se retrouve encore dans les écrits de Kant (1724-1804), Hegel (1770-1831) ou Goethe (1749-1832). Tous reconnaissent une *unité conçue dans les groupes et organismes, une loi universelle de l'harmonie*. Le biologiste E.H. Haeckel (1834-1919) voit une *unité orchestrée par des lois naturelles et universelles*. Charles Darwin (1809-1882) écrivait dans une lettre au botaniste américaine Asa Gray⁷⁶ (1810-1888) le 22 Mai 1860 qu'il était *enclin à regarder tout comme résultant des lois désignées*⁷⁷.

⁷⁶ <http://www.darwinproject.ac.uk/namedef-1957>

⁷⁷ <http://www.darwinproject.ac.uk/entry-2814>

L'un des premiers travaux importants, reconnu comme un élément fondateur de l'élaboration de la théorie des systèmes, est l'ouvrage de Ludwig von Bertalanffy⁷⁸ (1901-1972), *General System Theory: Foundations, Development, Applications* (1968). Bertalanffy affirme que la *théorie générale des systèmes* est « une science générale de la totalité [ou de la *globalité*]⁷⁹ » (ibid., p. 37).

La théorie générale des systèmes est basée sur une approche holistique visant la modélisation d'un tout appelé *système*, composé de parties appelées *éléments*, qui peuvent être elles-mêmes considérées à leur tour comme des systèmes. Les éléments qui constituent le système sont organisés dans une *structure* qui permet leur agrégation et assure sa consistance. Cette approche favorise une vue globale de la totalité des éléments qui composent les systèmes : organismes, phénomènes ou objets complexes.

Dans les sections suivantes, nous présentons les notions de base de l'approche systémique ainsi que les concepts fondamentaux permettant de comprendre ce qu'est un système : *interaction, globalité, organisation et complexité* (AFSCET, 2003).

2.1.1. Système : boîte noire ou ensemble d'éléments en relations

Le concept de *système* est la notion centrale de la théorie des systèmes. Le mot *système* est issu du grec ancien *systema*, qui signifie *ensemble organisé, ensemble cohérent* (« Henry George Liddell, Robert Scott, A Greek-English Lexicon, σύστημα », s. d.). Le Petit Robert définit le système comme : « (1) *dispositif ou ensemble d'éléments ayant une fonction déterminée ; (2) ensemble de méthodes et de procédés organisés qui concourent au même résultat ; (3) ensemble des principes théoriques et des méthodes concrètes de fonctionnement et d'organisation (du domaine politique, économique ou autre) d'un pays sur le plan institutionnel ; (4) ensemble d'éléments (de l'organisme) de même nature qui assurent une fonction commune* » (Le Robert et al., 2010, p. 2490).

⁷⁸ Ludwig von Bertalanffy, d'origine autrichienne, biologiste de formation, s'intéresse tôt à la conception de l'organisme comme système ouvert.

⁷⁹ « General system theory, therefore, is a general science of "wholeness" [...] a logico-mathematical discipline, in itself purely formal but applicable to the various empirical sciences » (L. V. Bertalanffy, 1968, p. 37)

Selon Bertalanffy (L. V. Bertalanffy, 1968, p. 73) un système est un « *complexe d'éléments en interaction* ». Joël de Rosnay (1975, p. 91) définit le système comme « un ensemble d'éléments en interaction dynamique, organisés en fonction d'un but » (v. Figure 2-2). Les éléments du système sont en interaction les uns avec les autres et le système interagit avec l'environnement à travers des entrées et des sorties. La présence du *but* permet et détermine l'organisation du système. Contrairement au point de vue de la mécanique classique, la *téléologie*⁸⁰ trouve des implications dans la systémique. L'organisation du système et les relations des éléments dans le cadre du système lui permettent d'atteindre son but.

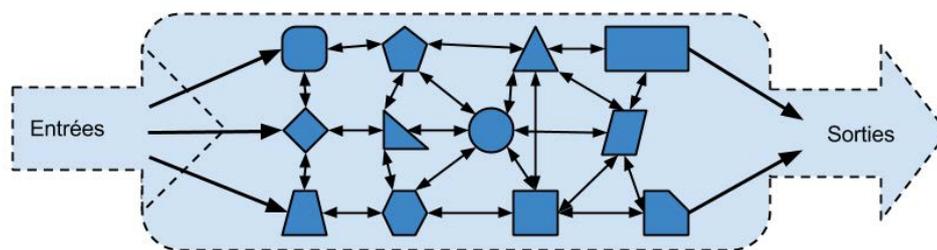


Figure 2-2. Système - ensemble d'éléments en interaction

En affirmant qu'« *un système est un unité globale d'interrelations entre éléments, actions et individus* », Edgar Morin (1977; cité par Donnadiou et Karsky, 2002, p. 29) met l'accent sur *la globalité* du système et sur *l'importance des relations* qui concernent aussi des *actions* sur les autres éléments du système. La notion de *globalité* d'un système sous-entend qu'un système doit être considéré dans son ensemble. Le système (S) est un ensemble des éléments (ϵ) en relation (σ) :

$$S = \{\epsilon, \sigma\} \quad \text{Équation 2-1}$$

Les relations apportent de la valeur au système. Bertalanffy soutient « *qu'on ne peut obtenir le comportement de l'ensemble comme somme de ceux des parties* » (« Concepts définissant un système | L'approche systémique », s. d.). Nous devons donc considérer les relations et les actions possibles entre les éléments du système pour comprendre le comportement du système. Cette propriété conduit au phénomène d'émergence : « les

⁸⁰ Téléologie : étude de la finalité. Le mot grec *telos* signifie « fin », « but » ou « objectif ». « Doctrine qui considère le monde comme un système de rapports entre moyens et fins ». (Le Robert et al., 2010, p. 2520).

systèmes présentent des qualités émergentes⁸¹ qui peuvent rétroagir sur les parties » (Morin, 2011, p. 123).

La disposition et les relations des éléments dans le système constituent l'organisation du système. Cette disposition peut être hiérarchique, dans des niveaux d'organisation, et peut permettre la décomposition d'un système en sous-systèmes. Dans l'organisation des systèmes, la notion de frontière doit être explicitée afin de permettre la composition des sous-systèmes (e.g. systèmes secondaires, modules). L'organisation est en relation avec l'information contenue dans le système. Selon Norbert Wiener⁸² « la somme d'information dans un système est la mesure de son degré d'organisation, l'entropie est la mesure de son degré de désorganisation ; l'un étant le négatif de l'autre » (Wiener, 1961, p. 11).

La considération du système dans son intégralité, son aspect *d'unité globale*, permet de l'utiliser sans connaître ses éléments composants. En fait, un système peut être lui-même élément d'une organisation supérieure. Dans ce cas, le système peut être vu comme une boîte noire. Les éléments du système ne sont pas visibles, mais nous connaissons sa fonction de transformation (f_T). Ainsi, il peut être considéré comme un transformateur de variables d'entrée (E_1, E_2, \dots, E_n) en variables de sortie (S_1, S_2, \dots, S_m), (v. Figure 2-3) (Donnadieu et Karsky, 2002, p. 38) et, dès lors, modélisé à travers une fonction de transformation :

$$\{S_1, S_2, \dots, S_m\} = f_T(E_1, E_2, \dots, E_n) \quad \text{Équation 2-2}$$



Figure 2-3. Système – boîte noire : fonction de transformation

⁸¹ « Ainsi l'eau a des qualités émergentes par rapport à l'hydrogène et l'oxygène qui la constituent. Par ailleurs, le tout est également moins que la somme des parties car les parties peuvent avoir des qualités qui sont inhibées par l'organisation propre au système » (Morin, 2011, p. 123).

⁸² Norbert Wiener (1894 – 1964) mathématicien, théoricien, chercheur et professeur au Massachusetts Institute of Technology est reconnu comme le fondateur de la cybernétique.

En fonction de leur capacité *d'interaction* avec leur environnement, les systèmes peuvent être ouverts ou fermés (v. Figure 2-4). Les systèmes fermés sont un cas particulier puisqu'ils ne permettent aucun échange avec l'extérieur. Ils sont presque inexistant dans le monde réel et leur représentation est seulement utilisée dans des approches théoriques.

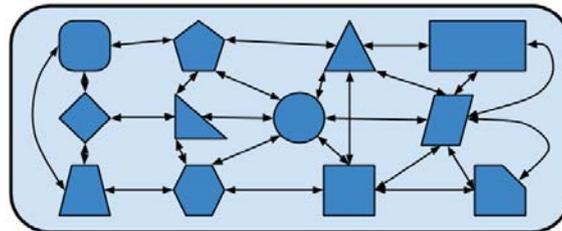


Figure 2-4. Système fermé

Dans le contexte de cette thèse, nous nous intéressons à la modélisation à travers des systèmes ouverts⁸³. En effet, ces systèmes ont la capacité d'interagir avec leur environnement : échange d'informations, d'énergie, etc. Ils peuvent s'adapter, s'intégrer, évoluer et influencer leur environnement.

Wiener souhaitait trouver les « lois communes à toutes les situations de contrôle et de régulation d'un système » (Heinderyckx, 2002, p. 16). S'interrogeant sur « la façon de prévoir l'état futur d'un système, à partir d'informations sur son état présent et passé », il introduisit la *boucle de rétroaction*, ou *feed-back* (v. Figure 2-5).

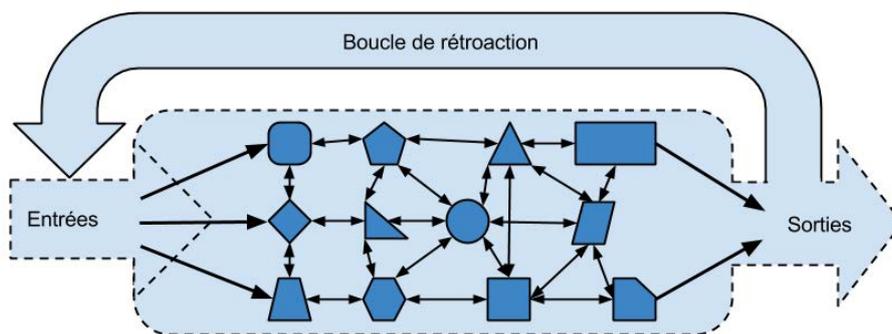


Figure 2-5. La boucle de rétroaction d'un système

⁸³ « Notre monde est intégralement constitué de systèmes, vivants ou non-vivants, imbriqués et en interaction. Peuvent ainsi être considérés comme des "systèmes" : une société, une économie, un réseau d'ordinateurs, une machine, une entreprise, une cellule, un organisme, un cerveau, un individu, un écosystème » (« Cybernétique, la science des systèmes », s. d.).

C'est un cas particulier d'interaction du système. La boucle de rétroaction permet d'ajuster l'entrée du système par rapport à la valeur de sortie. La boucle de rétroaction influence le comportement général du système et lui permet d'être finalisé et équilibré par le phénomène d'*autorégulation*.

En fonction du comportement résultant du système, la boucle de rétroaction peut être de deux types :

1. avec rétroaction *positive* ou *amplifiante* : caractéristique d'une dynamique du changement, la rétroaction positive facilite et accélère la transformation du système dans le même sens ; les effets sont cumulatifs (e.g. effet « boule de neige ») et le comportement du système s'accélère ;
2. avec rétroaction *négative* ou *stabilisante* : agissant en sens opposé aux résultats antérieurs du comportement du système, une boucle de rétroaction négative stabilise le comportement général du système jusqu'à ce qu'il atteigne une condition d'équilibre (CERTU, 2007, p. 22).

L'homéostasie⁸⁴ et la régulation des systèmes reposent sur l'efficacité de la boucle de rétroaction. Donnadieu et Karsky (2002, p. 47) définissent la régulation comme « *l'ensemble complexe des mécanismes d'ajustement que le système invente et met en œuvre en permanence pour maintenir son équilibre interne et dans le même temps, s'adapter à l'évolution de son environnement* ». La boucle de rétroaction introduit aussi une sorte de mémoire du système : l'état actuel du système est en relation avec son état d'évolution.

Dans le cadre de notre prototype de système d'apprentissage, la boucle de rétroaction est implémentée dans la fonction d'adaptation du contenu du parcours d'apprentissage en relation avec le retour obtenu de la part de l'apprenant : actions et activités, évaluation des résultats d'apprentissage, style d'apprentissage de l'apprenant.

⁸⁴ Homéostasie : terme créé par le physiologiste américain W.B. Cannon pour désigner la résistance au changement d'un système lorsque celui-ci réagit à une perturbation par une rétroaction de grandeur égale et de sens opposé (CERTU, 2007, p. 23).

2.1.2. Approche systémique : système et modélisation

Une des définitions de la systémique est la suivante : « *nouvelle discipline qui regroupe les démarches théoriques, pratiques et méthodologiques, relatives à l'étude de ce qui est reconnu comme trop complexe pour pouvoir être abordé de façon réductionniste, et qui pose des problèmes de frontières, de relations internes et externes, de structure, de lois ou de propriétés émergentes caractérisant le système comme tel, ou des problèmes de mode d'observation, de représentation, de modélisation ou de simulation d'une totalité complexe* » (AFSCET, 2003, p. 2-3).

En tant qu'adjectif, la *système* est définie en rapport avec le *système* : « *qui concerne un système ou qui agit sur un système* » (CNRTL, s. d.). Le nom *système* désigne une *méthode*, la pratique de la modélisation à travers les systèmes. L'approche systémique permet d'« entrer dans la complexité » des mondes à modéliser (AFSCET, 2003, p. 7), cette complexité étant déterminée par le nombre d'éléments (leur structure et fonction), le degré d'organisation du système, l'incertitude relative à son environnement et l'impossibilité de représenter tous les éléments et toutes les relations en jeu (CERTU, 2007, p. 18-19).

Pour Joël de Rosnay *l'approche systémique* « s'appuie sur la cybernétique et la théorie des systèmes » (Rosnay, 1975, p. 91). Toutefois elle présente aussi les caractéristiques suivantes :

- Elle dépasse et englobe l'approche cybernétique (N. Wiener, 1948) qui a pour but principal l'étude des régulations dans les organismes vivants et les machines construites par l'homme ;
- Elle se distingue de la *Théorie générale des systèmes* (L. von Bertalanffy, 1954) dont le but ultime consiste à décrire et à englober, dans un formalisme mathématique, l'ensemble des systèmes rencontrés dans la nature ;
- Elle s'écarte également de l'analyse des systèmes. Cette méthode ne représente qu'un des outils de l'approche systémique. Prise isolément, elle conduit à la réduction d'un système en ses composants et en interactions élémentaires.

Enfin, l'approche systémique n'a rien à voir avec une approche systématique, qui consiste à aborder un problème ou à effectuer une série d'actions de manière

séquentielle, détaillée, ne laissant rien au hasard et n'oubliant aucun élément. (ibid., p. 84).

Dans son état présent, la systémique résulte de trois générations de développement. La première, apparue au cours du 18^{ème} siècle, est principalement axée sur les systèmes matériels. On s'intéresse alors aux systèmes physico-chimiques, la nature elle-même étant un système assemblant des éléments (e.g. planètes, molécules). Cette génération de la systémique se construit autour du structuralisme, de l'organisation et de la régulation (CERTU, 2007; « Histoire de la Systémique », s. d.).

La systémique de deuxième génération est apparue dans les années 1920-1930. Le représentant le plus connu est L. von Bertalanffy. Il a formulé d'abord une « Théorie des Systèmes Vivants », puis la « Théorie Générale des Systèmes ». Bertalanffy s'intéresse aux systèmes biologiques : les organismes multicellulaires, les cellules vivantes, les écosystèmes, etc. Ce sont des systèmes ouverts qui échangent avec leur environnement de l'énergie, de la nourriture, de l'eau et de l'information. Ils s'organisent de façon autonome et hiérarchique (e.g. système nerveux). Ils ont besoin d'une boucle de rétroaction pour assurer les conditions d'équilibre à l'intérieur du système dans un environnement changeant. Ils ont leur propre mécanisme d'autorégulation (e.g. métabolisme, respiration) et leur finalité première est la survie. Ces systèmes donnent naissance à des propriétés émergentes. Ces propriétés ne sont pas réductibles aux éléments du système (e.g. composants chimiques ou physiques). Pour ces systèmes « la vie » est une propriété émergente.

Dans cette génération de systémique s'inscrit la cybernétique. Comme la science des systèmes autorégulés, la cybernétique s'intéresse à la modélisation des systèmes par l'étude de l'information et des principes d'interaction. C'est grâce à elle que des sciences comme l'informatique ou la robotique ont été inventées (ibid.).

La troisième génération de systémique, décrite dans les années 1970 par des chercheurs comme K.E. Weick et P. Checkland, répond aux questions liées aux systèmes qui n'ont pas une existence réelle. Le système n'est qu'une représentation, une carte mentale utilisée pour comprendre la réalité. C'est une approche adaptée aux systèmes sociaux

(e.g. entreprises, organisation, équipes, systèmes politiques, famille), qui sont des systèmes multipolaires. Weick affirme que la notion d'organisation « *en fonction d'un but* » est inadaptée à la description des systèmes sociaux (Weick, 1979). Ces systèmes possèdent plusieurs centres de décision, chacun ayant des intentions, des enjeux et des moyens d'action propres. En plus d'énergie, d'information ou de matière, ces centres échangent aussi du sens (significations des événements). Ces systèmes sont moins stables que les systèmes définis par la deuxième génération.

Pour comprendre la différence des modélisations dans une approche systémique, nous en présentons les principales caractéristiques par opposition à une approche analytique. La comparaison présentée a été réalisée par J. Rosnay (1975, p. 108). Nous la représentons dans le Tableau 2-1.

Tableau 2-1. Les approches analytique et systémique, selon (ibid.)

Approche analytique	Approche systémique
Isole: se concentre sur les éléments	Relie: se concentre sur les interactions entre les éléments.
Considère la nature des interactions.	Considère les effets des interactions
S'appuie sur la précision des détails.	S'appuie sur la perception globale.
Modifie une variable à la fois.	Modifie des groupes de variables simultanément.
Est indépendante de la durée : les phénomènes considérés sont réversibles.	Intègre la durée et l'irréversibilité.
Valide les faits par la preuve expérimentale dans le cadre d'une théorie.	Valide les faits par comparaison du fonctionnement du modèle avec la réalité.
Crée des modèles précis et détaillés, mais difficilement utilisables dans l'action (exemple: modèles économétriques).	Crée des modèles insuffisamment rigoureux pour servir de base de connaissances, mais utilisables cependant dans la décision et l'action (exemple: modèles du Club de Rome).

Propose une approche efficace lorsque les interactions sont linéaires et faibles.	Propose une approche efficace lorsque les interactions sont non linéaires et fortes.
Conduit à un enseignement par discipline (juxta-disciplinaire).	Conduit à un enseignement pluridisciplinaire.
Conduit à une action programmée dans son détail.	Conduit à une action définie par objectifs.
Permet une bonne connaissance des détails, mais les buts sont mal définis.	Permet une bonne connaissance des buts, mais les détails demeurent flous.

La modélisation du système d'apprentissage que nous proposons combine des éléments des deux approches : modèles précis et détaillés pour servir de base de connaissances, interactions non linéaires et fortes, tout en mettant l'accent sur plusieurs avantages offerts par l'approche systémique (e.g. interaction des éléments, effets d'interaction, perception globale, action par objectifs).

2.1.3. Typologie des systèmes. Système actif et expert

L'arborescence et le niveau d'organisation des systèmes sont les premiers concepts clés dans la construction d'une typologie des systèmes. Sur une échelle de complexité croissante, les premiers niveaux de l'arborescence correspondent aux systèmes physiques et chimiques. Dans (AFSCET, 2003, p. 3), Bunge propose une classification construite en fonction de la complexité des systèmes et de l'ordre supposé d'apparition dans le temps. La Figure 2-6 illustre cette classification.

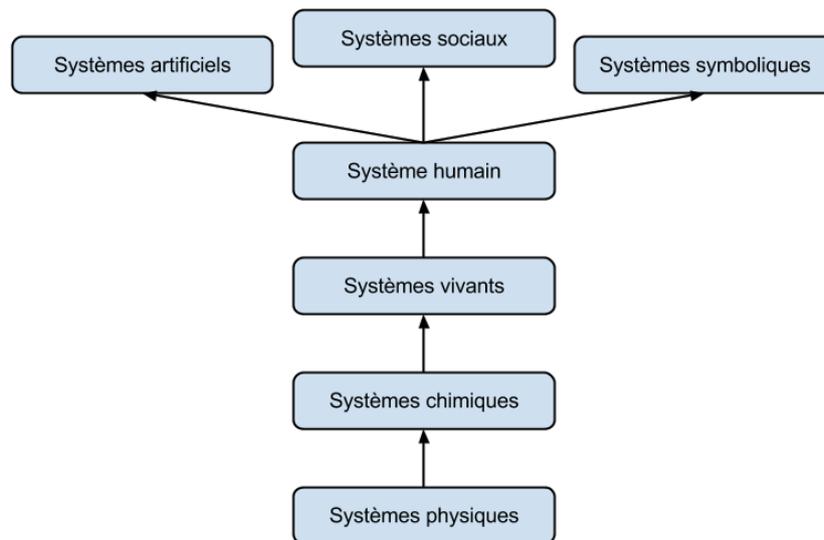


Figure 2-6. Typologie des systèmes, selon (ibid., p. 4)

H. Simon (Simon, 1974) et K. Boulding (Boulding, 1956) ont aussi contribué à l'effort de construction d'une typologie des systèmes. La complexité et le niveau d'organisation sont des critères évidents dans leur classification sur huit niveaux⁸⁵. Ces niveaux évoluent proportionnellement à la croissance de la complexité des environnements modélisés.

Dans la typologie proposée par J. Lesourne (Lesourne, 1976) on distingue :

- les systèmes à états : ils ont, comme leur nom l'indique, des états bien définis, l'état du système peut être déterminé, ils permettent les transformations entrées/sorties à travers les états du système ;
- les systèmes à buts : ils ont la capacité d'atteindre des objectifs ;
- les systèmes à apprentissage : caractérisés par mémoire, mécanismes de calcul, capacité de décision et d'adaptation, auto-organisation ;

⁸⁵ Les huit niveaux sont : (1) cadres dotés d'une structure statique (e.g. systèmes d'organisation, d'anatomie et géographie de l'univers, modèles d'électrons autour d'un noyau) ; (2) rouages, machines simples (e.g. système solaire, poulie, moteur) ; (3) thermostats, systèmes cybernétiques ou mécanismes de contrôle (e.g. les systèmes régulateurs) ; (4) cellules, systèmes ouverts et autorégulateurs : c'est le niveau où commencent les systèmes dits « vivants » ; (5) plantes, niveau société-génétique, groupes (sociétés) de cellules organisés dans des parties différentes mais interdépendantes (e.g. racines, feuilles) ; (6) animaux, systèmes capables de mobilité élevée et des récepteurs d'information spécialisés (e.g. yeux, oreilles) ; (7) êtres humains, qui, en plus des caractéristiques des systèmes animaux, possèdent une connaissance et une conscience plus complexes ainsi que la capacité de s'exprimer d'une façon intelligente ; et (8) organisations sociales, ensemble de systèmes vivants qui ont des rôles spécifiques et sont liés par des canaux de communication (AFSCET, 2003, p. 16).

- les systèmes à décideurs multiples sont des systèmes complexes composés par des sous-systèmes à buts, s'organisant de manière spontanée, hiérarchique.

Dans l'ensemble des typologies selon le niveau d'organisation des systèmes, chaque niveau supérieur englobe le niveau inférieur et projette sur celui-ci sa finalité. Le degré de complexité d'un système peut ainsi être mesuré en fonction du nombre et de la nature des différents niveaux d'organisation qui le composent.

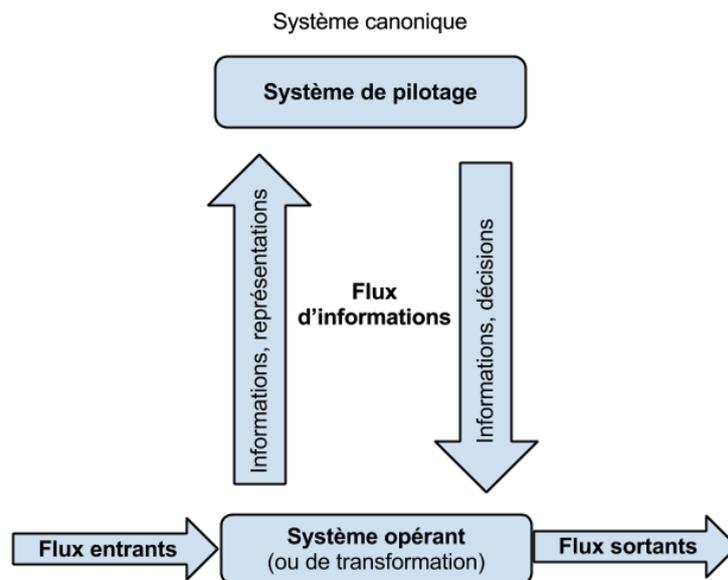


Figure 2-7. **Système canonique** (selon : Donnadiou et Karsky, 2002, p. 68)

Se basant sur le fait que les systèmes peuvent se décomposer en sous-systèmes, G. Donnadiou et M. Karsky (2002) présentent différentes organisations des systèmes en fonction de leur structure et composition. Les systèmes de degré un, dits *systèmes canoniques*, (v. Figure 2-7), correspondent à une organisation constituée d'un système de pilotage et d'un système opérant. Le but du système de pilotage est d'assurer la régulation du système, alors que le système opérant assure la transformation.

La complexité d'organisation des systèmes se poursuit avec les systèmes de degré deux, trois, quatre et cinq (v. ANNEXE A). Dans le niveau deux, le système de pilotage est constitué d'un système d'information et un système de décision. Le troisième degré de complexité d'un système correspond à l'accroissement de la densité d'interconnexions des divers processeurs décisionnels. Le système de degré quatre apporte l'auto-organisation dans l'ensemble du système. Il dispose pour cela de générateurs internes

(v. Figure 2-8), « *produisant de manière aléatoire des informations susceptibles de recevoir une interprétation dans le cadre du code sémantique qui gouverne le système* » (CERTU, 2007, p. 27).

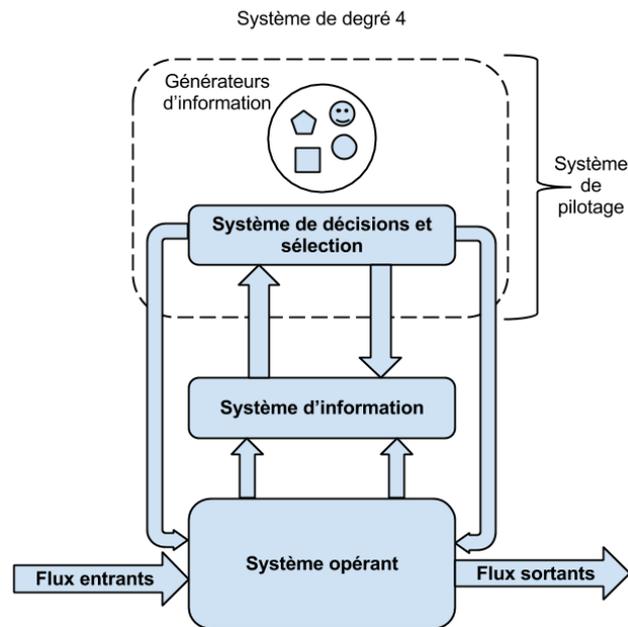


Figure 2-8. Système de degré quatre (selon : Donnadiou et Karsky, 2002, p. 73)

Afin de mieux atteindre sa finalité, le système de degré quatre organise les informations qu'il manipule conformément à une sémantique définie dans le cadre du système. Ces informations associées à un sens donné permettent au système de prendre des décisions.

Le système de degré cinq s'enrichit de la propriété d'auto-finalisation. Cette propriété est spécifique à l'Homme, aux constructions sociales et sociétales dont l'Homme est partie composante (e.g. familles, peuples, nations). Le système de degré cinq se donne à lui-même des objectifs et projets.

Notre prototype de système d'apprentissage s'inscrit dans la catégorie des systèmes de quatrième degré. Les principaux générateurs d'information sont l'apprenant et les créateurs de ressources d'apprentissage. L'interaction de l'apprenant avec le système d'apprentissage produit les informations nécessaires au système d'apprentissage afin de prendre une décision pour déclencher un ou plusieurs événements (e.g. recommandation d'une ressource pédagogique).

La capacité de déclencher des actions et des activités (e.g. réorganisation des éléments) dans le cadre du système qualifie le système d'*actif*. D'après Baroni *et al.* le *système actif* est un « *système dynamique qui interagit avec le monde extérieur et, en fonctionnement, soit il est quiescent, soit il réagit. Lorsque le système est quiescent, il ne change pas son état, et il ne génère aucun événement. Si, à un instant imprévisible, le système au repos reçoit un événement (venant nécessairement du monde extérieur), il commence à réagir, en d'autres termes, il commence à évoluer* » (Baroni *et al.*, 1999, p. 137-138). Au cours d'une réaction, le système actif peut générer des événements dirigés vers le monde extérieur qui, à son tour, peut générer des événements perçus par le système actif, ce qui affecte son évolution ultérieure. Lorsque la réaction est terminée, le système actif devient « quiescent » à nouveau.

Si cette définition se situe dans le courant de la théorie behavioriste (e.g. répondre aux stimuli), d'autres définitions éliminent la contrainte énonçant que les événements doivent venir de l'extérieur du système. Prenons par exemple l'organisme humain : si on le voit comme un système, ce système est intrinsèquement actif. L'activité autonome des différents systèmes (e.g. système nerveux) réside dans le système lui-même et rend le système actif. Dans l'évolution du système, « *les mécanismes de réaction semblent être super-imposés par les activités locomotrices rythmiques primitives, innées (e.g. la respiration, la marche). Les stimuli ne provoquent pas un processus dans un système par ailleurs inerte ou "fermé", il modifie uniquement de façon indépendante les processus dans un système actif* » (Mitchell, s. d.).

Pour la conception de notre prototype de système actif d'apprentissage, le terme de *système actif* est proposé dans le rapport avec *l'apprenant*. Quant à l'interaction du système relatif aux facteurs d'excitation, notamment de l'apprenant, deux scénarios existent :

- système réactif : le système répond aux différentes excitations et actions de l'apprenant, et modifie ensuite son état, organisation, etc. ;

- système proactif⁸⁶ : le système anticipe les actions prochaines et, avant de recevoir des excitations extérieures de la part d'apprenant, il modifie son état, organisation, etc.

Les comportements du système impliquent la coordination de plusieurs éléments le composant : facteurs d'excitation, générateurs et sources d'informations, règles de décisions, etc.

Si dans un domaine particulier, dans des circonstances spécifiques, un système est capable d'émuler des comportements (e.g. décisions, mécanismes cognitifs) comparables à ceux d'un expert humain, le système est appelé *système expert*. Conformément à Jackson, un système expert est « *un programme informatique qui représente et raisonne avec la connaissance d'un certain sujet spécialisé en vue de la résolution de problèmes ou pour donner des conseils* » (Jackson, 1999, p. 2). Le système expert est construit par « *l'assemblage d'une base de connaissances qui est interprétée par un logiciel qui contient un moteur d'inférence* » (ibid., p. 30). La base de connaissances est constituée par une base de faits et une base de règles.

Le prototype de système d'apprentissage proposé dans cette thèse s'inscrit dans la catégorie des systèmes actifs et des systèmes experts. Le comportement du système s'adapte aux nécessités de l'apprenant. Il a la capacité de proposer des ressources d'apprentissages adaptées aux besoins de l'apprenant. Le système comporte une base de connaissances et un ensemble de règles sur lesquelles s'appuie la composante décisionnelle du système dans le choix de la ressource d'apprentissage à proposer à l'apprenant. Le but du système se décline en deux principaux composants. Dans un premier temps, le but est *d'identifier et de représenter les entités participant à l'apprentissage* (l'apprenant, les ressources d'apprentissage, les objectifs d'apprentissage). Dans un second temps, le but est *d'émuler des décisions intelligentes* afin d'offrir une expérience d'apprentissage efficace, active et adaptée aux besoins de l'apprenant.

⁸⁶ À l'heure actuelle le Web se transforme en un Web proactif. Le comportement des applications Web (e.g. suggestions des recherches, proposition des produits) tend à anticiper et à réduire le travail humain dans ses tâches.

2.2. Systèmes d'information, de communication et d'apprentissage

Dans le cadre des sciences de l'information et de la communication, l'information et la communication sont intimement liées⁸⁷. L'objet de la communication est l'information et l'information fait sens et valeur lorsqu'elle est communiquée. L'évolution des systèmes d'apprentissage est étroitement liée au développement des systèmes de l'information et de la communication.

Le développement des technologies, et plus particulièrement des technologies de l'information et de la communication, a influencé la société, la façon de vivre et la manière d'apprendre. Avant d'être un système d'apprentissage, le SASA est un système d'information et de communication. Il identifie puis traite l'information qui s'échange dans le système, et gère la communication entre les éléments du système. Dans cette section, nous allons positionner le rôle des systèmes d'information et de la communication dans la construction des systèmes d'apprentissage.

2.2.1. Information et communication

B. Denis se demande si les mots « *information* » et « *communication* » sont des mots-clés ou sont des mots-omnibus, leur sens variant « *parfois du tout au tout selon l'auteur qui les emploie et le contexte d'utilisation* » (Denis, 1994, p. 26). *Le Petit Robert* (2010, p. 1328) donne trois sens principaux pour le mot *information* : « *I. ensemble des actes qui tendent à établir la preuve d'une infraction et à en découvrir les auteurs [...]* », « *II. Renseignement sur quelqu'un, sur quelque chose [...]* » et « *III. Élément ou système pouvant être transmis par un signal ou une combinaison de signaux (message), appartenant à un répertoire fini ; Ce qui est transmis (objet de connaissance, de mémoire) [...]* »

Les mots *informer* et *information* viennent du latin *informare*, qui signifie *former, donner une forme*, mais aussi *façonner l'esprit*. B. Denis continue et affirme que « *[...] aujourd'hui, dans le langage courant, la notion d'information renvoie plus spécifiquement à un ensemble de processus technologiques, de données techniques [...]* » (Denis, 1994, p.

⁸⁷ « Malheureusement, il ne suffit pas d'informer pour communiquer » (Wolton, p. 396). Avec la performance technique, on constate le passage « de la vitesse de l'information à la lenteur de la communication » (ibid.)

23-24). Le mot « *information* » est aujourd'hui utilisé avec une grande variété de sens : les *données* de l'informaticien, les *nouvelles* du journaliste, le *savoir* et la *connaissance*. « *Information* » sert à distinguer aussi *un symbole* ou *un signe* (e.g. 0 et 1 dans le codage binaire), *une quantité* transportée précisément mesurable (la théorie mathématique de la communication mesure par une formule invariable à base de logarithmes la « quantité d'information ») et *une qualité* (la signification des messages transmis).

F. Heinderyckx (2002, p. 7) distingue deux *familles sémantiques* dans la définition de l'information :

1. « *au sens restrictif: l'information est un élément qui peut être transmis par un signal. La cybernétique parle de "mesure de la réduction de l'incertitude qu'il peut y avoir au sujet de l'état d'une partie de l'univers, par l'intermédiaire d'un message"* »
2. « *sens journalistique: l'information est "nouvelle", "actualité", "renseignement ou événement qu'on porte à la connaissance d'une personne, d'un public"* ».

Comme une information « c'est un message qui apporte une précision dans une situation comportant un certain degré d'incertitude » (Roxin et Mercier, 2004, p. 33), on peut se demander quelle est la relation entre information et certitude ? D'une part, la quantité d'information et le niveau d'incertitude sont directement proportionnels. Ainsi, afin de préciser l'état courant d'un système comportant un très grand nombre d'états, nous aurons besoin d'une grande quantité d'information. D'autre part, l'information varie en sens contraire de l'incertitude : en augmentant l'information relative à un système, on diminue d'autant l'incertitude⁸⁸. En conséquence, pour déterminer la quantité d'information on peut utiliser la même mesure que pour le degré d'incertitude.

En 1948, C. Shannon⁸⁹ donne l'expression mathématique qui permet de calculer la quantité moyenne d'information en fonction du niveau d'incertitude. Soit un système ayant plusieurs états possibles, chaque état ayant une probabilité p_i , avec la somme des probabilités des états possibles égale à 1 (un système complet d'événements, Équation 2-3).

⁸⁸ « *La somme d'information dans un système est la mesure de son degré d'organisation [...]; l'entropie est la mesure de son degré de désorganisation; l'un étant le négatif de l'autre* » (Mattelart et Mattelart, 1995, p. 36).

⁸⁹ Claude Elwood Shannon (1916 – 2001) publie en 1948 le papier « A mathematical theory of communication » qui va poser les fondements de la théorie de la communication.

$$\sum_i^n p_i = 1, \text{ avec } 0 \leq p_i \leq 1$$

Équation 2-3

L'incertitude dans un tel système, représentée par une variable aléatoire (X) dépend essentiellement de ses probabilités (Équation 2-4)

$$X = \begin{bmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{bmatrix}$$

Équation 2-4

La quantité d'information ou *l'entropie*⁹⁰ *informationnelle* se calcule d'après la formule donnée par Shannon (Équation 2-5).

$$H(p_1, p_2, \dots, p_n) = - \sum_{i=1}^n p_i \times \log_2(p_i)$$

Équation 2-5

L'entropie, ou degré selon lequel une variable peut varier, est exprimée en fonction des probabilités des symboles qui peuvent apparaître dans une variable aléatoire donnée.

Si les probabilités des symboles sont égales ($p_i=1/n$), l'Équation 2-5 devient :

$$H(p_1, p_2, \dots, p_n) = - \sum_{i=1}^n \frac{1}{n} \times \log_2\left(\frac{1}{n}\right) = \log_2(n)$$

Équation 2-6

La formule résultante est la formule de R.V.L. Hartley⁹¹ et décrit le cas particulier dans lequel tous les états possibles sont équiprobables.

Le cas le plus simple est représenté par une variable qui ne peut avoir que deux états équiprobables. C'est le cas qui permet de définir l'unité de mesure d'information, désignée par le mot *bit*⁹². Un *bit* désigne l'incertitude qu'une variable binaire et aléatoire soit dans un des deux états possibles (e.g. *0* ou *1*, *vrai* ou *faux*, *oui* ou *non*) avec une probabilité égale d'être dans l'un ou l'autre de ces deux états (J. B. Anderson et Johnnesson, 2006, p. 153).

Ainsi pour une variable qui peut avoir deux états équiprobables ($p_1 = p_2 = 1/2$), l'entropie est :

⁹⁰ En analogie avec l'entropie physique de Boltzman.

⁹¹ Ralph Vinton Lyon Hartley (1888 - 1970) propose cette formule en 1928.

⁹² Le mot « bit » est la contraction des mots anglais **binary digit**, qui signifient « chiffre binaire », avec un jeu de mot sur bit, « morceau ». 8 bits forment un octet (byte en anglais).

$$H(x) = -\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{2} \log_2 \frac{1}{2} = -\log_2 \frac{1}{2} = \log_2 2 = 1 \text{ bit} \quad \text{Équation 2-7}$$

Dans le cas où tous les états sont équiprobables, l'incertitude est maximale et la quantité d'information pouvant y être représentée est dès lors maximale.

Si l'information concerne des *études scientifiques* (à l'aide des outils mathématiques et informatiques), la communication relève plutôt, quant à elle, d'un ensemble de « *pratiques sociales ou interindividuelles* » et se traite dans le champs des *études culturelles* (Denis, 1994, p. 25). A. Moles⁹³ et C. Zeltmann (cité par Heinderyckx, 2002, p. 8) parlent de « *l'action de faire participer un individu – ou un organisme –, situé à une époque, en un point donné R, aux expériences et stimuli de l'environnement d'un autre individu – ou d'autre système –, situé à une autre époque, en un autre lieu E, en utilisant les éléments de connaissance qu'ils ont en commun* ».

Les mots « *communication* » et « *communiquer* » viennent du mot latin *communicare* qui signifie *mettre en commun, être en contact avec, participer à*. La communication concerne des activités relatives au *transport* et *communiquer* renvoie à l'idée de transmettre, envoyer, diffuser, « *faire passer une personne, un objet ou une information d'un endroit à un autre* » (Denis, 1994, p. 23).

R. Escarpit (1991) suggère d'employer « le mot *transport* pour désigner le transfert de matière ou d'énergie et de réserver celui de *communication* au transfert de l'entité nommée *information* ». Pour lui, la communication est « *un acte, un processus de transmission issu du fonctionnement d'un appareil, d'une machine biologique ou non, (cerveau humain ou animal, ordinateur, etc.). Communiquer c'est procéder à un échange d'informations. L'information étant, quant à elle, l'entité mesurable, objet de l'échange, du processus de communication* ». C.E. Shannon (2001) a élaboré une représentation schématique de la communication, illustrée par la Figure 2-9. Le schéma, décomposé en étapes distinctes, présente une structure linéaire unidirectionnelle de la communication.

⁹³ Abraham Moles (1990 -1992) fut « le créateur de multiples liens entre les sciences de l'information et de la communication et les sciences psychologiques » (Devèze, 2004, p. 189)

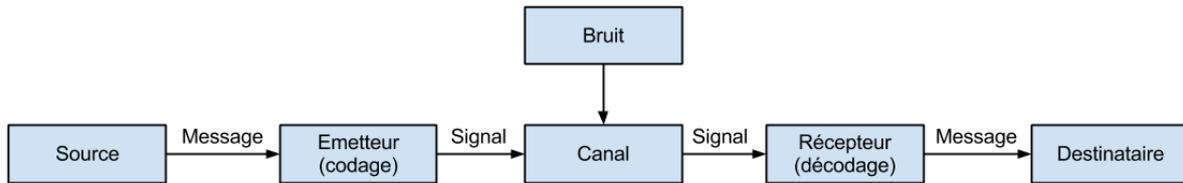


Figure 2-9. Schéma d'un système général de communication

L'émetteur, le canal et le récepteur constituent les *intermédiaires techniques* et les principales préoccupations de C.E. Shannon. Le bruit, « *c'est à dire toute source d'interférence susceptible de détériorer le signal et donc d'affecter la communication, occupe également une position centrale dans ce modèle. La communication ainsi modélisée place en tête des préoccupations la lutte contre le bruit [...]* » (Heinderyckx, 2002, p. 14).

Selon F. Heinderyckx, W. Weaver⁹⁴ a *humanisé* le schéma *purement technique* de C.E. Shannon : « *“Les aspects sémantiques de la communication ne relèvent pas de ses aspects technologiques”* écrivait Shannon, mais *“cela ne veut pas dire que les aspects technologiques n'interviennent pas dans les aspects sémantiques”*, ajoute Weaver » (*ibid.*, p. 15). W. Weaver introduit l'aspect sémantique dans le schéma de Shannon (v. Figure 2-10). Le *récepteur sémantique* est introduit entre le récepteur technique et le destinataire. Son but est « *d'accorder les caractères sémantiques des messages avec les possibilités sémantiques des destinataires* ». L'accord sémantique peut être perturbé par le *bruit sémantique*. Inséré entre la source et l'émetteur, il perturbe le message au niveau de sa signification.

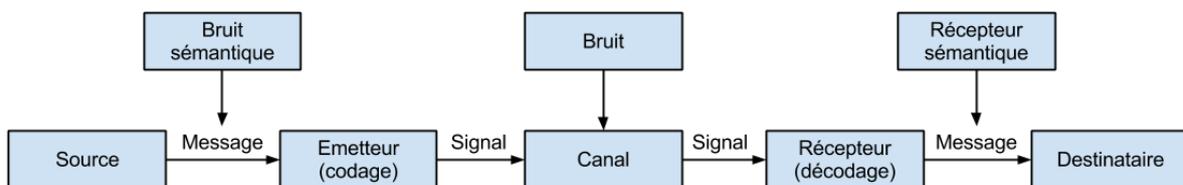


Figure 2-10. Schéma d'un système général de communication avec les apports de Weaver

W. Weaver présente les « *trois niveaux de problèmes de communication* » : (1) niveau technique (précision de transmission des symboles, système de codage de l'information); (2) niveau sémantique (la signification des symboles); (3) niveau efficacité (résultats espérés, influence sur les comportements et attitudes).

⁹⁴ Warren Weaver (1896 – 1978), linguiste américain, mathématicien, cosigne en 1949 avec Claude Shannon l'ouvrage « *Théorie mathématique de la communication* ».

Dans une communication technique⁹⁵ basée sur le Web, nous pensons que l'utilisation des technologies dites du « Web sémantique » peut apporter des améliorations au niveau de l'interprétation du message transmis lors d'une communication. En partageant une *spécification de la conceptualisation* d'un domaine tant au niveau de la source qu'à celui de la destination, l'émetteur et le récepteur améliorent leur capacité de compréhension du message. Les limites se retrouvent cependant dans le niveau d'expressivité des langages du Web sémantique et les capacités des agents à interpréter le message dans le contexte de la spécification commune.

2.2.2. Systèmes et technologies d'information et de communication : l'Internet et le Web

À la fin du 18^{ème} siècle, la France avait réalisé un progrès technique important : le premier système de télégraphie au monde. « *Avant le Web, avant l'ordinateur, avant le téléphone, même avant le code Morse, il y avait le système Chappe* » (« How Napoleon's semaphore telegraph changed the world », 2013). C. Chappe⁹⁶ avait inventé le *télégraphe*⁹⁷ *optique*⁹⁸. Si, en 1790, la transmission d'un message de Paris à Strasbourg prenait « *quatre jours [...], en 1799 le même message mettra moins... de 2 heures* » (« Communiquer à distance : le telegraphe Chappe », s. d.). Cette avancée technique réduit le temps de transfert de l'information d'environ 50 fois. La vitesse de transfert de l'information est, encore une fois, propulsée par le passage à la transmission électrique

⁹⁵ Dominique Wolton affirme que le considérable progrès technique « a créé l'illusion d'un continuum entre communication humaine et communication technique » (Wolton, 2012, p. 397). Mais « l'homme est définitivement plus compliqué que la technique ». L'homme communique « pour séduire, partager ou convaincre ». « La problématique de la communication est plus complexe que celle de l'information. L'information c'est le message, la communication, la relation, donc la prise en compte du point de vue de l'autre. Une information se transmet, une communication se construit et se négocie [...] On passe ainsi d'une théorie de communication liée à un modèle de transmission à un modèle de *cohabitation* » (ibid.). « *Le vrai mystère de la communication, c'est la réception* » (ibid., p. 130).

⁹⁶ Claude Chappe (1763 - 1805) est le premier entrepreneur des télécommunications dans l'histoire de l'humanité.

⁹⁷ Du grec *télos* : loin et *graphein* : écrire.

⁹⁸ Le système consiste en stations (ou postes) distribués sur le territoire et distants de 6 à 12 km. Le message était passé de station en station par observation visuelle. Chaque station se composait d'un télégraphe optique ou télégraphe sémaphore et était équipé de lunettes. À l'aide de trois bras en bois, l'opérateur avait la possibilité de produire 98 positions (signaux) de base, dont 6 utilisées pour le service et 92 pour le message. Les messages se transmettaient en groupe de deux signes et étaient ensuite confrontés à un répertoire de 92 pages. Chaque page ayant 92 lignes, le nombre total de significations étaient de $92 \times 92 = 8464$.

(e.g. télégraphe électrique, internet) puis une nouvelle fois grâce à un mode de transmission emprunté à l'optique (e.g. la fibre optique). Cette fois, la vitesse de transport de l'information s'approche de la vitesse de la lumière.⁹⁹ L'invention du transistor¹⁰⁰ puis le passage au numérique des dispositifs de traitement et de stockage de l'information, permettent la "superposition" des médias et l'agrandissement des capacités de stockage de l'information (« Byte-sized graphic guide to data storage », 2013). Toutes ces technologies ont contribué au développement des technologies dites des systèmes d'information et de communication. Les systèmes d'information et de communication présentent comme objet central d'intérêt *l'information*. Le traitement, le stockage et la diffusion de l'information sont des fonctions de base dans un système d'information et de communication. La prise en charge d'une partie de ces fonctions par la technologie a fait avancer l'ensemble du système d'information et de communication.

Même si, au départ, le système d'information était associé au système informatique, la différence est aujourd'hui devenue évidente (Boussagol, 1996). Un système d'information est « *bien plus qu'un parc informatique mis en réseau [...] il est intrinsèquement lié à l'organisation et aux flux économiques et sociaux qui font vivre une organisation* » (Leitzelman et Dou, 1998, p. 56). Un système d'information et de communication est un « *ensemble d'éléments humains, technologiques, organisationnels et financiers ayant comme fonction de traiter, stocker et diffuser les informations dans l'entreprise* » (Dibiaggio et Meschi, 2010, p. 181).

Actuellement, en raison de leur taux de propagation et d'utilisation, deux systèmes s'imposent : l'Internet¹⁰¹ et le Web. Notre attention se concentre dans un premier temps sur l'Internet et ensuite sur le Web comme système qui se construit sur les possibilités offertes par l'Internet. J. Rosnay affirme que « *l'Internet est un véritable système*

⁹⁹ En fonction du support à travers lequel passe le courant électrique, la vitesse du champ électromagnétique varie de 66% - pour les câbles coaxiaux typiques - et 97% pour les conducteurs en cuivre de la vitesse de la lumière. Étant données les limites des câbles, le débit de données peut aller jusqu'à 10 Gigabits/seconde. Les derniers résultats montrent qu'avec un nouveau type de fibre optique, la lumière peut atteindre des vitesses équivalentes à 99.7% de la vitesse de la lumière dans le vide (Poletti et al., 2013) : on réussit alors à transférer 40 Gigabit/seconde sur 37 canaux au sein de la même fibre.

¹⁰⁰ Le transistor a été inventé le 23 décembre 1947 par les américains William B. Shockley, J. Bardeen et W.H. Brattain. Ils ont reçu le prix Nobel de Physique en 1956 pour cette invention. Voir http://www.nobelprize.org/nobel_prizes/physics/laureates/1956/.

¹⁰¹ Internet – Dérivé du concept *d'internetting* (« interconnecter des réseaux »).

informationnel » (« Les quatre Web de Joel de Rosnay, du 1.0 au 4.0 », s. d.). L'Internet implique l'usage de la technologie, mais aussi des humains et des organisations pour la gestion de ce système. Parmi les canaux de transmission de l'information tels que la radio, la télévision ou l'Internet, c'est l'Internet qui a connu la propagation (l'adaptation, l'intégration) la plus rapide. Ce réseau de communication ayant l'avantage de permettre une communication bidirectionnelle, il est actuellement le support de base sur lequel s'appuient de nombreux systèmes de communication (e.g. le Web, l'e-mail). En tant que système de communication, l'Internet est structuré selon le modèle *Open Systems Interconnections (OSI)*. Produit dans le cadre d'un effort de standardisation des réseaux informatiques démarré en 1977 par l'*Organisation internationale de normalisation (ISO)*, le modèle OSI propose de standardiser et de grouper en couches les fonctions internes d'un système de communication. Ainsi, les applications Web et le protocole de transfert de l'hypertexte se trouvent à la septième couche de ce modèle.

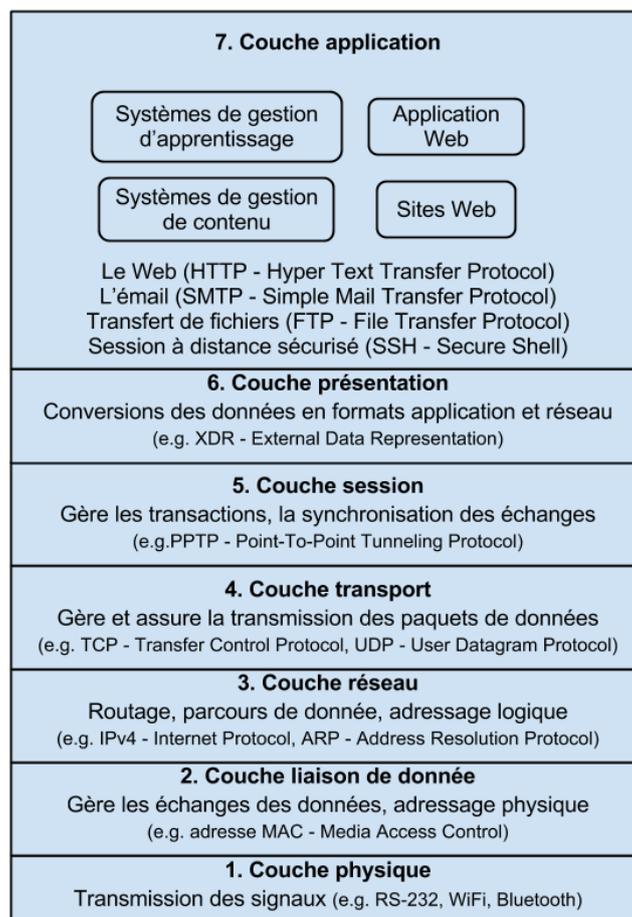


Figure 2-11. Le modèle OSI – l'Internet et le Web

L'Internet, le réseau des réseaux, connaît le début de son développement en 1964 au sein de l'agence ARPA¹⁰², qui va financer le projet ARPANET (ARPA Network). En 1968, J. Licklider et R. Taylor (1968, p. 21) affirment que « *en quelques années, les hommes seront en mesure de communiquer plus efficacement à travers une machine que face à face* ». Ils reconnaissent l'entrée dans une « *ère technologique dans laquelle nous serons en mesure d'interagir avec la richesse de l'information vivante, non seulement dans la manière passive que nous sommes habitués à utiliser pour les livres et les bibliothèques, mais en tant que participants actifs dans un processus continu, en apportant quelque chose à travers notre interaction avec elle, et pas simplement recevoir quelque chose d'elle par notre connexion* » (*ibid.*).

Développées par V. Cerf et R. Kahn, les spécifications de l'ensemble des protocoles de communication TCP/IP¹⁰³ sont publiées en décembre 1974 (Cerf, Dalal, et Sunshine, 1974). L'accès à l'Internet permet à chaque utilisateur d'accéder au contenu voulu (Web ou autre) d'une manière libre et immédiate. Cette pénétration de l'Internet, une nouvelle révolution culturelle, impose l'adaptation de l'apprentissage et une conception des systèmes d'apprentissage adaptée à une utilisation sur Internet.

Le système de communication le plus répandu et le plus utilisé, qui s'appuie sur l'Internet, est *le Web*¹⁰⁴. En 1990, T. Berners-Lee, propose le projet *WorldWideWeb*, un système hypertexte¹⁰⁵ qui lui permettra de publier des informations structurés sous la

¹⁰² *Advanced Research Project Agency*: plus tard transformé en DARPA (Defense Advanced Research Project Agency).

¹⁰³ En 1994, l'IETF (Internet Engineering Task Force) réalise que le protocole de communication sur internet IPv4, qui fournit presque 4,3 milliards (2^{32}) d'adresses internet (IP), ne pourra pas satisfaire les nouvelles demandes¹⁰³. Les nouveaux dispositifs demandant un nombre de connexions Internet de plus en plus élevé, l'IETF propose le protocole IPv6, qui permet d'allouer jusqu'à $3,4 \times 10^{38}$ d'adresses IP (2^{128}). Même si le transfert vers l'IPv6 a été réalisé d'une façon progressive, le 6 juin 2012 a marqué officiellement le passage au nouveau protocole. (Kaushik, s. d.; Vaughan-Nichols, 2012).

¹⁰⁴ Le World Wide Web (WWW), littéralement la « toile (d'araignée) mondiale », communément appelé le Web et parfois la toile, est un système hypertexte public fonctionnant sur Internet qui permet de consulter, grâce à un navigateur, des pages accessibles sur des sites. L'image de la toile d'araignée vient des hyperliens qui lient les pages Web entre elles. On retrouve aussi l'association de *Web spider* pour le robot d'indexation.

http://fr.wikipedia.org/wiki/World_Wide_Web

¹⁰⁵ Terme inventé dans les années 1965 par Ted Nelson, « hypertext » désigne un texte qui n'est plus nécessairement linéaire (texte + hyperliens). On trouvera plus d'information sur l'hypertexte dans le paragraphe 3.1.3.

forme d'une toile (*Web* en anglais) avec des nœuds connectés que l'utilisateur peut exploiter en suivant son propre parcours (Berners-Lee, 1990). Le premier site Web a été lancé le 6 août 1991 à l'adresse <http://info.cern.ch> (« About this project | Restoring the first Website », s. d.) et le 30 avril 1993, le Conseil Européen pour la Recherche Nucléaire (CERN) publie le document qui rend la technologie du *WorldWideWeb* (W3) disponible avec une licence libre de droits (« The document that officially put the World Wide Web into the public domain on 30 April 1993. », 1993).

Le Web est un système de publication et de diffusion d'informations, dont l'information effective est contenue dans des *documents Web*. Ces documents, écrits dans le langage *Hypertext Markup Language* (HTML)¹⁰⁶, sont déposés sur un serveur Web et accessibles pour les clients à travers un navigateur capable d'interpréter le langage HTML. Comme langage de balisage, HTML permet la mise en forme des documents (composition et structure du contenu, paragraphe, titre, image, hyperliens, éléments de saisie) et la construction des liens entre plusieurs documents. Le protocole *Hyper Text Transfer Protocol* (HTTP) est utilisé pour la communication et le transfert de documents Web entre le client et le serveur.

Sur le Web, le porteur principal d'information est le *document Web*¹⁰⁷. Selon W3.org « *un document Web est défini comme quelque chose qui a un URI¹⁰⁸ et peut retourner des représentations (les réponses dans un format tel que HTML ou JPEG ou RDF) de la ressource identifiée en réponse à des requêtes HTTP. Dans la littérature technique, telle que l'architecture du World Wide Web, Volume One [AWWW], le terme ressource d'information est utilisé à la place de document Web* » (Sauermann et Cyganiak, 2008).

Avec le développement des solutions d'administration, de création et d'annotation des documents Web, une nouvelle catégorie de systèmes a ainsi vu le jour : *les systèmes de*

¹⁰⁶ HTML5 est la dernière version du langage HTML : <http://www.w3.org/TR/html5/>

¹⁰⁷ Pour Ted Nelson, le document électronique (donc aussi le document Web) nous permettra d'échapper à la « *prison de la papier* ».

¹⁰⁸ *Uniform Resource Identifier, en anglais*. Chaque document Web a son propre URI. Un document Web n'est pas la même chose qu'un fichier: un seul fichier (e.g. script PHP) peut être responsable de la génération d'un grand nombre de documents Web avec ou non des URI différents. On trouvera plus de détails sur les URI dans le paragraphe 3.3.1.

*gestion de contenu*¹⁰⁹ (SGC). Ces systèmes permettent le partage du travail (e.g. éditeur, créateur, designer) et améliorent le processus de publication des documents Web. Ils reposent et sont construits sur la structure offerte par les technologies du Web. Les systèmes d'apprentissage basés sur le Web sont un cas particulier des systèmes de gestion de contenus créés pour le Web.

Dans le cas général, un SGC pour le Web est un logiciel informatique qui permet la gestion (e.g. création, édition, publication, stockage) des documents Web. Parmi les caractéristiques principales de ces systèmes se trouvent des fonctions qui permettent la gestion du contenu indépendamment de la gestion d'affichage (e.g. mise en forme, disposition, design) (Powel et Gill, 2003; Baxter et Vogt, 2002).

Les SGC pour le Web sont un des facteurs principaux de l'explosion des contenus générés par les utilisateurs sur le Web. Les fonctionnalités riches du SGC corrélées avec une grande simplicité d'utilisation fournissent un outil puissant à tous les utilisateurs désirant administrer (e.g. créer, éditer) leur propre contenu sur le Web. L'accès à un SGC se fait à travers un navigateur Web et son utilisation n'exige aucune connaissance spécifique des technologies du Web (e.g. langage HTML) (Powel et Gill, 2003, p. 45-46).

Les SGC¹¹⁰ permettent la réutilisation du contenu (e.g. texte, graphiques, images, vidéo, audio) stocké dans des espaces partagés sur l'ensemble du système. La base de données, au centre des SGC, est le moyen principal de stockage du contenu. Le système enregistre, dans une base de données, toute information nécessaire pour la construction des pages Web (e.g. des modèles d'affichage, des données concernant les utilisateurs, les métadonnées). Dans ces systèmes, la page Web statique est inexistante, toute page est dynamiquement générée en fonction de différents paramètres (e.g. utilisateur, temps).

¹⁰⁹ En anglais : *Content Management System* (CMS)

¹¹⁰ Parmi les SGC les plus répandus, nous pouvons citer *WordPress, Joomla, Drupal* ou *ModX* (Cawley, 2012).

2.2.3. Métadonnée : description des ressources d'apprentissage

Étant donnés les besoins de rendre le contenu exploitable par des systèmes externes (e.g. moteurs d'indexation), les SGC offrent des fonctions qui permettent d'enregistrer et d'intégrer des informations supplémentaires au contenu. En intégrant un vocabulaire spécifié à l'avance, les systèmes peuvent enregistrer et intégrer des *métadonnées* (e.g. créateur, date de création) au contenu des documents Web.

Facilitant l'administration des ressources distribuées, les métadonnées constituent un outil de plus en plus répandu et utilisé sur le Web. Les métadonnées sont des éléments-clés pour assurer la survie et la continuité de l'accessibilité à ces *ressources d'information* (National Information Standards Organization (U.S.), 2004). Le terme *ressource* est utilisé dans un sens général pour tout ce qui peut-être identifié sur le Web par une URI (Berners-Lee, Fielding, et Masinter, 2005).

Servant à décrire une autre donnée, la *métadonnée* peut être définie comme une *donnée à propos de donnée* (« DCMI Metadata Basics », s. d.). Le préfixe grec *méta* indique l'autoréférence, ou même : « méta : indique un niveau supérieur de référence. Un méta-X est un X qui parle d'autre X, et peut-être de lui-même. » (Tisseau, 1996, p. 252). La métadonnée est « *l'information structurée* qui décrit, explique, localise, et d'une certaine manière rend plus facile à extraire, à utiliser ou à gérer une *ressource d'information* » (National Information Standards Organization (U.S.), 2004).

Les métadonnées articulent le contexte des objets d'intérêt. Aux *ressources* (e.g. fichiers sons, livres, images) s'ajoutent *les descriptions de ressources*. Ces *descriptions* sont constituées par tout type d'information (e.g. titre, classification, information contextuelle) qui décrit, caractérise ou référence la *ressource* en question. Elles ont pour but de faciliter l'utilisation, l'extraction ou la gestion de la *ressource*.

Avant l'arrivée de l'informatique, les bibliothécaires utilisaient déjà des *descriptions* sous forme de fiches bibliographiques pour enregistrer les descriptions des documents qu'ils manipulaient. La *Fédération internationale des associations de bibliothécaires et des*

bibliothèques (FIAB)¹¹¹ a proposé en 1954, une normalisation du catalogage: le *standard international pour la description des livres* (*International Standard Book Description – ISBD*). L'ISBD spécifie « tous les éléments » (groupés en huit catégories¹¹²) qui sont « nécessaires pour *décrire et identifier* tous les types de matériaux qui sont susceptibles d'apparaître dans des collections de la bibliothèque, attribuent un ordre à ces éléments de description, et spécifient un système de ponctuation pour la description » (« ISBD(G): General International Standard Bibliographic Description », 1992). À partir de la spécification ISBD *Général* (*ISBD(G)*), plusieurs déclinaisons de descriptions ont été conçues. Pour répondre à l'évolution technologique, en 1986 a été formé le groupe de travail appelé *ISBD Computer Files* (*ISBD(CF)*) qui publie sa première version finale en 1988 (Saur, 1997a). Connue sous le nom d'*ISBD Electronic Resources*, ses éléments sont toujours groupés en huit catégories (Saur, 1997b, p. 2).

L'usage des *métadonnées* dans la pratique montre que leur utilisation n'est pas simplement la description des *ressources d'information*, mais s'applique à la description de tout type de *ressources* ou *d'entités*, numériques ou non (Nilsson, 2010, p. 10). Elles sont utilisées pour décrire le *réfèrent de la donnée* (e.g. personne, lieu, évènement, livre). Pour englober toutes les entités, M. Nilsson définit la métadonnée comme « *donnée descriptive* à propos d'entités identifiables » (ibid., p. 11). Il souligne le fait que le terme « *donnée* » implique d'être dans une forme exploitable par les ordinateurs. T. Berners-Lee souligne aussi que l'expression *exploitable par l'ordinateur* (en anglais « *machine understandable* ») est un élément clé dans sa définition : « la métadonnée est une information *exploitable par l'ordinateur* à propos de ressources Web ou d'autres entités » (Berners-Lee, 1997a). En plus de cette définition il donne trois axiomes pour les métadonnées : (1) « la métadonnée est une donnée »; (2) « les métadonnées à propos d'un document peuvent exister dans le document-même, ou dans un autre document et elles peuvent être transférées dans un document accompagné »; (3) « les métadonnées

¹¹¹La FIAB est connue plus sous son sigle anglais *IFLA* (*International Federation of Library Associations and Institutions*). Constituée le 30 septembre 1927, l'*IFLA* est « la principale organisation représentant les intérêts des bibliothèques et des services d'information et de leurs usagers. Au niveau mondial elle est le porte-parole de la profession des bibliothèques et de l'information » (<http://www.ifla.org/>).

¹¹² Les huit catégories sont : (1) titre et mention sur la responsabilité ; (2) édition ; (3) matériel ou type de la publication ; (4) publication, distribution ; (5) description physique ; (6) série ; (7) notes ; (8) numéro international normalisé (v. <http://archive.ifla.org/VII/s13/pubs/isbdg0.htm>)

peuvent décrire des métadonnées » (ibid.). S'exprimant sur leur architecture, T. Berners-Lee affirme que les métadonnées sont représentées comme un *ensemble d'assertions indépendantes* (ibid.). Ces *ensembles* se forment à partir d'ensembles de *spécifications des descripteurs* appelés *vocabulaires*, ou même *vocabulaires contrôlés* (Friesen, 2004). Le partage de ces *vocabulaires* permet l'interprétation pertinente de l'information dans des systèmes indépendants.

Dans l'architecture distribuée du Web, les métadonnées constituent un facteur essentiel pour assurer la découverte, l'interopérabilité et l'échange pertinent d'informations. À travers l'ajout de descriptions aux informations et la spécification explicite du sens des descripteurs, les métadonnées sont un des éléments principaux dans le déploiement du Web sémantique. À l'heure actuelle, parmi les *vocabulaires génériques* les plus répandus et utilisés sur le Web, nous pouvons citer *Dublin Core*¹¹³ et *Schema.org*¹¹⁴.

Le « *Dublin Core simple* »¹¹⁵ comprend 15 éléments¹¹⁶ de description permettant l'annotation de la source du document (e.g. titre, créateur, date, sujet), de son contenu (e.g. mots clés, résumé, langue, format, identifiant) ou de ses droits de propriété intellectuelle (e.g. droits d'auteur) (v. ANNEXE B).

Soutenu par les principaux fournisseurs de services de recherche sur le Web (e.g. Bing, Google, Yahoo !, Yandex), le vocabulaire *Schema.org* compte environ 1000 éléments (428 classes et 582 propriétés) (« Home - schema.org », s. d.). Les *extraits enrichis* (en anglais *rich snippets*) affichés par les moteurs de recherche exploitent ce vocabulaire. Couvrant

¹¹³ Le nom de *Dublin [Core]* est issu du nom de la ville de la première rencontre en 1995 à *Dublin (Ohio, États Unis)*, accueillie par l'*Online Computer Library Center (OCLC)* et par le *National Center for Supercomputing Application (NCSA)*. Le *Core (noyau)* fait référence aux termes de métadonnées qui doivent être *larges et génériques*, utilisables pour décrire un spectre large de ressources (Dublincore.org, 2012a).

¹¹⁴ *Schema.org* (<http://schema.org/>) fournit une collection de schémas pour l'annotation sémantique du contenu sur le Web.

¹¹⁵ Publié aussi dans *IETF RFC 5013*, *ANSI/NISO Standard Z39.85-2007* et *ISO Standard 15836:2009*, le *Dublin Core* est employé par l'*Organisation mondiale de la santé* et adopté par l'*Union Européenne* en 2002.

¹¹⁶ Le *Dublin Core simple* se réfère aux 15 propriétés de base qui sont définies dans l'espace de nom <http://purl.org/dc/elements/1.1/>. Pour introduire le *domaine* et la *portée (l'ensemble d'arrivée)* des propriétés, depuis janvier 2008, le *Dublin Core Metadata Initiative* introduit l'espace de nom <http://purl.org/dc/terms/> qui englobe 55 propriétés et 22 classes (Dublincore.org, 2012b)

plusieurs domaines d'application, les classes principales¹¹⁷ sont *CreativeWork*, *Event*, *Intangible*, *MedicalEntity*, *Organization*, *Person*, *Place*, *Product*, *Property* (« Full Hierarchy - schema.org », s. d.).

Dans le cas de l'apprentissage, le contenu à gérer est constitué par des ressources d'apprentissage. De ce point de vue, le système d'apprentissage doit intégrer des fonctions de gestion des ressources d'apprentissage et un vocabulaire permettant l'enregistrement des métadonnées associées. Les termes utilisés doivent être adaptés à la description des entités participant à l'apprentissage, notamment les objets pédagogiques.

Dans le cadre d'un système d'apprentissage, plusieurs entités (e.g. objet pédagogique, logiciel, ordinateur) collaborent dans le but de produire l'apprentissage chez l'apprenant. Dans ce sens, chaque entité participant à l'apprentissage peut être considérée en tant que *ressource d'apprentissage*. Comme présenté plus haut, nous considérons comme *ressource d'apprentissage* tout *objet d'information* qui peut être utilisé pour l'apprentissage (p. 65). Le terme *ressource* implique l'attribution d'un identificateur de type *URI (Uniform Resource Identifier)*. Plus particulièrement et, de par sa relation avec l'objectif d'apprentissage, *l'objet pédagogique* est la ressource conçue spécialement pour l'apprentissage et donc celle disposant des caractéristiques les plus pertinentes et adaptées au contexte de l'apprentissage.

Le vocabulaire spécifique le plus utilisé pour la description des objets pédagogiques (numériques ou non) est le vocabulaire *Learning Object Metadata (LOM)*. Proposé au départ par *IMS Global Learning Consortium*, le vocabulaire LOM est ensuite développé par le groupe *Comité des Normes en Technologies d'Apprentissage*¹¹⁸ de l'*IEEE* qui l'approuve en tant que standard en 2002. L'*Association française de normalisation (AFNOR)* publie, en décembre 2006, le standard *NF Z76-040 (LOM-fr)* dérivé du standard LOM d'*IEEE* (v. ANNEXE E). En 2011, l'*Organisation internationale de normalisation*¹¹⁹

¹¹⁷ Les sous-classes directes de la classe *Thing*.

¹¹⁸ <http://www.ieeeeltsc.org>

¹¹⁹ Plus précisément, il s'agit des groupes ISO/IEC JTC1/SC36 : *Joint Technical Committee, Information Technology, Subcommittee SC 36, Information technology for learning, education and training*.

publie la première partie¹²⁰ du standard international *ISO/IEC 19788 : Technologies de l'information – Apprentissage, éducation et formation – Métadonnées pour les ressources d'apprentissage*, qui intègre et harmonise des éléments de DublinCore et LOM en essayant de mettre en évidence les aspects d'*objet pédagogique* de LOM et le modèle *entités-relations du Web sémantique* utilisé dans *Dublin Core Abstract Model*¹²¹ (Hoel et Mason, 2011).

Le vocabulaire LOM comprend plus de 80 éléments organisés dans une structure hiérarchique et regroupés en neuf catégories¹²² (v. Figure 2-12). La sémantique d'un élément est déterminée par son contexte (e.g. parent ou autre élément conteneur dans la hiérarchie).

Les neuf catégories de la spécification LOM (IEEE, 2002; AFNOR, 2006) sont les suivantes :

1. Général : regroupe les informations générales qui décrivent la ressource pédagogique dans son ensemble (e.g. identificateur, titre, langue, description) ;
2. Cycle de vie : décrit l'histoire et l'état actuel de la ressource pédagogique (e.g. version), ainsi que les entités affectées au cours de son évolution ;
3. Méta-métadonnées : décrit les métadonnées de la ressource pédagogique ; ce ne sont pas des descriptions de la ressource pédagogique, mais seulement de ses métadonnées (e.g. comment identifier une instance spécifique de métadonnée, auteur de cette instance) ;
4. Ressource pédagogique (e.g. format, taille) ;
5. Pédagogie : décrit les caractéristiques de nature pédagogique de la ressource tel que le type (e.g. cours, questionnaire), le public ciblé ou le contexte et le niveau d'utilisation (e.g. enseignement scolaire, licence) ;
6. Droits : spécifie les droits de propriété intellectuelle de la ressource pédagogique et ses conditions d'utilisation (e.g. coût, droits d'auteur) ;

¹²⁰ Il s'agit d'un *standard multiparties*, assurant à la fois une approche intégrée ou modulaire. Facilitant l'utilisation et la maintenance, chaque partie est indépendante (*self-contained*) avec son propre objectif. Il est structuré en 5 parties. 1 : Charpente ; 2 : Éléments Dublin Core ; 3 : Profil d'application de base ; 4 : Éléments techniques ; 5 : Éléments éducatifs ; 6 : Éléments de disponibilité, de distribution et de propriété intellectuelle.

¹²¹ <http://dublincore.org/documents/abstract-model/>

¹²² http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf

7. Relation : définit les relations entre la ressource pédagogique utilisée et les autres ressources pédagogiques (e.g. est une partie de, contient, est prérequis de, a pour prérequis) ;
8. Commentaire¹²³ : cette catégorie permet aux enseignants ou autres personnes (ou organisations) de partager leurs expériences (e.g. suggestions d'utilisation) en lien avec la ressource pédagogique et de commenter son utilisation ;
9. Classification : permet de *localiser* la ressource pédagogique dans un système de classification (e.g. taxonomie des objectifs pédagogiques) ; plusieurs instances de cette catégorie sont nécessaires pour permettre une classification multiple.

¹²³ Le terme utilisé dans la version originale en anglais est *Annotation*. Dans le document fourni par AFNOR le nom de cette catégorie est *Commentaire*.

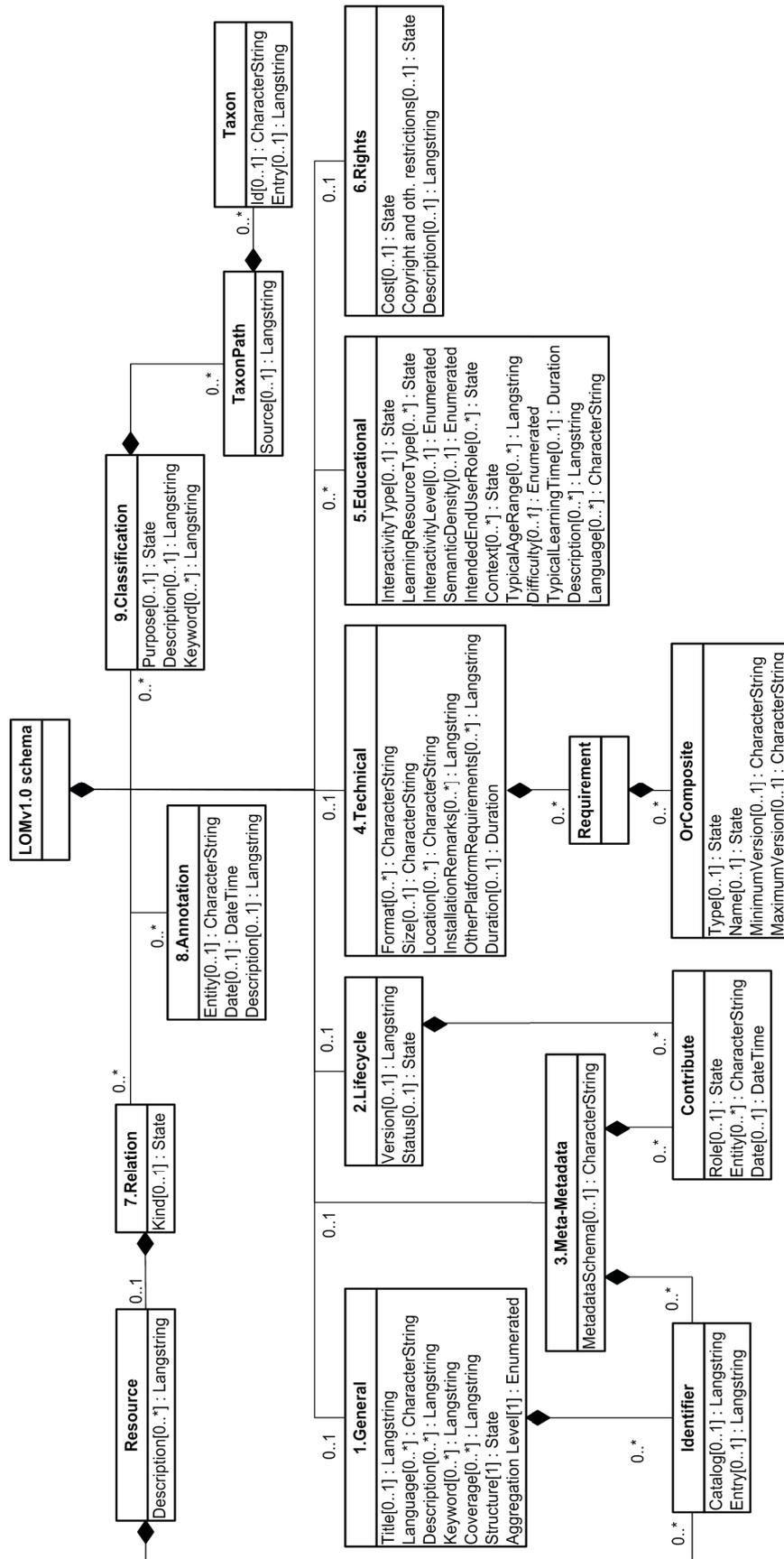


Figure 2-12. Les éléments et la structure hiérarchique du LOM

La catégorie *Pédagogie* présente les caractéristiques spécifiques pour la description de la ressource dans le contexte d'apprentissage. Ces caractéristiques constituent les éléments principaux permettant le choix et l'adaptation des ressources d'apprentissage aux spécificités de l'apprenant (e.g. style d'apprentissage). Les éléments constituant cette catégorie sont les suivants (IEEE, 2002; AFNOR, 2006) :

1. Type d'interactivité : quel est *le mode* d'apprentissage prédominant soutenu par cette ressource d'apprentissage ? les valeurs possibles pourront être : *active (pour un mode actif)*, *expositive (pour les objets de type exposé)* ou *mixtes (pour les objets qui combine les deux)* ;
2. Type de la ressource pédagogique : liste de types spécifiques de la ressource pédagogique, le plus pertinent devant être placé en premier (e.g. exercice, questionnaire, examen) ;
3. Niveau d'interactivité ou degré d'interactivité (le degré avec lequel l'apprenant peut influencer l'aspect ou le comportement de la ressource d'apprentissage) caractérisant la ressource pédagogique sur une échelle à 5 niveaux (très faible, faible, moyenne, élevé, très élevé) ;
4. Densité sémantique : le degré de *concision* de la ressource d'apprentissage, exprimé sur une échelle à 5 niveaux, peut être estimé en fonction du type de contenu, de la taille, ou de la durée nécessaire pour parcourir la ressource d'apprentissage (la densité sémantique est indépendante de la difficulté) ;
5. Public ciblé : l'utilisateur principal pour lequel la ressource a été conçue, le plus pertinent étant placé en premier (e.g. enseignant, auteur, apprenant, manager) ;
6. Niveau (contexte) : le contexte principal pour lequel la ressource a été conçue (e.g. école, université, stage, autre) ;
7. Tranche typique d'âge : l'âge de l'utilisateur pour lequel la ressource a été conçue (e.g. « 7-9 », « 15 », « 18- ») ;
8. Difficulté : le degré d'effort nécessaire pour travailler avec la ressource pédagogique en rapport avec l'audience type ciblée, sur une échelle à 5 niveaux ;
9. Durée d'apprentissage typique : le temps approximatif nécessaire pour travailler avec la ressource pédagogique en rapport avec l'audience type ciblée ;
10. Proposition d'utilisation (description) : commentaires sur *comment* la ressource pédagogique devrait être utilisée ;

11. Langue de l'utilisateur : langage humain utilisé par l'utilisateur typique auquel est destinée la ressource pédagogique ;
12. Activité induite¹²⁴ : activité proposée au public ciblé (e.g. communiquer, créer, observer, s'évaluer) ;
13. Validation des acquis¹²⁴ : quantité d'unités d'enseignement ou de crédits obtenus après avoir atteint les objectifs pédagogiques (e.g. « 40ECTS ¹²⁵ ») (« 5 Pédagogique - Sup-LOMFR », s. d.).

Dotées d'une structure hiérarchique, les métadonnées LOM sont généralement enregistrées en utilisant le langage XML¹²⁶. Les éléments du vocabulaire LOM ne possèdent pas un moyen d'identification unique et global (e.g. URI). Ainsi c'est leur position dans la structure XML qui détermine leur interprétation : l'élément *1.1. Identifier* est utilisé pour l'identification de la ressource d'apprentissage, et l'élément *3.1. Identifier* est utilisé pour l'identification de la métadonnée. L'utilisation du vocabulaire LOM avec XML est donc plus adaptée dans le cadre des SGA que sur le Web.

Cependant, l'utilisation du modèle *RDF*¹²⁷ (*Ressource Description Framework*) pour l'enregistrement de ces métadonnées permet d'associer des identificateurs uniques et globaux à chaque élément, indépendamment du contexte d'utilisation.

D'autres vocabulaires utilisés sur le Web peuvent être intégrés dans la description des objets pédagogiques. Même si le vocabulaire Dublin Core n'a pas pour objectif de proposer des termes spécifiques pour la description des ressources d'apprentissage, les 15 éléments trouvent des équivalences dans les catégories *Général* et *Cycle de vie* de la spécification LOM. Une correspondance des éléments LOM avec les éléments *Dublin Core* a été proposée par IEEE. Nous la présentons dans l'ANNEXE C.

Lancée en juin 2011, la *Learning Resource Metadata Initiative* (LRMI) se propose de faciliter la publication et la découverte des ressources d'apprentissage *sur le Web* en fournissant un vocabulaire dédié aux ressources d'apprentissage (« Learning Resource

¹²⁴ Élément présent seulement dans la version française proposée par AFNOR.

¹²⁵ ECTS : *European Credit Transfer and Accumulation System*.

¹²⁶ XML : *eXtended Markup Language*.

¹²⁷ <http://www.w3.org/RDF/>.

Metadata Initiative :: About the LRMI », s. d.). L'initiative, codirigée par l'*Association of Educational Publishers* et l'organisation *Creative Commons*, a développé un cadre de métadonnées pour la description des ressources d'apprentissage sur le Web. Stimulé par Schema.org, le projet LRMI se donne pour but premier l'intégration de son vocabulaire au lexique commun, soutenu dans cet effort par les acteurs majeurs de la recherche sur le Web. Soumise en mars 2012 sur la liste de diffusion publique du *World Wide Web Consortium* de vocabulaires pour le Web, la proposition est officiellement acceptée et intégrée dans le vocabulaire *Schema.org* en avril 2013 (« Learning Resource Metadata Initiative :: FAQ », s. d.).

Les nouveaux éléments proposés par LRMI et intégrés dans le vocabulaire Schema.org sont les suivants (« Learning Resource Metadata Initiative :: The Specification », s. d.) :

1. `educationalRole` : décrit le public cible du contenu d'apprentissage ;
2. `educationalAlignment` : cadre et contexte éducatifs ;
3. `educationalUse` : objectif de la ressource dans le contexte d'apprentissage (e.g. devoir, travail en groupe) ;
4. `timeRequired` : temps approximatif nécessaire à l'audience typique ciblée pour travailler avec la ressource d'apprentissage ;
5. `typicalAgeRange` : âge de l'utilisateur pour lequel la ressource d'apprentissage a été prévue
6. `interactivityType` : *mode* d'apprentissage prédominant soutenu par cette ressource d'apprentissage, les valeurs possibles étant *active*, *expositive* ou *mixte* ;
7. `learningResourceType` : type prédominant et caractérisant la ressource d'apprentissage (e.g. présentation) ;
8. `seeRightsUrl` : URL ou propriétaire spécifiant les droits d'utilisation de la ressource d'apprentissage ;
9. `isBasedOnUrl` : indique une ressource qui a été utilisée pour créer cette ressource d'apprentissage.

Dans l'approche LRMI, nous observons l'expansion des outils reconnus de l'apprentissage dans l'espace pervasif du Web. L'implémentation de ces descriptions de ressources d'apprentissage au niveau du Web améliore la découverte de contenus d'apprentissage et permet la personnalisation du processus d'apprentissage à l'échelle du Web. Grâce à l'utilisation de vocabulaires structurés spécifiques pour l'apprentissage,

l'organisation des ressources d'apprentissage sur le Web permet de considérer le Web tout entier comme un système d'apprentissage.

2.3. Exemples des systèmes d'apprentissage basés sur le Web

Les systèmes d'apprentissage fondés sur le Web ont émergé de la catégorie des *systèmes de gestion de contenu (SCG)*¹²⁸ pour le Web. Ces systèmes d'apprentissage sont spécialisés dans la gestion de ressources et de situations d'apprentissage. Ainsi nous parlons de *systèmes de gestion de cours*, de *système de gestion d'apprentissage (SGA)*¹²⁹ ou de *systèmes de gestion de contenu d'apprentissage (SGCA)*¹³⁰. Même si certaines fonctionnalités se retrouvent dans chaque déclinaison, ces systèmes ont des orientations différentes. Les systèmes de gestion de *cours* gèrent l'inscription des apprenants, leurs performances, ainsi que la création et la diffusion de contenus de cours (« Course Management Systems and the Reinvention of Instruction », 2004).

Les systèmes de gestion de *l'apprentissage* sont conçus pour permettre l'administration automatique des *activités d'apprentissage*. Ils permettent d'intégrer et d'administrer des contenus et des cours provenant de différents fournisseurs. Le SGA ne fournit pas des outils permettant la création de contenus d'apprentissage, mais il se concentre sur l'administration de cours fournis par plusieurs distributeurs (« LMS and LCMS Demystified », 2006).

Les systèmes de gestion de *contenu d'apprentissage* sont principalement utilisés pour la création, le développement, la maintenance, l'étiquetage et le stockage du contenu d'apprentissage. Ces systèmes fournissent des outils permettant l'annotation des objets pédagogiques à l'aide de métadonnées ainsi que des outils pour créer dynamiquement un objet pédagogique à partir d'objets d'information. Même si les SGCA peuvent délivrer du contenu d'apprentissage, ils ne fournissent pas les mêmes fonctionnalités qu'un SGA (« E-Learning Glossary », s.d.-b). Parmi les systèmes de gestion de contenu et

¹²⁸ Le SGC peut être utilisé pour *système de gestion de contenu* ou *système de gestion de cours*.

¹²⁹ En anglais *Learning Management System (LMS)*.

¹³⁰ En anglais *Learning Content Management System (LCMS)*. LCMS = LMS +CMS.

d'apprentissage les plus connus, on trouve *Moodle*¹³¹, *Blackboard*¹³², *Sakai*¹³³ ou *WebCT*¹³⁴ (Grant, 2012).

La diversité et l'abondance de ressources d'apprentissage ont incité la mise à disposition de ressources d'apprentissage sous une licence libre de droits : « Les termes "cours ouverts" font référence à la volonté de proposer des cours académiques disponibles librement, notamment par le moyen de l'Internet et du Web » (UNESCO, 2002). Lors du premier forum mondial des *Ressources Éducatives Libres*¹³⁵ (REL) organisé à Paris en 2002, l'UNESCO¹³⁶ a proposé le terme de *ressources éducatives libres*, pour décrire sa volonté de stimuler « l'accès universel à une éducation de qualité » (UNESCO, s. d.). L'acronyme *REL* désigne « des matériels d'enseignement, d'apprentissage et de recherche sur tout support, numérique ou autre, existant dans le domaine public ou publiés sous une licence ouverte permettant l'accès, l'utilisation, l'adaptation et la redistribution gratuits par d'autres, sans restrictions ou avec des restrictions limitées. Les licences ouvertes sont fondées dans le cadre existant du droit à la propriété intellectuelle, comme défini par les conventions internationales concernées, et respectent la paternité de l'œuvre » (UNESCO, 2012). Les REL constituent un pas important dans l'apparition des *MOOCs* (*Massive Open Online Courses*, en français : *cours en ligne ouverts et massifs*) et la « démocratisation de l'éducation » (*Welcome to the Brave New World of MOOCs (Massive Open Online Courses)*, 2013). En même temps, se sont développés des systèmes de diffusion de ce type de cours, tels que *Coursera*, *Udacity*, *edX* ou *Khan Academy*, *Iiversity*, *FUN* (France Université Numérique).

Un système informatique d'apprentissage *basé sur le Web* utilise l'architecture et les technologies du Web (e.g. protocole HTTP, langage HTML). Comme les applications Web, ces systèmes sont accessibles depuis n'importe quel dispositif connecté à Internet à travers un navigateur Web. Les technologies standardisées du Web assurent le

¹³¹ Modular Object-Oriented Dynamic Learning Environment : <https://moodle.org/>.

¹³² <http://www.blackboard.com>.

¹³³ <https://sakaiproject.org/>.

¹³⁴ Web Course Tools : <http://www.webct.com>

¹³⁵ Les *Ressources Éducatives Libres* - en anglais *Open Educational Resources - OER*. La plateforme UNESCO des REL se trouve à l'adresse <http://www.oercommons.org/>.

¹³⁶ L'Organisation des Nations Unies pour l'éducation, la science et la culture : <https://fr.unesco.org/>.

déploiement des SGA sur les différents navigateurs et systèmes d'exploitation. Les SGA et le contenu d'apprentissage doivent être construits en respectant les standards et le modèle de diffusion du Web. Ces systèmes suivent l'évolution de ce dernier.

Le contenu d'apprentissage, constitué de cours, est le premier à s'adapter à la diffusion massive de l'apprentissage. Ainsi, soulignée par D. Cormier et B. Alexander (R. L. Herman, 2012; *What is a MOOC?*, 2010), l'expression *Massive Open Online Course* exprime la tendance et l'approche actuelle de l'apprentissage en ligne. L'appellation est inspirée d'un cours enseigné à l'*Université de Manitoba* en 2008 par G. Siemens et S. Downes. En plus des 25 étudiants inscrits (qui ont payé la participation), le cours était ouvert et gratuit pour encore environ 2300 étudiants supplémentaires (R. L. Herman, 2012). Le cours était *en ligne*, il était *ouvert* pour la majorité des participants et, étant donné le nombre de participants, nous pouvons dire qu'il était *massif*. En détaillant chaque élément composant l'acronyme, un MOOC est :

- massif : en principe, le cours peut accueillir un nombre non limité de participants ;
- ouvert : le cours est ouvert à tous les internautes, sans distinction d'origine, de niveau d'études ou d'un quelconque critère ;
- en ligne : l'ensemble du contenu du cours (e.g. activités, devoirs, examens) peut être suivi en ligne ; les échanges avec les participants au cours peuvent se faire depuis leur propre espace virtuel (e.g. blog personnel) ;
- cours : un contenu structuré d'apprentissage, avec des objectifs pédagogiques et non simplement des ressources diffusées en ligne (*What is a MOOC?*, 2010; Stevens, 2013; *Business and MOOCs*, 2013, « France Université Numérique - FUN - Découvrir, apprendre et réussir », s. d.).

Un MOOC intègre *la connectivité des réseaux sociaux, la facilitation d'un expert reconnu dans le domaine d'étude* et une collection de *ressources en ligne librement accessibles*. Le MOOC s'appuie sur la participation active de plusieurs centaines à plusieurs milliers de participants qui s'auto-organisent en fonction d'objectifs d'apprentissage, de connaissances, de compétences et d'intérêts communs.

Le MOOC existe en réseau selon un modèle distribué qui s’aligne avec l’architecture et les paradigmes du *Web sémantique* (Masters, 2011). Dans ce réseau, l’importance des SGA est réduite à des nœuds participant à la connexion avec les autres éléments du cours. Le Web entier devient la plateforme sur laquelle reposent les éléments du MOOC.

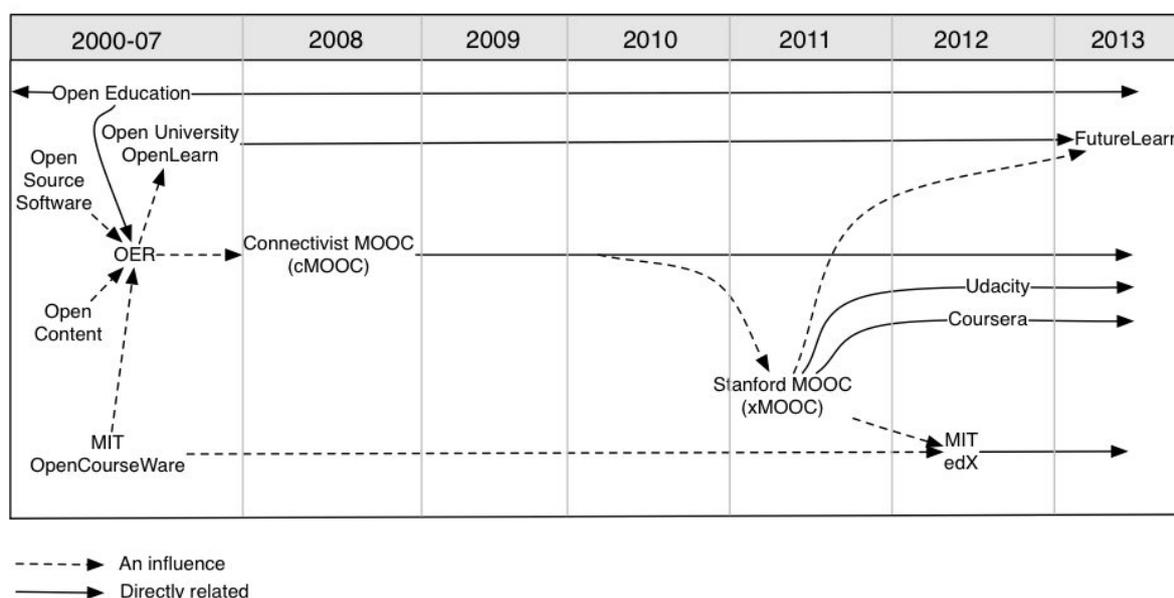


Figure 2-13. La chronologie des MOOCs

À l’heure actuelle, les acteurs principaux dans l’éducation en ligne sont *Coursera*, *Udacity*, *EdX*, *Iversity*, *FutureLearn* et *Khan Academy*¹³⁷ (v. Figure 2-13). Afin de « proposer une offre innovante de formations en ligne » et « de favoriser la réussite des étudiants grâce à une pédagogie renouée par le numérique », en octobre 2013, à l’initiative du ministère de l’enseignement supérieur a été lancée en France la plateforme FUN (France Université Numérique). Il s’agit d’une « plateforme nationale pour héberger, sur un site unique, les formations numériques, MOOC certifiants et formations en ligne diplômantes ou qualifiantes ».

2.3.1. Contenu et ressources d’apprentissage dans les SGA

Afin de favoriser la réutilisabilité des contenus d’apprentissage, la plupart des systèmes conventionnels de gestion de l’apprentissage utilisent des standards qui permettent l’interopérabilité et l’échange de ces contenus. Ces derniers sont « encapsulés » dans des

¹³⁷ *Blackboard* propose son catalogue en ligne des cours de type MOOC à l’adresse <http://www.coursesites.com/mooccatalog>

objets partageable appelés « sco » (*sharable content object*). Chaque objet contient une unité d'apprentissage conforme au profil d'application SCORM¹³⁸ (Sharable Content Object Reference Model). Le profil SCORM comprend un ensemble de standards utilisés dans l'e-learning pour faciliter l'accessibilité, la réutilisabilité et l'interopérabilité des ressources d'apprentissage dans le contexte des SGA (ADL, 2001). Cette collection est proposée à l'initiative d'*Advanced Distributed Learning* en 1997 avec l'objectif de grouper et de moderniser le management de l'éducation et des ressources électroniques d'apprentissage. SCORM spécifie les deux principaux aspects suivants: comment l'unité d'apprentissage doit être « emballée » (e.g. structure hiérarchique, métadonnées, séquence du contenu) et comment s'effectue l'échange de données entre cette unité d'apprentissage et le SGA (e.g. SCORM API). SCORM utilise le standard LOM pour décrire les métadonnées du contenu d'apprentissage.

Ce modèle impose l'existence de ressources d'apprentissage dans le SGA (le contenu et le SGA sur le même serveur). Cette limitation est dépassée dans la dernière génération SCORM, appelée Tin Can API¹³⁹ (ou *Experience API*). Cette nouvelle spécification permet l'agrégation des traces d'apprentissage à partir de plusieurs sources et donc l'utilisation des ressources d'apprentissage existantes sur d'autres serveurs. Les traces des apprenants sont enregistrées dans un *Learning Record Store* sous forme de « substantif, verbe, objet »¹⁴⁰. Cela permet d'évaluer l'évolution de l'apprentissage d'un apprenant dans son ensemble. Afin d'enregistrer le contexte de l'apprenant pour permettre la personnalisation du processus d'apprentissage Butoianu *et al.* (Butoianu *et al.*, 2010) présentent un outil permettant l'acquisition des *métadonnées d'attention contextualisée* des apprenants.

Ainsi les systèmes d'apprentissage élargissent leur espace d'apprentissage et peuvent intégrer des ressources d'apprentissage stockées sur des plateformes différentes (e.g. vidéos – Youtube, Vimeo ; présentations – Slideshare). La majorité des contenus d'apprentissage de type *présentation* de ces nouveaux MOOCs se trouve sur Youtube (e.g. Khan Academy, EdX, Udacity). Il s'agit de cours enregistrés et organisés dans des

¹³⁸ <http://scorm.com/>

¹³⁹ Le projet est lancé en Octobre 2011 et la version 1.0 est lancée en Avril 2013.

¹⁴⁰ La structure des déclarations renvoie au format des triplets RDF. <http://tincanapi.com/overview/>

unités de 5 à 15 minutes environ afin de faciliter l'apprentissage. Dans le processus d'apprentissage, ces présentations sont enchaînées avec d'autres types de contenu, notamment des exercices pratiques. Les MOOCs proposent des espaces de collaboration et d'interaction entre apprenants. Toutefois ces systèmes ne permettent pas une adaptation du parcours d'apprentissage et un accompagnement intelligent de l'apprenant dans ses démarches. Ces tâches sont prises en charge par les autres participants à l'apprentissage (enseignants ou collègues).

2.3.2. Caractéristiques des systèmes d'apprentissage actuels

Afin d'identifier les fonctionnalités des systèmes d'apprentissage, nous avons analysé six des principaux acteurs du domaine: MITx (le pré-EdX, v. ANNEXE D), Udacity, Khan Academy, Coursera et Open HPI. Sans insister sur l'évaluation de ces systèmes, nous rappelons quelques caractéristiques communes de fonctionnement.

Le parcours d'apprentissage est linéaire et le scénario pédagogique est prédéfini à l'avance par l'enseignant. Les lectures vidéo sont alternées avec des exercices ou devoirs qui sont évalués de façon automatique. Le co-constructivisme est le modèle de base exploité dans ces systèmes dans le cadre des échanges entre les participants à l'apprentissage. Ils sont encouragés à échanger, à interagir entre eux dans les espaces destinés (e.g. discussion, forum, wiki). Dans le cadre du cours pilote de MITx (*Circuits and Electronics*), le système fournit aussi un outil adapté pour des exercices pratiques (« Circuit Sandbox »). Ces systèmes n'intègrent pas des méthodes d'adaptation ou de personnalisation du contenu d'apprentissage. Enfin, la réussite à ces cours donne droit à un certificat qui atteste la participation et la validation du cours tout en respectant un « code d'honneur ».

Le nombre des participants aux MOOCs est parfois de l'ordre de centaines de milliers. Toutefois, le taux d'accomplissement de ces cours se situe en dessous de 13 pour cents¹⁴¹. La dynamique des cours, une grande différence entre les niveaux des

¹⁴¹ <http://www.katyjordan.com/MOOCproject.html>. Pour le cours pilote de MITx les statistiques sont : « 6.002x had 154,763 registrants. Of these, 69,221 people looked at the first problem set, and 26,349 earned at least one point on it. 13,569 people looked at the midterm while it was still open, 10,547 people got at least one point on the midterm, and 9,318 people got a passing score on the midterm. 10,262 people looked at the final exam while it was still open, 8,240 people got

apprenants ou l'interaction directe limitée avec l'enseignant peuvent souvent avoir un effet décourageant sur les apprenants. La majorité de systèmes d'apprentissage actuels sont de type « one size fits all ». Le cours se présente de la même manière pour chaque apprenant. Il a la même composition pour tous les apprenants, sans prendre en compte les capacités ou le profil de compétences de ceux-ci.

Ces systèmes ne fournissent pas des moyens d'adaptation ou de personnalisation du processus d'apprentissage. Si plusieurs services et domaines sur le Web implémentent déjà des approches d'adaptation ou de personnalisation (e.g. commerce, recherche Web, tourisme, films), dans le cas de l'apprentissage la personnalisation est un aspect plus complexe.

Il est important de noter que la classification et l'identification pertinente des ressources d'apprentissages selon des critères pédagogiques imposent l'enregistrement des métadonnées correspondantes. En effet, sans une description détaillée exploitable par la machine, un contenu d'apprentissage reste peu réutilisable en dehors du contexte dans lequel il a été créé.

Une catégorie de systèmes répondant à ces limites est celle des *tuteurs intelligents* (*ITS – Intelligent Tutoring System*). Ces tuteurs sont des systèmes d'information qui utilisent des techniques de l'intelligence artificielle pour interagir avec l'apprenant. Ils fournissent des instructions immédiates et personnalisées aux actions des apprenants (Psootka, Massey, et Mutter, 1988). Le développement de ces systèmes se trouve à l'intersection de l'informatique, de la psychologie cognitive et des sciences de l'éducation (Nwana, 1990). Les premiers ITS ont été conçus comme un complément aux autres méthodes d'apprentissage. Leur rôle principal était de soutenir l'apprenant dans la résolution des problèmes. Il était assumé que la connaissance nécessaire à la résolution des problèmes est acquise à l'extérieur du système, donc ces systèmes n'offrent pas de contenus d'apprentissage. Toutefois, il est devenu évident que le

at least one point on the final exam, and 5,800 people got a passing score on the final exam. Finally, after completing 14 weeks of study, 7,157 people have earned the first certificate awarded by MITx, proving that they successfully completed 6.002x» (<https://6002x.mitx.mit.edu/info>).
<http://halfanhour.blogspot.fr/2013/04/moocs-for-credit-current-state-of-art.html>

couplage d'un système hypermédia avec un système ITS peut bénéficier des avantages des deux catégories de systèmes (Brusilovsky, 2000).

Basé sur le modèle *Dexter*¹⁴², un des premiers modèles d'adaptation dans le domaine de l'e-learning est le modèle AHAM (Adaptive Hypermedia Application Model) (De Bra, Houben, et Wu, 1999). Il distingue trois parties distinctes dans l'adaptation : un modèle de domaine, un modèle d'utilisateur et un modèle d'adaptation qui consiste dans des règles d'adaptation. Étant donnée la structure du *document hypermedia* (ensemble des *nœuds* connectés par des *liens*), Brusilovsky présente deux approches de l'adaptation : (1) l'adaptation au niveau de la présentation et (2) l'adaptation au niveau du support de navigation (Brusilovsky, 2001).

Lefevre *et al.* présentent les résultats d'une étude menée sur plusieurs dispositifs de formation. Cette étude avait pour buts « d'une part d'identifier les besoins actuels en terme de personnalisation de l'apprentissage et d'autre part de comparer des approches permettant cette personnalisation » (Lefevre *et al.*, 2012). Ils évaluent quatre systèmes, notés de A1 à A4¹⁴³, permettant la personnalisation automatisée de l'apprentissage et présentent une synthèse mettant en évidence la nécessité d'exploiter soit un profil d'apprenant, soit ses traces dans le système. Toutefois, « la principale question posée par l'automatisation du processus de personnalisation de l'apprentissage, fût-elle partielle, relève de *l'ingénierie des connaissances* » (ibid.).

D'un point de vue systémique, l'adaptation de l'apprentissage est une boucle de régulation pédagogique. Altet souligne l'importance de la « dynamique de la régulation pédagogique » (Altet, 1997, p. 15) réalisée entre le processus *enseigner* et le processus *apprendre*. Il propose en conséquence un modèle systémique de l'apprentissage (v. Figure 2-14).

¹⁴² Le modèle Dexter (Halasz et al., 1994) est un modèle de référence pour l'architecture des systèmes hypermédia. Il structure l'application hypermédia en trois couches : composante interne (le contenu ou la structure intérieure aux nœuds), stockage (une base de données contenant un réseau des nœuds et liens) et exécution (la présentation de l'hypertexte, l'interaction de l'utilisateur).

¹⁴³ A1 : (Szilagyi, Greffier, et Domenget, 2011) ; A2 : (Butoianu et al., 2011) ; A3 : (Daubigney, Geist, et Pietquin, 2011) ; et A4 : (Lefevre, Jean-Daubias, et Guin, 2011).

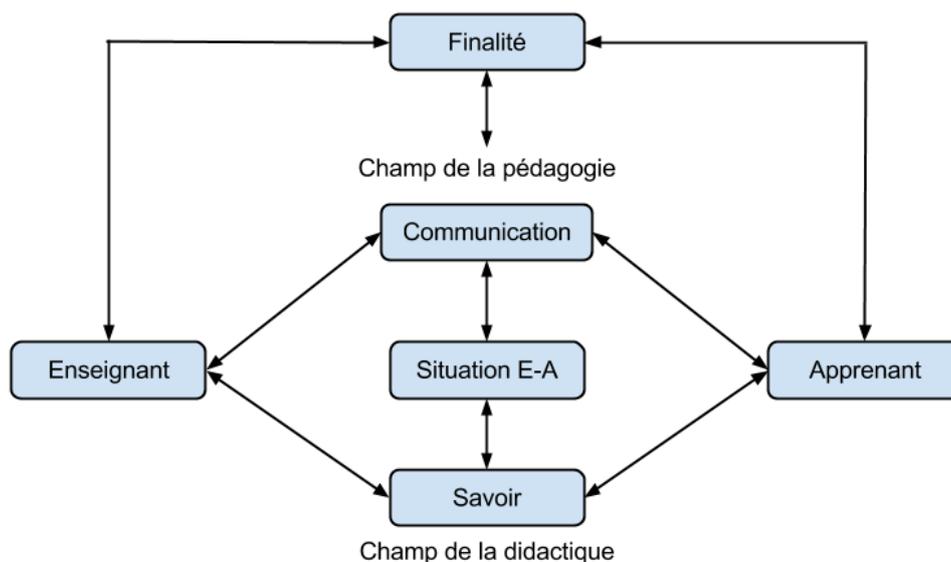


Figure 2-14. Un modèle systémique de l'apprentissage (selon de Rézeau, 2001, p. 42)

Aux trois entités contenues dans le triangle pédagogique de J. Houssaye (v. page 26), s'ajoutent ici des caractéristiques spécifiques aux *systèmes*. À l'organisation des éléments, aux relations de l'apprenant avec le savoir (champ didactique) et avec l'enseignant (champ pédagogique), s'ajoutent les notions de *finalité* et de *boucle de rétroaction*. Les éléments du système sont organisés ayant comme but l'apprentissage. Ces éléments sont variés et peuvent eux aussi être des systèmes (e.g. apprenant, objet pédagogique, enseignant, système de communication, système informatique).

Le système d'apprentissage proposé dans cette thèse (Système Actif et Sémantique d'Apprentissage - SASA) est dans un premier temps un *système conceptuel*. Nous considérons les *éléments* de l'apprentissage et proposons ensuite une modélisation, une organisation et une interaction entre ces éléments. Le système conceptuel, ainsi que les éléments le composant, sont présentés en détail dans le Chapitre 5. Modélisation du Système Actif et Sémantique d'Apprentissage (SASA). Ce *système conceptuel* se matérialise à travers une *modélisation* implémentée dans le cadre d'un *système informatique*. En intégrant les notions de finalité (objectifs d'apprentissage) et de rétroaction (interaction avec l'apprenant, système actif), le système informatique proposé est plus qu'un système de gestion de contenu d'apprentissage. Il permet l'expression *du système conceptuel* de l'apprentissage. On trouvera plus de détails sur ce système informatique dans le Chapitre 6. Implémentation du SASA.

Nous considérons qu'un système d'apprentissage efficace englobe des caractéristiques existantes dans les systèmes énumérés (LMS, ITS, MOOCs). Comme la plupart des LMS, un système d'apprentissage efficace doit bénéficier des avantages du Web, doit intégrer un certain niveau d'intelligence et doit être disponible aux apprenants, peu importe leur nombre. Nous considérons les avantages et les possibilités offertes par les technologies sémantiques pour la construction d'un système d'apprentissage construit sur la nouvelle génération du Web. Dans un premier temps, l'idée est d'aider l'enseignant dans son choix et de déléguer à la machine la fonction de recherche des contenus d'apprentissage répondant au choix formulé. Dans un second temps, le défi est de rendre intelligente cette fonction de recherche de contenu d'apprentissage, afin que le système puisse assister l'apprenant dans son parcours d'apprentissage.

Dans ce qui suit, nous nous concentrons sur l'utilisation des technologies du Web sémantique, d'une part, pour la description des ressources d'apprentissage (déclaration des métadonnées) et, d'autre part, pour la construction des modèles nécessaires pour la personnalisation de l'apprentissage.

Chapitre 3. Le Web sémantique et la représentation de l'information

« Les dogmes du paisible passé ne sont plus à la hauteur de l'orageux présent. L'occasion offre un amoncellement de difficultés et il faut nous élever à la hauteur de l'occasion. De même que notre position est nouvelle, il nous faut de nouvelles pensées et des actes nouveaux ; il faut nous affranchir nous-mêmes, et alors nous sauverons le pays ».
Abraham Lincoln, 1^{er} décembre 1862

Plusieurs idées, apparues au fil des années, renvoient aux efforts de construction d'un savoir universel organisé, ouvert et accessible en temps réel. Dans les démarches d'acquisition de ce savoir, plusieurs moments clés, projets et propositions apportent leur contribution. La représentation du monde, la représentation de l'information, l'organisation de l'information et l'accès rapide à l'information ont toujours été essentiels au développement de la société. La construction des relations entre les entités d'information, l'élaboration des agents intelligents qui peuvent calculer et raisonner sur l'ensemble des informations ont constitué l'étape suivante dans l'évolution des systèmes de stockage et gestion du savoir. Plusieurs de ces idées se retrouvent dans l'évolution des systèmes d'information et du Web. Le Web sémantique s'inscrit quant à lui dans les démarches d'amélioration du transfert, de l'enregistrement et de l'accessibilité du savoir.

3.1. L'utopie d'un savoir universel

Après l'invention de l'imprimerie, au début du 17^{ème} siècle, le philosophe et chancelier d'Angleterre Francis Bacon¹⁴⁴ eut l'idée et l'intention de *réorganiser les savoirs*¹⁴⁵. Il

¹⁴⁴ Sir Francis Bacon (1561 –1626) est reconnu comme le premier théoricien de la méthode expérimentale.

¹⁴⁵ Cette idée renvoie à la devise de Google : « organiser l'information à l'échelle mondiale ».

présenta ce projet dans l'ouvrage *La nouvelle Atlantide*¹⁴⁶ (Bacon, 1627). Il y décrit une île, appelé *Bensalem*¹⁴⁷, qui est gouvernée par une société philosophique savante : *La Maison de Salomon*¹⁴⁸. La science, au cœur des préoccupations de cette île, vise au progrès scientifique et à l'amélioration de la vie¹⁴⁹.

Dans l'organigramme de la planification de la science utile¹⁵⁰ de cette île, seuls « douze "Marchands de lumière" ont le droit de sortir du pays pour parcourir le monde à la recherche de livres, de résumés et de modèles d'expérience » (Mattelart, 2009, p. 7) « Nous avons douze collègues qui voyagent à l'étranger et qui nous rapportent des livres, des échantillons et des exemples d'expériences de toutes les régions du monde, ceci en se faisant passer pour des gens d'autres nationalités, puisque nous cachons la nôtre. Nous les appelons les Marchands de Lumière. » (Bacon, 1627, p. 129) L'idée ne semble-t-elle pas évoquer les robots d'indexation qui parcourent le web pour récupérer d'information (« Recherche d'information & utopie », 2012) ? L'idée d'organiser tout le savoir dans une encyclopédie inspire aussi D. Diderot¹⁵¹ et Jean le Rond d'Alembert¹⁵². Leur idée est, encore une fois, d'organiser toute la connaissance humaine. Ils publient l'Encyclopédie (Diderot et d' Alembert, 1751), un ouvrage en 28 volumes comprenant 71.818 articles et 3.129 illustrations, ouvrage auquel contribuent environs 100

¹⁴⁶ Écrit en latin (*Nova Atlantis*) en 1622, l'ouvrage était traduit en anglais en 1627 sous le titre de *New Atlantis*. *La nouvelle Atlantide* renvoie à l'île légendaire décrite dans le *Critias* de Platon.

¹⁴⁷ *Bensalem* signifie « fils parfait » en hébreu. L'idée de l'île et son nom même renvoient au mythe platonicien de l'Atlantide.

¹⁴⁸ « En 1662, le plan baconien pour la réorganisation des connaissances se matérialisa dans la *Royal Society of London for Improving Natural Knowledge by Experiments* » (Mattelart, 2009, p. 7)

¹⁴⁹ « C'est une société parfaite et utopique totalement rationalisée, poussée par une évolution interne grâce aux progrès historiques des sciences et des arts. Elle est dirigée par une élite très hiérarchisée de savants aidés de techniciens. Les habitants de l'île *Bensalem* se déplacent à l'aide de machines volantes ou encore explorent les océans dans des engins sous-marins. Ils étudient les phénomènes météorologiques et utilisent l'énergie hydraulique. Ils pratiquent également la chirurgie, les greffes d'organes et analysent le sang pour mieux comprendre les maladies. Dans l'île de *Bensalem*, la médecine fait effectivement de grands progrès, les recherches ont d'ailleurs éliminées toutes les maladies. C'est une utopie monarchique qui admet la propriété privée, le commerce et l'argent. Le mariage est l'institution primordiale de l'île dont les moeurs excluent la prostitution, l'adultère, la polygamie et l'homosexualité. » (« Recherche d'information & utopie », 2012).

¹⁵⁰ Les scientifiques de l'île de *Bensalem* travaillent selon une division rigoureusement hiérarchisée du travail, dans les activités de collecte, de classement et de traitement d'information (Mattelart, 2009, p. 7).

¹⁵¹ Denis Diderot (1713-1784) écrivain, philosophe et encyclopédiste français.

¹⁵² Jean-Baptiste le Rond d'Alembert (1717-1783), mathématicien, philosophe et encyclopédiste français.

personnes. « *Le but d'une Encyclopédie est de rassembler les connaissances éparses sur la surface de la terre ; d'en exposer le système général aux hommes avec qui nous vivons, et de le transmettre aux hommes qui viendront après nous ; afin que les travaux des siècles passés n'aient pas été des travaux inutiles pour les siècles qui succéderont ; que nos neveux, devenant plus instruits, deviennent en même temps plus vertueux et plus heureux, et que nous ne mourions pas sans avoir bien mérité du genre humain.* » (*Encyclopédie*, art. *Encyclopédie*, Diderot), (Blom, 2005, p. 139).

Aujourd'hui, ces efforts d'organisation de la connaissance renvoient aux moteurs de recherche sur le Web et aux encyclopédies en ligne. Nous rappelons trois projets dans lesquels se retrouvent ces idées centrales dans la construction du Web.

3.1.1. Mundaneum : « savoir au bout des doigts »

P. Otlet¹⁵³ « *l'homme qui voulait classer le monde* » (Levie, 2006), fondateur de l'*Office International de Bibliographie Sociologique (OIBS)*, propose avec H. La Fontaine¹⁵⁴ en 1905 le système de *classification décimale universelle (CDU)*, basé sur le système de *classification décimale de Dewey*¹⁵⁵ (CDD). Précurseur de la science de l'information, qu'il a appelé *documentation*, sa proposition d'*organisation de la documentation* renvoie à la pyramide de la sagesse (donné – information – connaissance - sagesse) (v. Figure 3-1).

¹⁵³ Paul Otlet (1868 – 1944), juriste belge, visionnaire, avait un grand intérêt pour la bibliothèque et la bibliographie. Il est souvent considéré comme le père du management de l'information (*Alle Kennis van de Wereld (Biography of Paul Otlet)*, 1998) et il est à l'origine de l'Internet en Europe (« The origins of the Internet in Europe - Google Cultural Institute », s. d.)

¹⁵⁴ Henri-Marie La Fontaine (1845 – 1943), docteur en Droit, professeur de droit international, sénateur, président du Bureau International de la Paix, reçoit le prix Nobel de la paix en 1913 (« Henri La Fontaine - Biographical », s. d.). H. La Fontaine et P. Otlet organisent en 1895 à Bruxelles la « Première Conférence de Bibliographie, qui décidera la création de l'Institut International de Bibliographie, dont l'objectif est l'organisation du Répertoire Bibliographique Universel (RBU) [...] Le RBU devait comprendre les notices de tous les livres, revues et articles contenus dans les publications périodiques depuis l'invention de l'imprimerie » (*Otlet et la bibliologie*, 2011, p. 74). Cette indexation et organisation de connaissance donne naissance au projet *Mundaneum*. Proposé par P. Otlet et H. La Fontaine, l'objectif de *Mundaneum* est de collecter toutes les connaissances du monde et de les indexer dans un système de fiches bibliographiques.

¹⁵⁵ La classification décimale de Dewey (CDD) était développée en 1876 par Melvil Dewey.

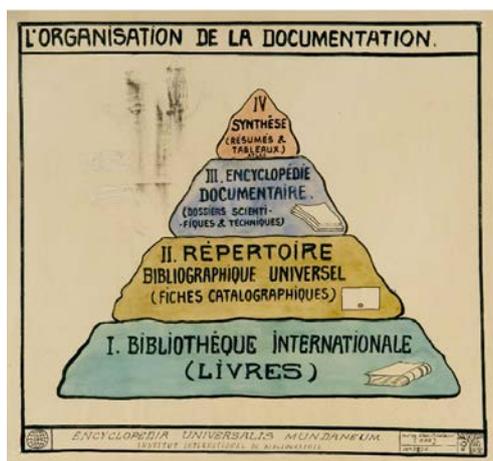


Figure 3-1. L'organisation de la documentation selon P. Otlet (source : « The origins of the Internet in Europe - Google Cultural Institute », s. d.)

Dans la perspective de rendre plus accessible l'information, le répertoire universel bibliographique *Mundaneum* devait permettre de répondre à toute question quel que soit le sujet. En effet, l'objectif de *Mundaneum* est de collecter toutes les connaissances du monde et de les indexer dans un système de fiches bibliographiques. Pour ce faire, des niveaux de traitement de l'information sont définis (Otlet, 1934). Les éléments de son schéma, voir la Figure 3-2, renvoient à l'architecture en couches des technologies du Web sémantique.

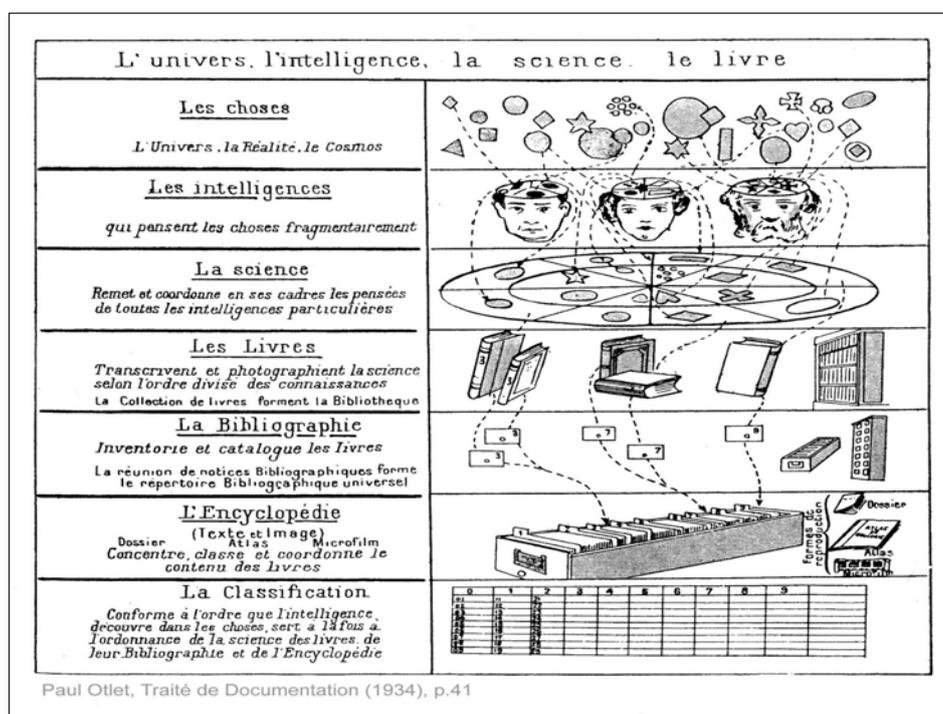


Figure 3-2. L'accès à l'information selon P. Otlet

Il utilise des termes comme *lien* et *réseau de connaissances* pour décrire sa vision d'un endroit central contenant toute la connaissance humaine. « *Le réseau, de quelque manière que ce soit, doit relier les uns aux autres les centres producteurs, distributeurs, utilisateurs, de toute spécialisation et de tout lieu. Il s'agit pratiquement, que tout producteur ayant quelque donnée à faire connaître, quelque proposition à présenter ou à défendre - tout utilisateur ayant quelque information à recueillir pour avancer son travail théorique ou pratique - toute personne enfin puisse au moindre effort et avec un maximum de sûreté et d'abondance, entrer en possession de ce qui lui est offert.*» (ibid., p. 415)

Il anticipe ainsi l'invention de l'Internet, du Web et des moteurs de recherche : « *De loin, tout le monde sera en mesure de lire le texte, agrandi et limité à l'objet désiré, projeté sur un écran individuel. De cette façon, tout le monde à partir de son fauteuil sera en mesure de contempler la création, dans son ensemble ou dans certaines de ses parties* » (Otlet, 1935, p. 391).

3.1.2. Memex : mémoire associative

Toujours dans l'anticipation des technologies modernes et du Web, Vannevar Bush (1890 – 1974), coordinateur de la recherche scientifique des États-Unis pendant la seconde guerre mondiale, tire la leçon de son expérience et propose plusieurs projets révolutionnaires pour son temps¹⁵⁶. Ainsi, en s'inspirant du modèle de fonctionnement du cerveau humain¹⁵⁷, il propose en 1945 le *Memex (memory extender)*.

¹⁵⁶(1) *Cylops Camera* : porté sur le front en face de la tête, ce caméra pourrait photographier tout ce que vous voyez et voulez enregistrer ; (2) *Microfilm* – l'enregistrement condensé de l'information, « il pourrait réduire le volume de l'encyclopédie britannique à la taille d'une boîte d'allumettes », « une bibliothèque entière pourrait être stockée dans un bureau » ; (3) *Vocoder* : une *machine* qui pourrait taper à notre place quand on lui parle ; (4) *Thinking machine* : une calculatrice mathématique développée ; en lui fournissant les prémisses elle sera capable de calculer la conclusion conformément à une logique ; (5) *Memex* : (*memory extender*) un dispositif d'aide à la mémoire, comme le cerveau, *Memex* pourrait associer l'information et ensuite vous accompagner dans la navigation sur le chemin des informations résulté de ces associations (Bush, 1945, p. 113). Plusieurs concepts et idées de ses projets sont aujourd'hui implémentés : *Google Glass*, *Siri* de l'*iPhone*, les moteurs de raisonnement développés dans le Web sémantique, etc.

¹⁵⁷ Sections de texte original de V. Bush (*As we may think*) : "When data of any sort are placed in storage, they are filed alphabetically or numerically, and information is found (when it is) by tracing it down from subclass to subclass [...] The human mind does not work that way. It operates by association [...] Selection by association, rather than indexing, may yet be mechanized. (Texte complet dans <http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/4/>)

L'idée est de procéder par association entre les entités d'information plutôt que par indexation, conformément au fonctionnement naturel de l'esprit humain.

Le Memex est « *un dispositif dans lequel un individu stocke tous ses livres, ses enregistrements ou ses communications, et qui est mécanisée afin qu'il puisse être consulté avec une rapidité et flexibilité supérieure [...] il s'agit d'un supplément intime et agrandi de sa mémoire* » (Bush, 1945, p. 121). Le matériel peut être projeté sur les deux écrans translucides dont le meuble memex, une espèce de gros bureau, est équipé sur le dessus. Il permet l'interaction à l'aide des boutons, des leviers et d'un clavier. L'utilisateur peut lui-même ajouter ses annotations à travers une platine transparente. L'information est enregistrée sur des *microfilms* et, selon V. Bush, « si l'utilisateur insérait 5000 pages par jour il lui faudrait des centaines d'années pour remplir le répertoire »¹⁵⁸ (ibid.). Le contenu pourrait être acheté sous forme de microfilms déjà enregistrés et prêts à être utilisés par le Memex.

L'association des entités d'information grâce à des liens explicites permet la construction de *chemins de pensée* dans le Memex. Par sa capacité à relier les documents selon les associations de l'utilisateur et par la possibilité qu'offre le *Memex* à d'autres utilisateurs – et bien sûr à l'utilisateur originel – de rappeler à volonté ces documents, cette machine anticipe le développement de *l'hypertexte*. Ce type de liaison entre les entités d'information et l'accès rapide à l'information se retrouvent à la base du développement du Web. Plusieurs concepts et idées proposés dans ce projet se retrouvent dans les technologies actuelles¹⁵⁹. On retrouve conceptuellement *l'ordinateur personnel*¹⁶⁰ moderne, ainsi que le Wikipédia (les encyclopédies enregistrées dans les microfilms et la navigation dans leur contenu). Basé sur une structure dynamique des annotations d'utilisateurs, le système hypertexte *Porphyry 2001* (Bénel *et al.*, 2002) visant à faciliter la création, la récupération et le partage des documents et

¹⁵⁸ Le texte de V. Bush : "Yet if the user inserted 5000 pages of material a day it would take him hundreds of years to fill the repository, so he can be profligate and enter material freely. Most of the memex contents are purchased on microfilm ready for insertion."

¹⁵⁹ L'existence du bouton *Home*, présent aujourd'hui dans la plupart des téléphones intelligents portatifs ou dans les logiciels (e.g. navigateurs web, pages web), est spécifiée dans l'article présentant le Memex : « Un *bouton spécial* le transfère immédiatement à la première page de l'index » (Bush, 1945, p. 121).

¹⁶⁰ V. Bush nous invite à considérer « un dispositif futur pour une utilisation individuelle, une sorte de bibliothèque et fichiers privés mécanisés » (ibid.)

d'annotations. Comme dans le « memex » de V. Bush, ce système vise à offrir à l'utilisateur des possibilités de création et de récupération des « chemins de pensée ».

3.1.3. Xanadu : implémentation de l'hypertexte

Inventé par T. Nelson en 1965, le terme d'*hypertexte*¹⁶¹ connaît le succès dans le développement et l'expansion du Web. La définition courte proposée par le W3C¹⁶² affirme que l'*hypertexte* est « un texte qui n'est pas contraint d'être linéaire », « un texte qui contient des *liens* vers d'autres textes » (W3.org, s. d.). D'après (Clément, 1995) l'*hypertexte*¹⁶³ est un « ensemble complexe dans lequel chaque élément, chaque "texte" est relié aux autres de façon implicite ou explicite ». L'*hypertexte* est un « système de renvois permettant de passer directement d'une partie d'un document à une autre, ou d'un document à d'autres documents choisis comme pertinents par l'auteur » (Commission générale de terminologie et de néologie, 1999), ou encore une « combinaison de *texte en langage naturel* avec les *capacités de l'ordinateur* de construire des ramifications interactives, ou d'afficher dynamiquement un texte non linéaire [...] qui ne peut pas être imprimé facilement sur une page conventionnelle » (T. Nelson cité par Conklin, 1987, p. 17).

XANADU¹⁶⁴, « projet original d'hypertexte » (« Project Xanadu® », s. d.), proposé par T. Nelson en 1960 se base sur « *les connexions latérales* entre les documents et les fichiers » (« Xanadu FAQ », s. d.). Dans une présentation de Xanadu, T. Nelson met en avant les opportunités du document électronique qui permet au texte « d'échapper à la prison du papier » (*Ted Nelson demonstrates Xanadu Space*, 2008).

Par rapport à celui utilisé dans le Web, l'*hypertexte* tel qu'il est développé dans Xanadu est *profond*. Il propose en effet les caractéristiques suivantes (qui ne se retrouvent pas dans le Web) : (1) des liens incassables ; (2) une simplification du droit d'auteur ; (3) une connexion au contenu d'origine, (4) des liens à deux sens ; (5) une comparaison côte

¹⁶¹ Le préfix *hyper* vient du grecque *huper* : au-dessus, au-delà.

¹⁶² W3C : World Wide Web Consortium <http://www.w3.org>.

¹⁶³ Pour les formats qui ne sont pas contraints d'être du *texte* (e.g. graphiques, vidéos, son), T. Nelson introduit le terme *hypermedia* (W3.org, s. d.)

¹⁶⁴ Xanadu est le nom du « lieu magique de la mémoire littéraire » décrit dans le poème *Kubla Khan* (publié en 1816) de Samuel Taylor Coleridge.

à côté des documents connectés ; (6) une gestion profonde de la version du document ; (7) une publication incrémentale (« THE XANADU MODEL », s. d.).

Formalisé en 17 règles, le projet Xanadu vise à respecter les exigences suivantes :

1. Chaque serveur Xanadu est unique et identifié d'une façon sécurisée.
2. Chaque serveur Xanadu peut être mis en service séparément ou en réseau.
3. Chaque utilisateur est unique et identifié d'une façon sécurisée.
4. Chaque utilisateur peut rechercher, récupérer, créer et stocker des documents.
5. Chaque document peut consister en un nombre quelconque de parts dont chaque partie peut être constituée de n'importe quel type de données.
6. Chaque document peut contenir des liens de tout type, voire des copies virtuelles (« transclusions ») d'un autre document accessible par son propriétaire.
7. Les liens sont visibles et peuvent être suivis depuis les deux extrémités.
8. La permission de créer un lien vers un document est explicitement garantie par l'acte de publication même.
9. Chaque document peut contenir un mécanisme de rétribution, à un degré quelconque de granularité, pour assurer le paiement de chaque portion à laquelle l'utilisateur a eu accès, en incluant les copies virtuelles (« transclusions ») de tout ou partie d'un document.
10. Chaque document est identifié, unique et sécurisé.
11. Chaque document peut avoir des règles d'accès sécurisé.
12. Chaque document peut rapidement être recherché, stocké et récupéré sans que l'utilisateur ne sache où il est physiquement situé.
13. Chaque document est automatiquement situé sur un moyen de stockage approprié en fonction de sa fréquence d'accès depuis n'importe quel point de consultation.
14. Chaque document est automatiquement stocké de façon redondante, pour maintenir la disponibilité même en cas de désastre.
15. Chaque fournisseur de service Xanadu peut facturer à sa discrétion ses utilisateurs pour le stockage, la récupération, et la publication de documents.
16. Chaque transaction est sécurisée et reste perceptible seulement par les parties l'effectuant.

17. Le protocole de communication client-serveur Xanadu est un standard librement publié. Le développement et l'intégration de tierces parties sont encouragés (« Xanadu FAQ », s. d.).

Actuellement, le Web est conforme aux règles 1 à 5, 11, 12 et 17. Le standard XHTML supporte également la règle numéro 6.

3.2. Le Web et la sémantique

Proposé par T. Berners-Lee en 1989 au sein du CERN, le système d'hypertexte distribué qui deviendra le Web (appelé *Mesh* à ce moment-là) intégrait *le sens des liens* entre les *entités représentées* dès sa proposition (v. Figure 3-3). En 1990 le projet prend le nom de *World Wide Web*. Constitué des *cercles* et des *flèches*, ou plus spécifiquement, de *nœuds* et de *liens*, le Web est le système qui offre « une liberté de représentation que le tableau¹⁶⁵, par exemple, n'offre pas » (Berners-Lee, 1989). Chaque nœud peut « représenter ou décrire une personne spécifique ou un objet » (e.g. modules logiciels, groupes de personnes, projets, concepts, documents). Les liens peuvent représenter les relations entre les nœuds : dépend de, fait partie de, réfère à, décrit, par exemple (ibid.).

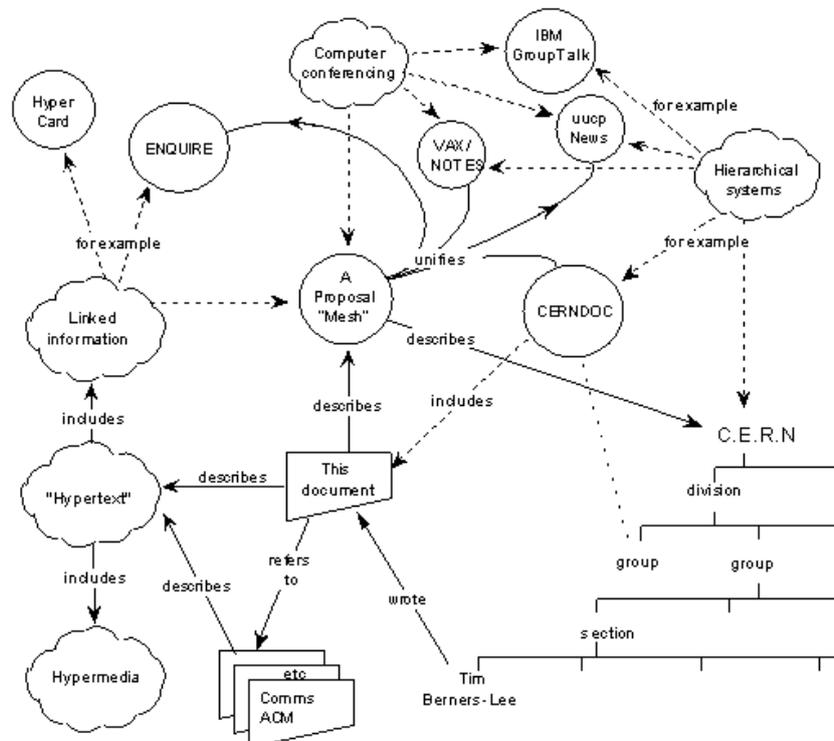


Figure 3-3. Le système d'hypertexte proposé par T. Berners-Lee (ibid.)

¹⁶⁵ Tableau dans le sens informatique (e.g. tableau de données)

Les nœuds et les liens peuvent donc porter du sens. Berners-Lee affirme qu'« *il est utile pour le système d'être au courant des types génériques des liens entre les éléments (dépendances, par exemple), et des types de nœuds (e.g. les personnes, les choses, les documents) sans imposer de limites* » (ibid.).

Opérant dans un « sous-domaine spécifique de l'hypertexte », le *World Wide Web* visait à : (1) fournir un protocole commun permettant à un humain de demander des informations lisibles stockées dans un système à distance, en utilisant des réseaux ; (2) fournir un protocole dans lequel l'information peut être automatiquement échangée dans un format commun au fournisseur et au consommateur ; (3) fournir une méthode de lecture au moins de type texte (si ce n'est graphique) en utilisant une grande partie des écrans d'ordinateur en service au CERN ; (4) fournir et maintenir au moins une collection de documents, dans laquelle les utilisateurs peuvent insérer leurs propres documents ; (5) fournir une option de recherche par mot clé, en plus de la navigation par références, en utilisant des index nouveaux ou existants ; le résultat d'une recherche par mot clé est tout simplement un document hypertexte constitué d'une liste de références à des nœuds qui correspondent aux mots-clés [...] ; (6) fournir les logiciels énumérés sous une licence libre de droits (Berners-Lee, 1990).

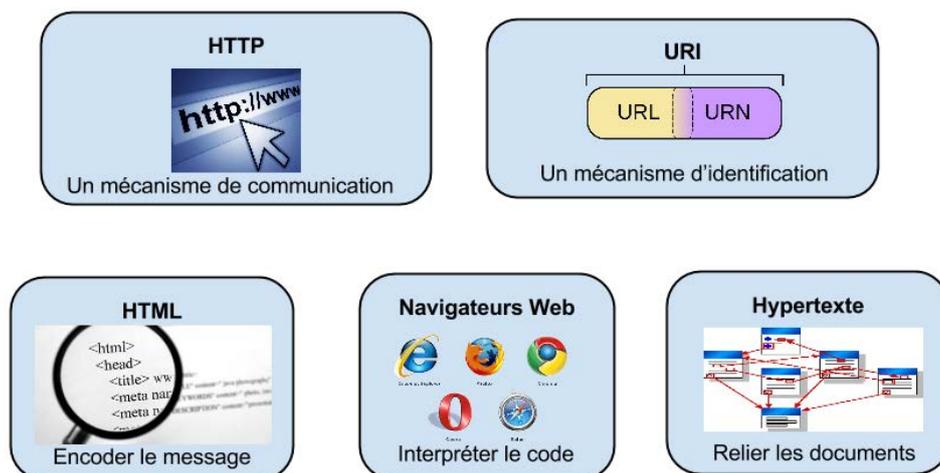


Figure 3-4. Les technologies du Web

Nous retrouvons dans les objectifs énumérés par T. Berners-Lee, les technologies et les protocoles utilisés aujourd'hui sur le Web (v. Figure 3-4) : le langage HTML (HyperText Markup Language) pour encoder les messages (document Web), les navigateurs Web

pour interpréter le code HTML, le protocole HTTP (HyperText Transfer Protocol) pour assurer l'échange des messages et les URI (Uniform Resource Identifier, v. page 147) pour permettre l'identification des documents Web.

3.2.1. Comment en-est on arrivé là ?

Dans le § 2.2.2 nous avons parlé du Web d'un point de vue technique. Considérons maintenant l'usage du Web relativement à son contenu. Le niveau d'adoption et d'évolution du Web est étroitement lié aux *pratiques* d'usages du Web. L'usage du Web a été influencé par la variété croissante des types de contenus véhiculés et des possibilités d'interaction avec ce contenu. Ainsi, les différents paradigmes d'usages du Web ont généré jusqu'à aujourd'hui trois scénarios différents qui correspondent aux trois générations du Web.

La première génération du Web (1991 - 1999), le Web 1.0, appelé encore « Web traditionnel », était un web statique. Les standards et les outils étant à leurs débuts, les documents Web ont été créés et stockés dans la forme sous laquelle ils devaient être « servis ». De plus, le niveau élevé de connaissances nécessaire à la création des documents Web a eu un impact direct sur le nombre des créateurs de contenu pour le Web. Le coût de création des documents web était alors élevé et l'intervention des utilisateurs faible. C'était un modèle centré sur la distribution de l'information, une approche *un-à-plusieurs*.

Avec l'apparition des solutions SGC (Systèmes de Gestion de Contenu) et les facilités apportées par ces systèmes dans la création des documents web, le Web devient plus accessible aux créateurs de contenu, ce qui conduit à une nouvelle génération en ce qui concerne son usage. Dans cette transition, un des aspects techniques qui ont facilité le développement du Web dynamique est l'objet JavaScript `XMLHttpRequest`¹⁶⁶. Cet objet offre une interface qui permet au code interprété par les navigateurs web de lancer des requêtes HTTP asynchrones sur les serveurs web afin de modifier le document Web sur le client (ou d'envoyer des données sur le serveur). Le document Web est donc

¹⁶⁶ (I)XMLHttpRequest créé par les développeurs de *Outlook Web Access* (de Microsoft) était présent dans Internet Explorer 5.0 (mars 1999), puis intégré dans la version 1.0 de *Gecko* (2002, *Gecko* est un moteur de rendu pour présenter les pages Web utilisé par Firefox Mozilla) et implémenté dans Safari 1.2 (février 2004)

dynamique. Pouvant interagir avec le serveur (e.g. modifier son contenu, modifier sa forme, mettre à jour le serveur) sans l'intervention explicite de l'utilisateur, le document Web se transforme en une application Web. C'est le début des *applications internet riches* (RIA)¹⁶⁷, qui offrent des expériences riches en interaction avec l'utilisateur.

Un autre facteur facilitateur est fourni par le langage XML¹⁶⁸ (Extensible Markup Language). Permettant « *de définir, de valider et de partager des formats de documents sur le Web* », ce langage de balisage contribue à « *l'interopérabilité du web* ». Les spécifications de *XML 1.0* deviennent une recommandation du W3C (World Wide Web Consortium) le 10 Février 1998¹⁶⁹. Avec d'autres technologies (e.g. XHTML¹⁷⁰, CSS¹⁷¹, DOM¹⁷², JavaScript, XSLT¹⁷³), JavaScript et XML forment AJAX (*Asynchronous JavaScript And XML*).

La dynamisation des pages web et les RIA ont préparé la migration vers le *Web 2.0* (2000-2009). Même si son nom suggère une nouvelle version du Web¹⁷⁴, il « *ne renvoie pas à une mise à jour de spécifications techniques quelconques mais à des changements dans la façon dont les développeurs de logiciels et les utilisateurs finaux utilisent le Web* »

¹⁶⁷ En anglais Rich Internet Application (RIA)

¹⁶⁸ Un des avantages du langage XML est que son contenu peut être lu relativement facilement par les humains.

¹⁶⁹ <http://www.w3.org/Press/1998/XML10-REC>

¹⁷⁰ XHTML : Extensible HyperText Markup Language : une « reformulation du HTML 4 en tant qu'application XML » (« XHTML 1.0: The Extensible HyperText Markup Language (Second Edition) », s. d.).

¹⁷¹ CSS : Cascading Style Sheets : « un mécanisme simple pour ajouter un style (e.g. police, couleurs) aux documents Web » (« Cascading Style Sheets », s. d.).

¹⁷² DOM : Document Object Model : « une interface indépendante de tout langage de programmation et de toute plateforme, permettant à des programmes informatiques et à des scripts d'accéder dynamiquement et de mettre à jour le contenu, la structure et le style des documents [web] » (« W3C Document Object Model », s. d.).

¹⁷³ XSLT : Extensible Style Sheets Transformation : « un langage permettant la transformation des documents XML en d'autres documents XML » (« XSL Transformations (XSLT) », s. d.).

¹⁷⁴ Le terme « Web 2.0 » est devenu populaire avec l'organisation de la première conférence *Web 2.0* en Octobre 2004 par O'Reilly Media. Un an plus tard, en octobre 2005, Tim O'Reilly donne une définition du Web 2.0 : « *le réseau en tant que plate-forme, couvrant tous les périphériques connectés ; les applications Web 2.0 sont celles qui contribuent le plus aux avantages intrinsèques de cette plateforme: la fourniture de logiciels en tant que services constamment mis à jour et qui s'améliorent avec l'augmentation du nombre de gens qui l'utilisent¹⁷⁴, consomment et remixent les données provenant de plusieurs sources, y compris des utilisateurs individuels, tout en fournissant leurs propres données et services sous une forme qui permet de le remixer par d'autres, créant des effets de réseau à travers une "architecture de participation", et allant au-delà de la métaphore de page du Web 1.0 pour offrir des expériences utilisateur riches* » (O'Reilly, 2005b).

(Floridi, 2009). Appelé *Web social*¹⁷⁵, il existe sous un modèle de partage de type *plusieurs-à-plusieurs*.

Le terme *Web social* est utilisé pour illustrer davantage l'aspect humain que l'aspect technique du Web. Il est utilisé pour décrire « *un sous-ensemble d'interactions Web qui sont hautement sociales, conversationnelles et participatives* » (Breslin, Passant, et Decker, 2009, p. 12). La participation et la création de contenus sont encouragées pour tous.

Les forums, les blogs¹⁷⁶, les wikis¹⁷⁷, les plateformes de partage de vidéos et ces « *réseaux numériques dits sociaux* »¹⁷⁸, accentués par leur taux de pénétration sur le Web, mettent l'utilisateur et le contenu au centre du Web. La création de contenu pour le Web (e.g. documents, textes, images, vidéos) n'est plus une tâche attribuée exclusivement aux experts. Le contenu désormais généré par tout type d'utilisateur¹⁷⁹ démocratise le Web et a des influences majeures sur la vie et la société¹⁸⁰. « *L'avis du consommateur est sollicité en permanence et il prend goût à cette socialisation virtuelle. Toutefois, la prolifération de contenus de qualité inégale engendre une infobésité difficile à contrôler* » (Waterschoot, 2012).

Le Web social a déclenché une explosion de la quantité d'information et les chiffres montrent une croissance soutenue du contenu généré par l'utilisateur. C'est l'exercice du socioconstructivisme à l'échelle globale. Les dispositifs mobiles élargissent encore la présence du Web et les moyens de création de contenu (e.g. photos, vidéos). Les statistiques montrent que d'ici 2015, le nombre de dispositifs mobiles qui accèdent à

¹⁷⁵ « *Le Web social est un ensemble de relations qui lient les personnes sur le Web.* » (« A Standards-based, Open and Privacy-aware Social Web », s. d.)

¹⁷⁶ Le terme de *blog* (ou *weblog*) fut proposé par Jorn Barger en 1997 (Aghaei, 2012; Murugesan, 2007).

¹⁷⁷ Wiki, signifie « vite » en hawaïen. Ward Cunningham crée le *WikiWikiWeb* en 1994.

¹⁷⁸ Intitulé « *Ces réseaux numériques dits sociaux* », le numéro 59 de la revue *Hermès*, traite les aspects sociaux apportés par ces réseaux : « *formes de lien social et de communication, la visibilité de la vie quotidienne, l'effacement des frontières publiques et privées [...]* » (Stenger et Coutant, 2011).

¹⁷⁹ User Generated Content (UGC).

¹⁸⁰ Un exemple est le rôle du Web 2.0 dans *le printemps arabe* (*Cases on Web 2.0 in developing countries*, 2012).

l'Internet va dépasser le nombre d'ordinateurs personnels (Weide, 2010). L'adoption de dispositifs mobiles ouvre le chemin vers un apprentissage pervasif¹⁸¹ ou ubiquitaire.

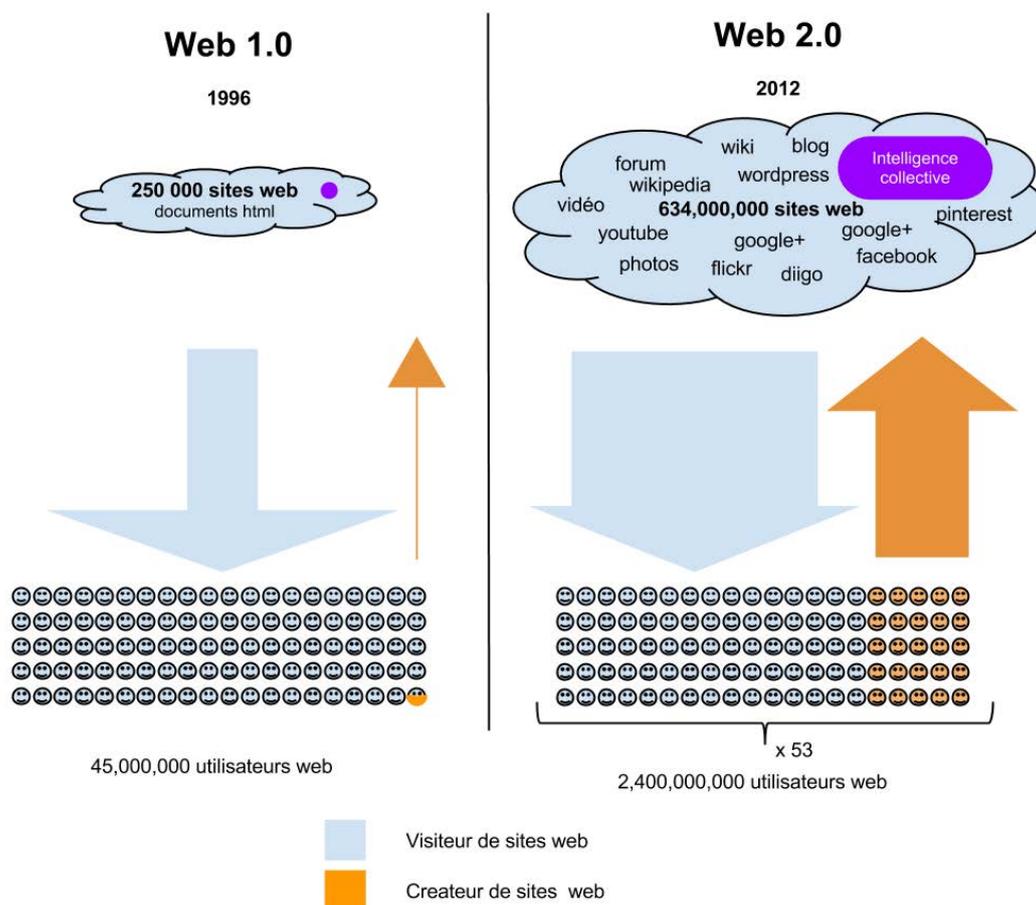


Figure 3-5. La participation¹⁸² dans le Web 1.0 et le Web 2.0, adaptée selon (Halpin, 2006)

La Figure 3-5 présente approximativement le rapport entre les consommateurs et les créateurs des sites web, dans les contextes du Web 1.0 et du Web 2.0. En réalité la quantité de *contenu* généré n'est pas strictement liée au *nombre de sites web*. Dans un cas simple, le nombre des blogs web (e.g. SGC WordPress installés) peut fournir un indice sur la quantité des articles créés sur le web. Les chiffres montrent que les principales plateformes de partage¹⁸³ (e.g. YouTube, DailyMotion, Flickr, Pinterest) et les

¹⁸¹ Dans l'informatique : diffusion à travers toutes les parties du système d'information.

¹⁸² Sources : <http://www.internetworldstats.com/stats.htm> et <http://news.netcraft.com/archives/2012/12/04/december-2012-web-server-survey.html>

¹⁸³ Une plate-forme de partage est un site d'hébergement de fichiers permettant qu'ils soient vus, lus, entendus ou encore téléchargés par des internautes. Il s'agit aussi d'un site web social permettant de sauvegarder et de partager ses références (signets web, articles en ligne, bibliographies) ainsi que de les classer selon le principe de folksonomie, par des mots-clés libres (tags).

réseaux socio-numériques (e.g. Facebook, Google+, Twitter) absorbent la plus grande partie du contenu généré par les utilisateurs.

Un exemple de la puissance du Web 2.0 et de la co-construction du savoir est *Wikipédia*. En effet, la participation au cœur du Web 2.0 a fait naître la plus grande encyclopédie gratuite. Le principe de la théorie des systèmes—le tout est plus grand que la somme des parties— est illustrée par le Web 2.0. L'agrégation de l'information créée par un grand nombre d'utilisateurs conduit à des contextes plus favorables au développement de la sagesse. James Surowiecki affirme dans son livre *La sagesse des foules* (Surowiecki, 2004) que les décisions résultant de l'agrégation de l'information dans des groupes sont souvent meilleures que celles d'individus isolés du groupe.

Dans le Web 2.0 les systèmes de classification se sont transformés. Si dans le Web traditionnel on parle de *taxonomies*¹⁸⁴, c'est-à-dire de classifications des entités faites dans des catégories définies en *amont* (Chekuri *et al.*, 1997), dans le Web 2.0 cette classification est *collaborative*, *décentralisée* et faite *après* l'enregistrement des entités. Ce type de classification est appelé *folksonomie*¹⁸⁵. Contrairement aux *taxonomies*, les folksonomies encouragent la génération des classifications en fonction du contenu et après la publication du contenu. La folksonomie n'est pas contrainte¹⁸⁶ à utiliser un vocabulaire ou une terminologie spécifique, déclarés en amont, l'élément central étant *l'étiquette* ou *le tag*. Le *tag* est un « mot-clé librement déterminé » (Sinclair et Cardew-Hall, 2008, p. 15) associé à une ressource. Bergson affirme que les mots sont comme des « *étiquettes* » que nous mettons sur les choses (Bergson, 1900, p. 67). La liste de mots clés émergents résultant de la participation des utilisateurs est communément appelée *folksonomie*. L'objectif est d'améliorer la recherche et l'identification des ressources pour référence ultérieure. L'organisation des tags selon leur nombre d'occurrences en relation avec une ressource est appelée « *tag cloud* » (Sinclair et Cardew-Hall, 2008, p.

¹⁸⁴ Voir les annuaires thématiques fondés sur la classification arborescente et proposés par les premiers moteurs de recherche (e.g. Yahoo Directory). Dans l'article « *Ontology is Overrated: Categories, Links, and Tags* » (http://shirky.com/writings/ontology_overrated.html) Clay Shirky critique les catégorisations et les classifications de type Dewey ou Yahoo qu'il qualifie « d'ontologique ». Mais nous ne pouvons pas mettre sur le même plan les classifications et les ontologies informatiques.

¹⁸⁵ Folksonomie, de l'anglais « *folksonomy* », combinaison entre les mots *folk* – *peuple*, *gens* et « *taxonomy* » : *taxonomie*.

¹⁸⁶ Renvoie à la liberté de l'XML pour la déclaration des nœuds.

15). Un exemple de « *tag cloud* » est présenté dans la Figure 3-6, la taille du tag donnant son importance (e.g. nombre d'association à la ressource).



Figure 3-6. Tag Cloud: les mots clés¹⁸⁷ du Web 2.0 (source : <http://www.journaldunet.com/diaporama/0610-dicoweb2/>)

La liberté de définition des étiquettes et le peu de contraintes dans leur utilisation peuvent être vus comme un avantage (e.g. liberté de choix), mais peuvent aussi se transformer en inconvénient (e.g. interprétation erronée). Le défaut principal des folksonomies consiste dans le fait que, d'une part, il n'existe pas de vocabulaire commun des étiquettes et que, d'autre part, les étiquettes peuvent avoir plusieurs sens. Les *spécifications explicites et formelles* (e.g. ontologies, schémas) déclarées dans le cadre du Web sémantique, se basent sur des taxonomies de concepts et termes utilisés dans le domaine de discours. L'acceptation de ces spécifications dans une communauté d'intérêt valide la structure de la taxonomie.

¹⁸⁷ Une phrase (en français) qui emploie les mots clés du Web 2.0 : "Nous allons crowdsourcer notre réseau social en proposant un mashup qui permettra aux blogueurs syndiqués de taguer nos podcasts afin de tirer la quintessence de notre long tail. Une API ajax sera également proposée" (<http://www.journaldunet.com/diaporama/0610-dicoweb2/>).

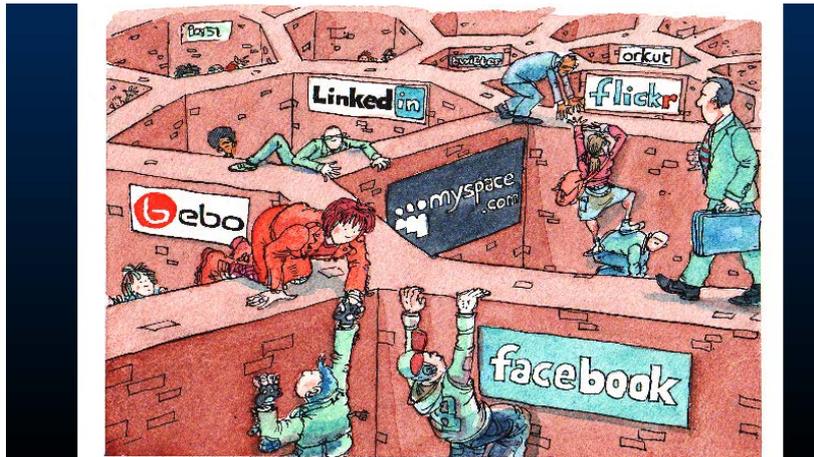


Figure 3-7. La disjonction des réseaux sociaux, source : (Berners-Lee, 2009, p. 22)

Les plateformes de partage se sont formées en fonction du type des données partagées (e.g. vidéo – Youtube, DailyMotion ; photos – Flickr ; blogs – WordPress) ou des types d'intérêts fondant les communautés (e.g. apprentissage de langues - Babbel.com). Comme le montre la Figure 3-7, la principale limite de ces plateformes est leur cloisonnement par rapport aux autres plateformes (Berners-Lee, 2009). Cette disjonction est due au manque d'implémentation des standards pour les *données liées*. Des informations complémentaires sur les *données liées* sont présentées dans le paragraphe 3.2.3.

Toutefois, les données ne sont pas générées que par les utilisateurs, mais aussi par différents agents informatiques (e.g. ordinateurs, machines, logiciels, capteurs, sondes). Elles sont de plus en plus riches et variées en contenu, structure et forme. C'est ce qu'on appelle le *Big Data*. IBM constate que « *les discussions concernant les volumes des données changent de téraoctets en pétaoctets, avec un passage inévitable à zettaoctets*¹⁸⁸ » (Zikopoulos *et al.*, 2012, p. 7). Selon les estimations d'IBM le *Big Data* s'enrichit¹⁸⁹ chaque jour avec 2,5 exaoctets de données, quatre-vingt pourcents ayant été créés dans les deux dernières années (« IBM What is big data? », 2013). Si ce rythme se maintient, la prédiction pour le volume de données stocké d'ici à 2020 est de 35 zettaoctets (Zikopoulos *et al.*, 2012, p. 5).

¹⁸⁸ 1 Pétaoctet (Po) = 10^{15} octets ; 1 exaoctet (Eo) = 10^{18} octets ; 1 zettaoctet (Zo) = 10^{21} octets.

¹⁸⁹ « *Ces données proviennent de partout: des capteurs utilisés pour recueillir des informations sur le climat, les messages vers des sites de médias sociaux, photos et vidéos numériques, enregistrements de transactions d'achat, et les signaux GPS des téléphones cellulaires pour n'en nommer que quelques-uns.* » (« IBM What is big data? », 2013).

IBM caractérise le Big Data par les trois *V*, ou le *V*³ : volume, vitesse et variété, voir la Figure 3-8. Le volume représente la quantité des données ; la variété représente la diversité des données (« *tout type de données* »). Les systèmes d'analyse doivent s'adapter pour traiter aussi des données non-structurées provenant des différentes sources. La vitesse, quant à elle, est la vitesse à laquelle les données circulent. Cette vitesse du flux des données, en combinaison avec la durée de vie courte de données nécessite leur traitement en temps réel.

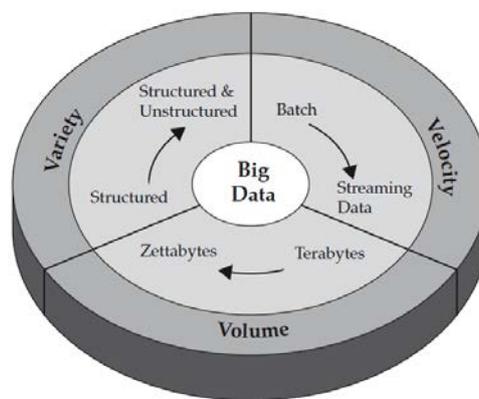


Figure 3-8. Caractérisation de la Big Data selon IBM (ibid.)

Dans un article récent sur la mesure de la valeur commerciale de *Big Data*, aux trois *V* précédents s'ajoute un quatrième : la *véracité* (Kobielus, 2013). L'authenticité, l'exactitude et la vérité de la donnée ajoute un cinquième *V* : *valeur*. Le Big Data crée des nouveaux moyens pour générer de la *valeur* et la capacité d'utilisation de ces données est le facteur clé qui détermine le succès d'une entreprise dans notre siècle (*In the 21st Century, Data Is King*, 2013). Le « nouveau "intel inside" » (O'Reilly, 2005a) sont les données. Il ne s'agit plus d'applications, il s'agit de données : comment gérer ces données (vu la quantité générée), comment exploiter les données (e.g. extraire de l'information utile, efficace), comment les intégrer dans différentes applications (e.g. utilisation sur différents applications et dispositifs), comment les stocker (e.g. cloud), comment y accéder (e.g. accès mobile) ?

Les trois *V* – volume, vitesse et variété – poussent les algorithmes de recherche vers leurs limites. Dans cette perspective, les métadonnées et les technologies du Web sémantique, par l'intégration du sens dans les documents Web (e.g. RDFa, Microdata),

The Web	The Social Web	The Social Semantic Web
Personal Websites	Blogs	Semantic Blogs: semiBlog, Haystack, Structured Blogging, Zemanta
Content Management Systems, Britannica Online	Wikis, Wikipedia	Semantic Wikis: Semantic MediaWiki, SemperWiki, Platypus, DBpedia, Rhizome
AltaVista, Google	Google Personalised, Searchles	Semantic Search: SWSE, Swoogle, Intellidimension, Powerset, Hakia
CiteSeer, Project Gutenberg	Google Scholar, Book Search	Semantic Digital Libraries: JeromeDL, BRICKS, Longwell
Message Boards	Community Portals	Semantic Forums and Community Portals: SIOC, OpenLink Data Spaces, Talis Engage
Buddy Lists, Address Books	Online Social Networks	Semantic Social Networks: FOAF, PeopleAggregator, Social Graph API

Figure 3-10. L'évolution des plateformes dans le cadre de l'évolution du Web

Les limites du Web 2.0 poussent le Web vers une nouvelle génération, dite Web 3.0. Nommée aussi *Web sémantique*, cette nouvelle génération vise à organiser la masse d'informations en ajoutant du sens aux liens et aux données représentés. Son but est de rendre le Web plus facile à exploiter, à la fois par les hommes et les ordinateurs. Ne concernant pas les mêmes problèmes, le Web sémantique ne remplace pas le Web social, mais il l'augmente. Les plateformes et les systèmes reposant sur le Web tirent avantage des possibilités offertes par la génération du Web dont ils sont issus, voir la Figure 3-10.

Présentant l'évolution du Web (v. Figure 3-11), N. Spivack considère la *sémantique des connexions sociales* sur l'axe horizontal, et la *sémantique des connexions de l'information* sur l'axe vertical. Dans sa vision, le passage vers le Web 2.0 a mis l'accent sur la participation de l'utilisateur et sur une *mise à niveau de l'interface et de l'expérience de l'utilisateur*. L'innovation apportée par le Web 3.0 consiste dans la *mise à niveau du « back-end »*¹⁹⁰ du Web (Spivack, 2007).

¹⁹⁰ Selon N. Spivack, le Web 2.0 concernait la mise à jour du « front-end » du Web (e.g. interface utilisateur, expérience utilisateur). Dans ce contexte, le « back-end » du Web se réfère aux données et à la logique avec laquelle on exploite ces données.

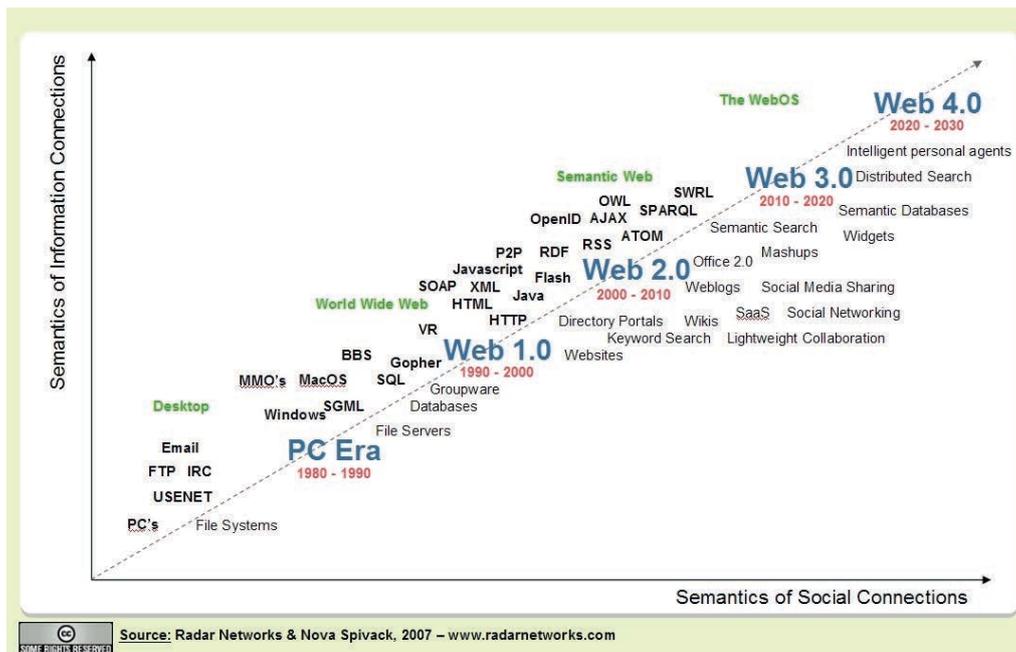


Figure 3-11. La sémantique dans l'évolution du Web (source: ibid.)

La sémantique des connexions, qu'elle soit du côté des utilisateurs ou du côté de l'information est un facteur déterminant dans l'évolution du Web. Dans la partie suivante, nous abordons la question de la sémantique dans le Web.

3.2.2. Sens et signification dans le Web

Véritable support et medium de communication, le Web fait l'objet de démarches d'amélioration continue des procédures de communication et d'échange d'informations. Dans une communication réussie, l'information doit être correctement transmise (e.g. syntaxe, règles orthographique) et son sens doit être correctement interprété (e.g. comprise).

Il s'agit donc de développer et de fournir des outils et langages capables d'enregistrer le sens et la signification de l'information, afin de permettre une interprétation et une exploitation correctes. C'est ici qu'intervient le domaine de *la sémantique*¹⁹¹ qui étudie la construction du sens et de la signification des signes du langage.

¹⁹¹ Le mot sémantique, dérivé du grec *sēmantikos*, désigne « l'étude du langage considéré du point de vue du sens ; théorie visant à rendre compte des phénomènes signifiants dans le langage » (Le Robert et al., 2010, p. 2345). Elle veut comprendre comment le sens et la signification des concepts complexes peuvent être obtenues à partir des concepts simples.

Le signe¹⁹² « est signe parce qu'il signifie », parce qu'il renvoie à un sens à interpréter. Porteur de signification, il est « quelque chose qui tient lieu pour quelqu'un de quelque chose sous quelque rapport ou à quelque titre » (Peirce, 1978, p. 121). Dans les années 1890, C.S. Peirce¹⁹³ analyse le signe en tant que composé d'une triade : *representamen* (signe matériel), *objet* (objet de pensée) et *interprétant* (représentation mentale de la relation entre les deux précédents, fondement de la signification) (« 76 Définitions du signe relevées dans les écrits de C.S. Peirce », s. d.).

Le terme *sémantique* fut inventé par le linguiste français Michel Bréal en 1878 dans une lettre adressé à Angelo de Gubernatis¹⁹⁴ (Bréal, 1991). Il introduit officiellement le terme dans un article de 1883 intitulé « Les lois intellectuelles du langage. Fragment de sémantique » : « En effet, c'est sur le corps et sur la forme des mots que la plupart des linguistes ont exercé leur sagacité : les lois qui président à la transformation des sens, au choix d'expressions nouvelles, à la naissance et à la mort des locutions, ont été laissées dans l'ombre ou n'ont été indiquées qu'en passant. Comme cette étude, aussi bien que la phonétique et la morphologie, mérite d'avoir son nom, nous l'appellerons la *SÉMANTIQUE* (du verbe *sêmainô*¹⁹⁵), c'est-à-dire la science des significations » (Bréal, 1883).

Parlant de l'importance du sens, Bréal affirme que « ce qui manque le plus à ces études [linguistiques] jusqu'à présent, c'est le côté sémantique : il semblerait, à lire ces travaux, que les questions d'accentuation, de voyelle de liaison, d'ordre des termes fussent tout. Je crains qu'on n'ait oublié l'essentiel, à savoir le sens, car c'est le sens, et non autre chose qui fait le composé et qui, en dernière analyse décide de la forme » (Bréal, 1897, p. 173-174).

¹⁹² Le signe : chose mise pour une autre, « *indice, élément qui permet de percevoir quelque chose* ».

¹⁹³ C.S. Peirce (1839 – 1914) est reconnu comme le fondateur de la *sémiotique*, ou science des systèmes de signes de communication.

¹⁹⁴ P. Ciureanu, « Lettere inedite di Michel Bréal, Gaston Paris e Émile Littré », *Convivium*, Juillet-Août 1955, 452-66.

¹⁹⁵ En grec *sêmainein* : montrer, montrer par un signe, indiquer, signaler ; *sêmainô* signifie *je montre*.

Le fondateur du mouvement *sémantique générale*, A. Korzybski¹⁹⁶, affirmait que « *la carte n'est pas le territoire* » et que « *le mot n'est pas la chose* » (Korzybski, 1958). L'identification, la distinction et les relations des éléments présents dans l'expression des mots (signes) sont mises en évidence dans le *triangle sémiotique* proposé par C. Ogden et I. Richards (Ogden et Richards, 1923). *Le triangle sémiotique* représente les relations entre *la chose*, *le signe* et *le concept* (v. Figure 3-12).

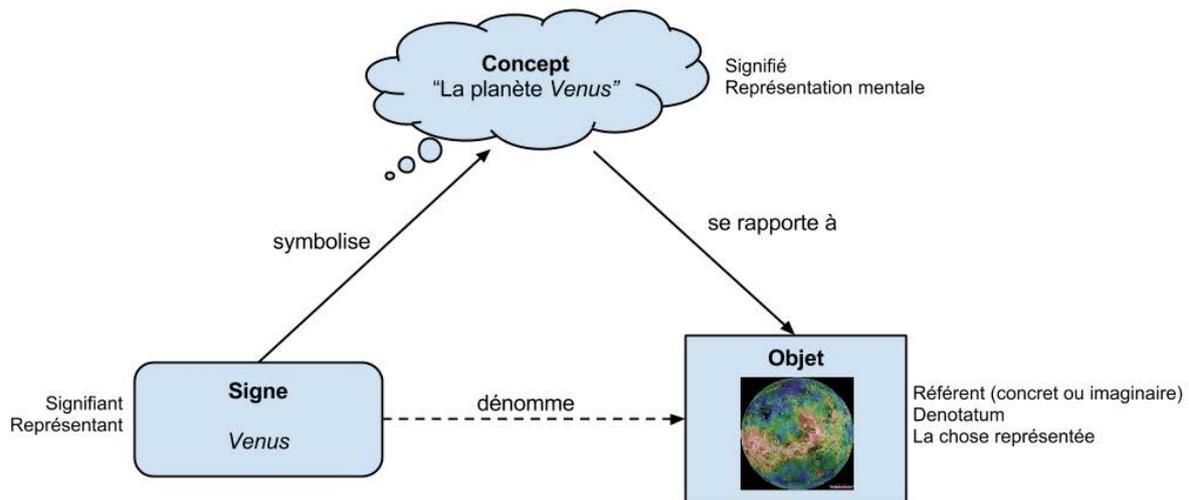


Figure 3-12. Le triangle sémiotique (selon : ibid.)

Se référant à Aristote (e.g. parole <-> concept/état d'âme <-> chose) il propose trois notions principales :

1. *la chose, l'objet* : c'est-à-dire l'objet du monde, aussi qualifié de *réfèrent* ;
2. *le signe* : aussi qualifié de forme ou de *signifiant*, c'est-à-dire le terme, la chaîne de caractères ou la photo qui symbolise cet objet particulier du monde ;
3. *le concept* : encore appelé *signifié*, correspond à l'idée que dénote le signe.

Dans sa sémiotique, C.W. Morris (1938) distingue trois dimensions fondamentales : (1) la sémantique, ou relation entre le signe et ce qu'il signifie ; (2) la syntaxique¹⁹⁷, ou

¹⁹⁶ Alfred Korzybski (1879 - 1950), fondateur de la « *Sémantique générale* », tenait pour un fondement essentiel que la connaissance humaine du monde est limitée à la fois par le système nerveux de l'homme et par la structure du langage.

¹⁹⁷ Syntaxe (du grecque : « arrangement », « ordre »). « *Étude des relations entre les formes élémentaires du discours (mot, syntagme) Étude des règles qui président à l'ordre mots et à la construction des phrases, dans une langue ; ces règles* » (Le Robert et al., 2010, p. 2488). Dans l'informatique, la syntaxe définit la structure normative des données.

relation des signes entre eux et (3) la pragmatique¹⁹⁸, ou relation entre le signe et les utilisateurs des signes. L'interprétation et le sens de l'information dépendent d'un part du contexte¹⁹⁹ de l'émetteur et du récepteur (e.g. connaissance du monde) et d'autre part de la pragmatique de l'émetteur (e.g. l'intention de l'émetteur).

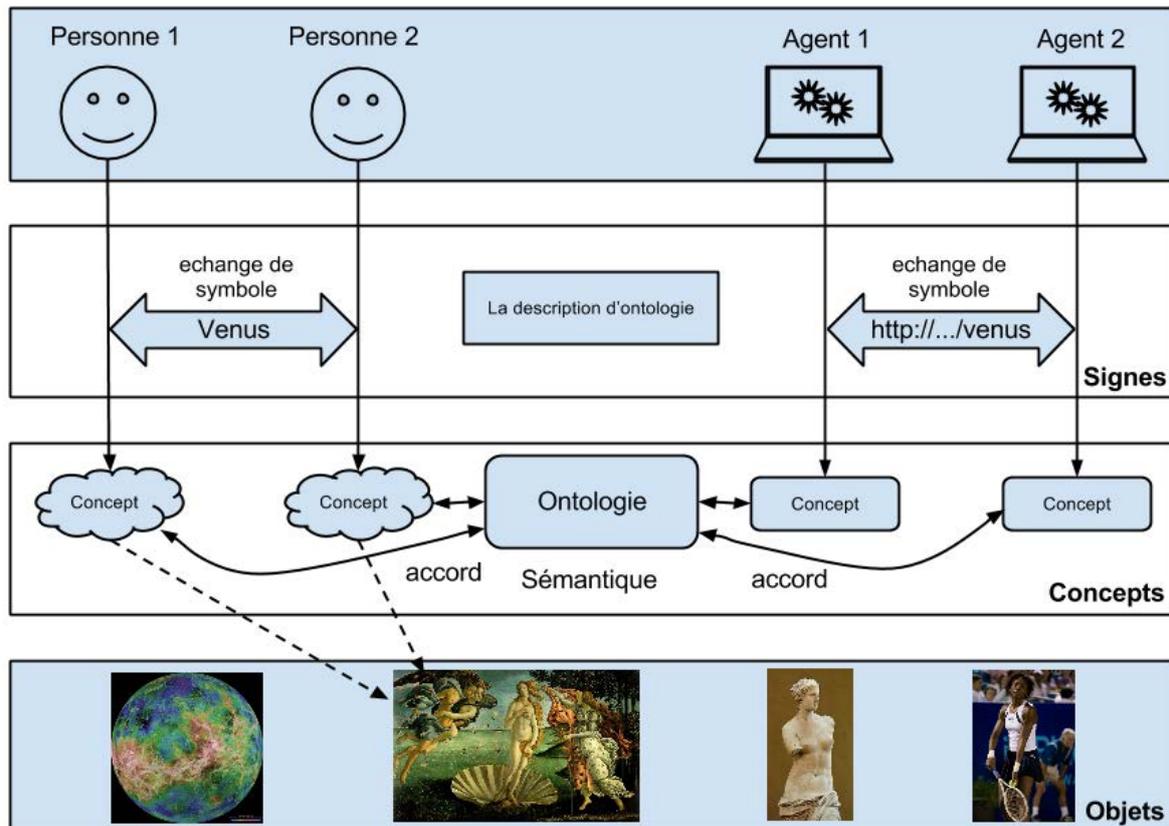


Figure 3-13. La transmission de la sémantique dans la communication (humaine et technique)

Si dans la communication humaine²⁰⁰ la sémantique est implicite, dans l'échange d'information entre les ordinateurs, la sémantique doit être explicitée (v. Figure 3-13). Le Web sémantique s'intéresse donc aux modalités d'explicitation et de représentation

¹⁹⁸ Pragmatique (du grecque *pragmatikos* : « relatif à l'action ») « *Études des signes en situation. L'énonciation fait partie de la pragmatique. Syntaxe, sémantique et pragmatique.* » (ibid., p. 1992) Elle reflète l'intention par laquelle la langue est utilisé pour communiquer un message. L'étude de l'application de la langue dans des situations différentes, le but visé par l'émetteur.

¹⁹⁹ Contexte (lat. *contextus* : « assemblage »), désigne l'entourant d'un symbole, ses relations avec l'environnement et d'autres éléments connexes. Il peut être générale (e.g. lieu, temps) ou personnel (e.g. relation entre l'émetteur et le récepteur du message).

²⁰⁰ L'accord sur les concepts ou la signification des signes échangés ne sont pas nécessairement transmis au moment de la communication. Le sens est dérivé, dans un contexte donné, d'une connaissance préalable, d'un accord antérieur, de règles et de spécifications antérieures partagées entre l'émetteur et le récepteur. La communication implique l'acceptation d'un langage commun, « *les gens ne peuvent pas partager leurs connaissances s'ils ne parlent pas une langue commune* » (Davenport et Prusak, 1998, p. 98).

de la sémantique dans l'échange d'informations sur le Web. Selon W3C, « le Web sémantique fournit un modèle qui permet aux données d'être partagées et réutilisées entre plusieurs applications, entreprises et groupes d'utilisateurs » (source : Semantic Web Activity, <http://www.w3.org/2001/sw/>). À un niveau plus pragmatique, il consiste à développer un cadre de référence (e.g. formats, langages, technologies) permettant d'enregistrer à la fois l'information et sa sémantique (e.g. signification, contexte) dans une forme adaptée aux ordinateurs. Il vise également à supprimer le bruit sémantique lors d'un échange d'information.

Dans la communication (e.g. écrite, orale), les signes doivent respecter les règles syntaxiques. La sémantique n'impose pas le même type de contraintes dans l'association du sens. Un signe peut porter en même temps plusieurs « contenus », « sens » ou « sèmes »²⁰¹. Tout énoncé peut avoir plusieurs interprétations. Comme l'affirme S. Ullmann, « la polysémie²⁰² nous permet d'exploiter rationnellement le potentiel des mots [...] Le prix de cette rationalisation est le risque d'ambiguïté » (Le Robert et al., 2010, p. 1958). De plus, « la polysémie affecte plus particulièrement les signes de la langue qui sont les plus fréquents dans le discours [...] » (Pergnier, 1997, p. 42). Pour répondre à cette source de « malentendus et de conflits », Ogden s'assigne la tâche d'imaginer un nouveau langage où « chaque mot renverrait à une notion précise et nette, sans ambiguïté » (Ogden et Richards, 1923; Poibeau, 2008, p. 31). Le résultat est le *Basic English* (Ogden, 1930), un vocabulaire de 850²⁰³ mots anglais qui devrait suffire à « une communication simple et non ambiguë²⁰⁴ ».

²⁰¹ Sème : (du grecque *sêmeion* « signe »). « Unité minimale différentielle de signification », « trait distinctif » (Le Robert et al., 2010, p. 2346). En sémiotique : « Unité constituant un acte de communication minimal » (Alain Rey, *Dictionnaire culturel en langue française*, Tome IV, p. 686)

²⁰² Polysémie : (*poly* + grec *semaînem* (sémantique)). « Caractère d'un signe qui possède plusieurs contenus, plusieurs sens » (ibid., p. 1958)

²⁰³ 100 mots pour des opérations (e.g. come, get, give, go) ; 400 mots généraux pour des choses (e.g. account, act, addition, adjustment) ; 200 mots *images* (e.g. angle, ant, apple) ; 100 mots pour des qualités (e.g. able, acid, angry) et 50 mots pour les contraires (e.g. awake, bad, small).

²⁰⁴ « L'idée de définir des mots compliqués au moyen de mots plus simples est évidente et naturelle. Il ne s'ensuit évidemment pas que l'on puisse trouver un sous-ensemble de sens de base au moyen duquel tout le reste s'exprime » (Poibeau, 2008, p. 32).

Le plus éloquent exemple d'ambiguïté dans l'échange d'information sous une forme de chaîne de caractères est le cas des *homographes*²⁰⁵. Dans ce cas, le sens est déduit du contexte en utilisant des méthodes statistiques. Dans le cas des systèmes de reconnaissance vocale, où l'échange d'information se transmet sous une forme sonore, d'autres types d'ambiguïtés interviennent : *l'homonymie*²⁰⁶ , *l'homophonie*, *la paronymie*²⁰⁷. L'ordinateur doit donc « résoudre de nombreuses ambiguïtés dont l'individu n'a même pas conscience » (Poibeau, 2008, p. 32).

Pour résoudre le problème d'ambiguïté et permettre l'interprétation du contenu par l'ordinateur, le Web sémantique associe des identifiants uniques et globaux (e.g. URI) aux *entités nommées*. Le terme d'*entité nommée*, largement utilisé dans le domaine du *traitement automatique du langage naturel*, fut inventé pour la sixième *Message Understanding Conference* (MUC-6) (Nadeau et Sekine, 2007; Grishman et Sundheim, 1996, p. 467). Au cœur « de la problématique de l'extraction de l'information d'un document », les *entités nommées* représentent les *éléments atomiques de sens* du document (Poibeau, 2008, p. 34). Utilisés comme étiquettes (tags) dans un langage de balisage, l'*ENAMEX* (*entity name expression*), ces éléments désignent « les noms de personnes et d'organisations », alors que le *NUMEX* (*numeric expression*) est utilisé pour « les unités monétaires ou pourcentages » (Grishman et Sundheim, 1996, p. 467). Le Web sémantique permet la spécification du sens des relations entre les entités nommées identifiées au sein d'un document.

3.2.3. Le Web sémantique

Conçu comme un espace d'information, le Web devrait être utile « non seulement pour la communication interhumaine, mais aussi pour que les machines soient en mesure de participer et d'aider » (Berners-Lee, 1998). Le fait que la plupart des données présentes sur le Web soient conçues pour la « consommation » humaine représente un obstacle dans l'interprétation et le traitement de ces données par la machine. Pour répondre à ce

²⁰⁵ Homographes : mots qui ont la même orthographe mais un sens différent. Les homographes à prononciation différente sont des allophones.

²⁰⁶ Homonymes : de mots qui ont la même prononciation, une orthographe identique ou différente (homophones) et un sens différent. (ex. sang, cent, sans, sent)

²⁰⁷ Paronymes : mots de sens différent mais dont la graphie et/ou la prononciation sont fort proches. (ex. « Qui se ressemble s'assemble » .

problème, le Web sémantique propose des langages qui permettent d'exprimer le sens de l'information dans une forme traitable par la machine. En septembre 1998, T. Berners-Lee publie le schéma de la feuille de route du Web sémantique (v. Figure 3-14).

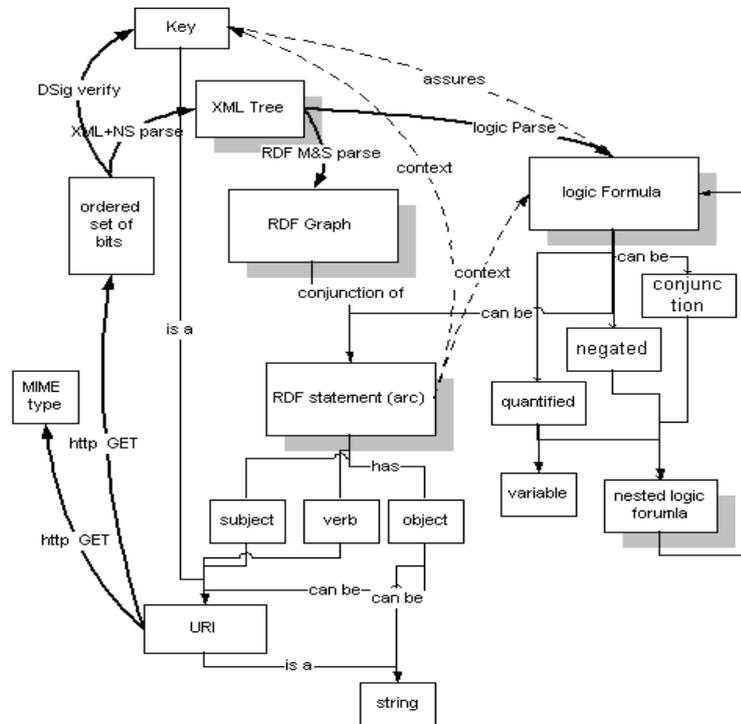


Figure 3-14. Schéma de la feuille de route du Web sémantique (source : ibid.)

Le schéma présente des éléments du Web (e.g. URI, HTTP), le format XML et, au centre du Web sémantique, *le cadre de description de ressources (RDF)*, ce qui permet d'exprimer l'information sous forme de graphe et d'ajouter des *formules logiques*. Ensuite, ces formules logiques permettent l'exploitation du sens de l'information par les machines (ibid.).

En 2001, Tim Berners-Lee et ses collègues publient dans le journal *Scientific American Magazine* l'article *The Semantic Web* : « *Le web sémantique est une extension du Web actuel dans laquelle l'information se voit associée à un sens bien défini améliorant la capacité des ordinateurs et des hommes à travailler en coopération* » (Berners-Lee, Hendler, et Lassila, 2001). Selon W3C, le Web sémantique « fournit un cadre commun qui permet aux données d'être partagées et réutilisées entres plusieurs applications, entreprises et communautés d'utilisateurs. » (« W3C Semantic Web Activity Homepage », s. d.).

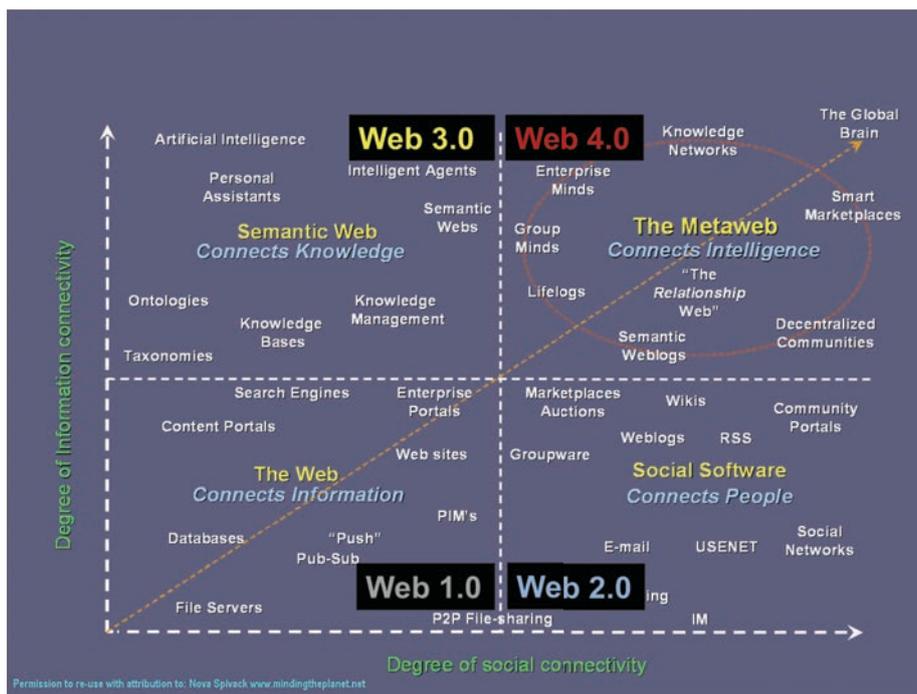


Figure 3-15. L'évolution du Web (source : Spivack, 2004)

Selon N. Spivack, le Web 1.0 *connecte l'information* et le Web 2.0 *connecte les personnes* (v. Figure 3-15). Le facteur provocateur d'évolution vers le Web 2.0 est l'augmentation des connexions des utilisateurs, phénomène que l'on peut constater dans le degré de leur participation, dans la création de contenus et dans les interactions. Le passage au Web 3.0 (ou sémantique) accroît et enrichit la qualité de la connexion de l'information, conduisant ainsi à la *connexion de connaissances*. Enfin, le *MétaWeb* ou *Web 4.0* permettra une connectivité au niveau de *l'intelligence* (ibid.).

Dans le Web sémantique, l'enrichissement des connexions des informations est réalisé à travers *les annotations sémantiques*. Le sens de ces annotations est défini et rendu explicite par des *représentations de connaissances formelles et standardisées*. Le contenu en langage naturel doit alors être *explicitement* annoté avec des *annotations sémantiques*. Encapsulant le sens et la signification du contenu, ces annotations permettent aux machines (e.g. ordinateurs, agents logiciels) d'interpréter correctement le contenu.

Trois thèmes principaux apparaissent dans le cadre du Web sémantique (Hitzler, Krötzsch, et Rudolph, 2009, p. 1-2) :

- la construction des modèles : décrire le monde dans des termes abstraits pour permettre la compréhension de la réalité dans sa complexité ;
- le calcul des connaissances : tenter de construire des machines à raisonner capables de tirer des conclusions éloquentes à partir de connaissances encodées ;
- l'échange d'information : transmettre des ressources complexes entre les ordinateurs, permettant de distribuer, de lier et construire la connaissance à une échelle globale.

Deux caractéristiques essentielles peuvent être relevées lors de l'utilisation des technologies du Web sémantique conduisant à deux types d'applications (v. ANNEXE R). Dans un premier temps les interconnexions sémantiques fournissent le cadre de création des *données ouvertes et liées*. Le Web n'est plus alors un Web de documents mais un *Web de données*, en quelque sorte une *base de données globale* constituée d'un *réseau sémantique* d'assertions. La seconde caractéristique du Web sémantique, rendue possible par les *descriptions formelles* (du monde, de l'environnement, d'un domaine spécifique) se rapproche du domaine de l'Intelligence Artificielle. *La sémantique formelle*, c'est-à-dire la dimension *logique* de la sémantique, permet d'exprimer les spécifications (e.g. schémas RDF, ontologies) dans le cadre du Web sémantique d'une manière précise, utilisant une formulation mathématico-logique (ibid., p. 73). Ainsi, les *représentations formelles de connaissances* permettent aux agents intelligents d'inférer et de raisonner sur les *connaissances* en se basant sur ces descriptions formelles.

L'existence et l'accès à l'information dans un format adapté aux agents logiciels ainsi que les possibilités d'inférence et de raisonnement sur les données ont conduit au développement des *agents intelligents*. Par *agent intelligent* nous nous référons aux logiciels capables d'inférer et de raisonner sur un ensemble de données selon des règles établies à l'avance. Dans le domaine de l'Intelligence Artificielle, un *agent intelligent* est « toute entité qui peut percevoir son environnement grâce à des capteurs et peut ensuite agir sur l'environnement par les effecteurs afin de réaliser ses objectifs » (Russell et Norvig, 2003, p. 31). Dans le contexte de la représentation de connaissances, *intelligent* se réfère à « la capacité de trouver les conséquences implicites de ses connaissances explicitement représentés » (Nardi et Brachman, 2003, p. 6). Les systèmes qui intègrent ce type d'intelligence s'appellent *systèmes à base de connaissances*.

Dans le Web traditionnel, l'identification des *documents web* se fait par le biais d'*URI (Uniform Resource Identifier)*. Or, le Web sémantique est un Web de *données*. L'association des URI aux données, qui décrivent toutes sortes d'*objets ou concepts* (e.g. personnes, lieux), conduit à la construction du *Web de données*. Les liaisons entre ces données permettent la création de ce qu'on appelle les *données liées (Linked Data - LD)*. Dans un article de 2006, T. Berners-Lee publie les quatre principes des *données liées* (Berners-Lee, 2006) :

1. utiliser les URI pour le nommage des ressources (e.g. personnes, lieux, concepts) ;
2. utiliser les URI accessibles via le HTTP pour permettre l'accessibilité à ces ressources ;
3. donner accès aux données utiles en utilisant les standards (e.g. RDF, SPARQL) ;
4. inclure des liens vers d'autres URI identifiant les mêmes ressources dans d'autres vocabulaires.

En 2010, il propose un système d'évaluation des données liées en cinq niveaux²⁰⁸ correspondant à un nombre d'étoiles attribuées par niveau. Pour rendre possible l'exploitation des données liées, elles doivent être aussi publiées sous une licence libre de droit. On parle dans ce cas là de *données ouverts et liées (Linked Open Data - LOD)*. Les technologies du Web sémantique simplifient l'intégration des données provenant de différentes sources et l'utilisation des standards rend possible la construction des applications intégrant des sources variées de données. La combinaison de ces données conduit à une représentation plus complète du domaine exploité.

Données liées est un « terme utilisé pour décrire une pratique recommandée pour exposer, partager et connecter les données, l'information et la connaissance dans le Web sémantique en utilisant les URI et le RDF » (Linkeddata.org, s. d.). Du point de vue technique, le terme de *Données liées* se réfère aux « données publiées sur le Web de telle

²⁰⁸ Les cinq niveaux sont les suivantes : (1) publication sur le Web (quel que soit leur format) avec une licence libre de droit, afin d'être données *ouvertes (open data)* ; (2) publication sous forme de données brutes structurées, traitables par l'ordinateur (e.g. les données tabulaires sont préférées en format *Excel* plutôt que sous forme d'une image scannée) ; (3) règles précédentes plus utilisation d'un format non-proprétaire (e.g. le format *csv*²⁰⁸ est préféré au format *xls*) ; (4) règles précédentes plus utilisation des standards proposés par W3C (RDF et SPARQL), afin que d'autres peuvent référencer les données ; et (5) règles précédentes plus construction des liaisons vers les données des autres personnes pour fournir un contexte à ces données.

manière qu'elles soient *exploitables par les ordinateurs* ; les *significations* sont *explicitement définies*, elles sont *liées* à d'autres ensembles de données externes et, à leur tour, elles peuvent être liées dans les ensembles de données externes » (Bizer, Heath, et Berners-Lee, 2009).

3.3. Représentation de l'information dans le Web sémantique

Le Web sémantique est « un ensemble de technologies et standards pour réaliser le Web de données » (I. Herman, 2008, p. 54). Nous présentons dans cette section les principales technologies utilisées pour la représentation de l'information dans le cadre du Web sémantique. Proposée à l'origine par T. Berners-Lee (Berners-Lee, 2000, p. 17), l'illustration de *l'architecture en couches (Layer Cake en anglais)* regroupe les technologies utilisées dans le cadre du Web sémantique (v. Figure 3-16) et recommandées par le Consortium W3.

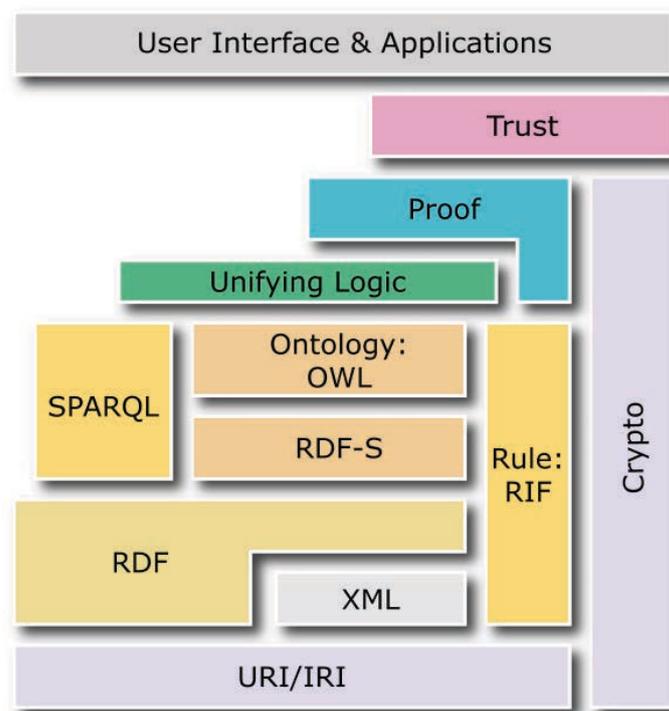


Figure 3-16. L'architecture en couche des standards du Web sémantique (source : Bratt, 2007, p. 24)

En tant qu'extension du Web actuel, le Web sémantique repose sur les URI²⁰⁹ (*Uniform Resource Identifier*), mécanisme d'adressage unique et global du Web. Ensuite, le langage XML permet d'échanger d'une manière standardisée les documents sur le Web

²⁰⁹ IRI – Internationalized Resource Identifier

en assurant l'interopérabilité entre les applications. La couche du *cadre de description de ressources (RDF – Resource Description Framework)* est la première couche propre aux technologies du Web sémantique. Le RDF fournit une grammaire, un modèle d'encodage des données et des métadonnées sur n'importe quelles ressources identifiées par des URI sur le Web. De plus, le RDF est un langage qui permet de représenter l'information dans le cadre du Web sémantique.

Les couches *RDF-S (Schéma RDF)* et *Ontology (Ontologie)* permettent de spécifier le vocabulaire utilisé dans une représentation d'informations. Le modèle RDF et l'ontologie permettent de définir les termes du vocabulaire, la sémantique des termes et les relations entre les termes.

SPARQL (*Simple Protocol And RDF Query Language*²¹⁰) est le langage de base qui permet d'exécuter des requêtes sur l'ensemble des données enregistrées en format RDF. Le RIF (*Rule Interchange Format*) fournit un format commun pour « permettre l'échange des règles sur le Web en général et sur le Web sémantique en particulier » (Gandon, Faron-Zucker, et Corby, 2012, p. 134).

La couche supérieure permet *d'unifier la logique* exprimée dans les couches inférieures. Dans le contexte d'une architecture distribuée (e.g. le Web), l'information existe sous différentes formes (e.g. ontologies, bases de règles) et dans différents endroits (e.g. points d'accès SPARQL). Ce niveau permet d'incorporer les connaissances et les axiomes provenant de sources différentes. Le résultat obtenu est traité au niveau supérieur dans la couche *Preuve (Proof)*. Ce niveau fournit des moyens de validation sur l'ensemble des connaissances. Le niveau *Trust (Confiance)* a pour but de garantir la véracité et la fiabilité des connaissances et des auteurs. La confiance sera assurée grâce à la cryptographie, présente à chaque niveau pour signer numériquement les métadonnées et l'ensemble des documents Web (*Berners-Lee, 1997b*).

Enfin, le niveau *User Interface & Applications (Interface utilisateur et applications)* permet à l'utilisateur et aux applications d'exploiter ces technologies. Les applications

²¹⁰ Actuellement SPARQL est un acronyme récursif de *SPARQL Protocol And RDF Query Language*. Toutefois, à son origine, l'acronyme vient de *Simple Protocol And RDF Query Language*. <http://lists.w3.org/Archives/Public/public-rdf-dawg/2004JulSep/0500.html>

peuvent intervenir et exploiter les données à n'importe quel niveau de l'architecture du Web sémantique.

Puisque nous proposons la modélisation des entités du système d'apprentissage à travers les ontologies et le langage RDF, nous utiliserons les technologies illustrées ici dans l'implémentation de notre système actif d'apprentissage.

3.3.1. Identification des ressources dans le Web sémantique

Dans le Web sémantique, il « s'agit d'associer des identifiants abstraits à des notions, puis de relier ces identifiants entre eux pour former un réseau de connaissances avérées, partagées et universelles » (Poibeau, 2008, p. 33). Nous avons déjà vu que la ressource (p. 94) est définie en rapport avec son URI. En fait, un URI « définit des moyens simples et extensibles pour identifier une ressource » (Berners-Lee *et al.*, 2005). T. Berners-Lee affirmait que l'aspect de nommage et d'identification est « probablement le plus crucial de la conception et de la normalisation dans un système hypertextuel ouvert » (Berners-Lee, 1991). Les URI ont « deux fonctions très puissantes : (1) assigner des noms uniques et sans ambiguïté pour des choses et (2) spécifier le lieu où se trouve la chose [...] fonctions qui les rendent particulièrement utiles pour les ontologies et métadonnées » (Bratt, 2007, p. 25).

Selon la note *RFC3986*²¹¹ de l'IETF²¹², un URI est un « *identifiant* consistant en une *séquence de caractères*²¹³ correspondant à des règles spécifiques [...] permettant l'identification uniforme des ressources via des schémas de nommage [...] chaque schéma spécifiant comment l'identification est accomplie, affectée ou déléguée ». (Berners-Lee *et al.*, 2005).

Les principales caractéristiques intrinsèques d'un URI sont visibles dans son appellation :

²¹¹ *RFC (Request for Comments)* sont des publications des principaux organismes de normalisation pour l'Internet: Internet Engineering Task Force (<http://www.ietf.org>) et Internet Society (<http://www.internetsociety.org>). Le *RFC3986* est devenu standard en Janvier 2005 (<http://www.ietf.org/rfc/rfc3986>).

²¹² Internet Engineering Task Force

²¹³ Pour les URI, les caractères appartiennent au jeu de caractères US-ASCII (*American Standard Code for Information Interchange*).

- un URI est *uniforme*, car il permet une interprétation sémantique uniforme des conventions syntaxiques communes spécifiées pour des types différents d'identificateurs de ressources ;
- un URI est une *ressource*, car la spécification ne limite pas la portée de ce qui peut être considéré comme ressource et identifié par un URI ;
- un URI est un *identifiant*, car il englobe l'information nécessaire pour distinguer ce qui est identifié de toutes autres choses existant dans son contexte (champ, portée) d'identification (ibid.).

Un URI peut ensuite être classifié comme étant un *locateur*, un *nom* ou les deux. L'*Uniform Resource Locator (URL, RFC 1738)* est un type spécifique d'URI car il définit un moyen de localisation de la ressource en décrivant son mécanisme primaire d'accès. Un URL spécifie l'endroit où se trouve la ressource et cet endroit peut changer durant son cycle de vie.

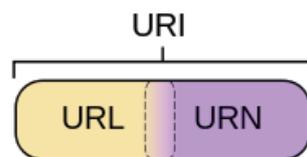


Figure 3-17. Relation entre l'URI, l'URL et l'URN

L'*Uniform Resource Name (URN, RFC 2141)* est un identifiant qui a les propriétés d'un schéma de nommage (e.g. les URI sous le schéma *urn*). Les URN sont des identifiants persistants ; ils sont indépendants de leur localisation et ne changent pas pendant leur cycle de vie. La syntaxe générale d'un URI est représentée dans la Figure 3-18.

URI = schema:"://" [userinfo"@"] host[:port] [path] ["?" query] ["#" fragment]

- schema : e.g. http, ftp, ssh ;
- userinfo : e.g. username:password ;
- host : e.g. nom de domaine, adresse IP ;
- port : e.g. 80 pour le port standard http, 22 pour ssh ;
- path : e.g. le chemin dans le système du serveur Web ;
- query : e.g. paramètre pour l'application Web ;
- fragment : e.g. identifie la partie spécifique d'un document Web.

Figure 3-18. La syntaxe générique de l'URI (source : <http://www.ietf.org/rfc/rfc3986>)

Dans la Figure 3-19 sont présentés deux exemples d'URI avec leurs composantes, la dernière étant également un URN.

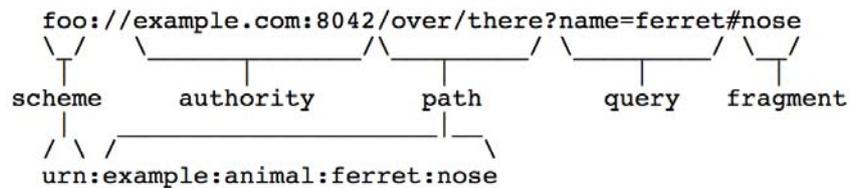


Figure 3-19. Exemples d'URI (source : <http://www.ietf.org/rfc/rfc3986>)

Le format *IRI* (*Internationalized Resource Identifier*), spécifié dans le standard RFC3987, étend la syntaxe des URI à un répertoire beaucoup plus large de caractères. Un IRI est « une séquence de caractères appartenant au jeu *Universal Character Set* (*Unicode/ISO 10646*²¹⁴) » (*Duerst et Suignard, 2005*). Sauf explicitations spécifiques, nous utilisons dans la suite de cette thèse les deux acronymes *IRI* et *URI* pour désigner les identifiants des ressources dans le cadre du Web. Les IRI identifient les *référents* (v. Figure 3-20).

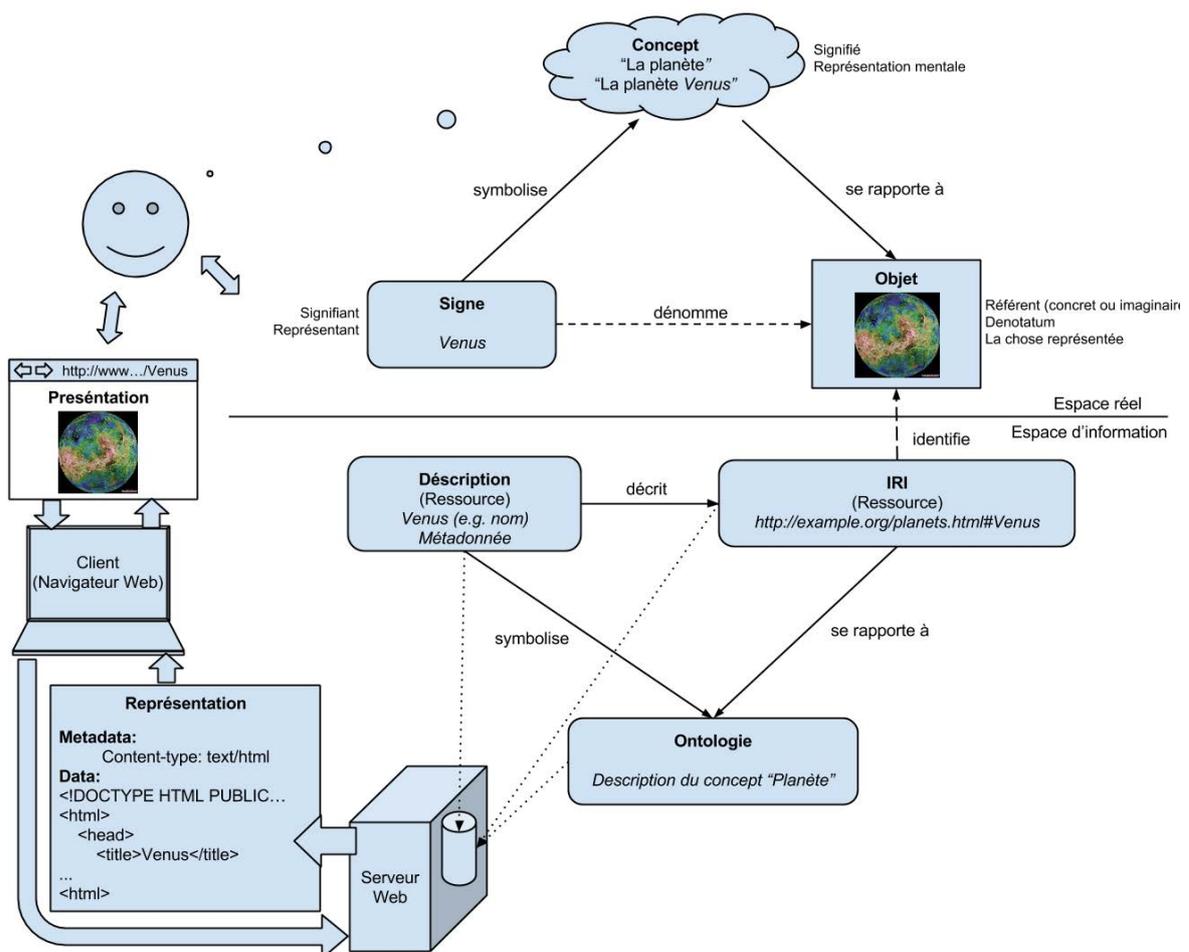


Figure 3-20. Triangle sémiotique dans l'espace d'information du Web

²¹⁴ http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=56921

Les identifiants IRI ont une portée globale (global scope) : l'apparition d'un même IRI dans des contextes différents dénote la même ressource (Cyganiak et Wood, 2013). Le procédé par lequel on récupère la *représentation* d'une *ressource* identifiée par un IRI s'appelle le *déréférencement* (Lewis, 2007). En tant qu'espace d'information, le Web ne peut fournir que des *représentations* ou des informations sur les ressources identifiées par un IRI (en particulier pour les objets existant en dehors d'espace d'information du Web).

Dans la pratique, les autorités²¹⁵ (e.g. les propriétaires d'un domaine) peuvent créer des IRI pour tout type de ressource (information ou non). Les bonnes pratiques recommandent de fournir toujours des informations sur les IRI créés. Si l'IRI identifie une ressource sous forme d'information, le propriétaire (créateur de l'IRI) devrait fournir des *représentations* de cette ressource. Si l'IRI identifie une ressource qui n'est pas donnée sous forme d'information, son propriétaire devrait fournir des *informations associées* qui, une fois déréférencées, fournissent des informations additionnelles sur la ressource originale. Les informations associées devraient être accessibles via le mécanisme de négociation du type de contenu fourni par le protocole HTTP (e.g. réponse avec le code 303). Dans le Tableau 3-1 sont présentées les interprétations possibles suite aux codes des réponses du HTTP.

Tableau 3-1. Sommaire des inférences suite à la déréférencement d'un IRI (source : ibid.)

Code de la réponse HTTP	Matériel retourné	Inférence
200 (Success)	Une représentation	La ressource est présentée sous forme d'information et une représentation de cette ressource est retournée.
303 (See Other)	Un IRI	La ressource peut avoir n'importe quelle forme. Il existe une ressource associée et son IRI est retourné. La ressource associée peut être ou non une ressource sous forme d'information.
4XX ou 5XX (error)	Rien	On ne peut pas s'exprimer sur la nature de la ressource.

²¹⁵ La possession d'un URI est « une relation entre un URI et une entité sociale, telle qu'une personne, une organisation ou une spécification » (Jacobs et Walsh, 2004). Cette relation donne au propriétaire de l'URI le droit (1) de déléguer (faire passer la possession de toutes ou d'une partie de ses URI à un autre propriétaire) et (2) d'affecter des URI (associer une ressource à un URI qu'il possède). Par convention sociale, la possession des URI est déléguée par IANA (*Internet Assigned Numbers Authority*) à d'autres propriétaires désignés, qui à leur tour peuvent subdéléguer la possession de ces URI. (ibid.)

Le mécanisme de *négociation du contenu http* permet au client de spécifier dans la *requête http* le format du contenu qu'il *accepte* (e.g. text/html, application/rdf+xml). Cela permet d'utiliser un seul identifiant IRI pour la requête HTTP et d'avoir la réponse dans le format adapté. La négociation du contenu permet de rendre les représentations des ressources disponibles à la fois pour l'utilisateur humain et pour les ordinateurs (e.g. logiciels, agents intelligents) (v. Figure 3-21).

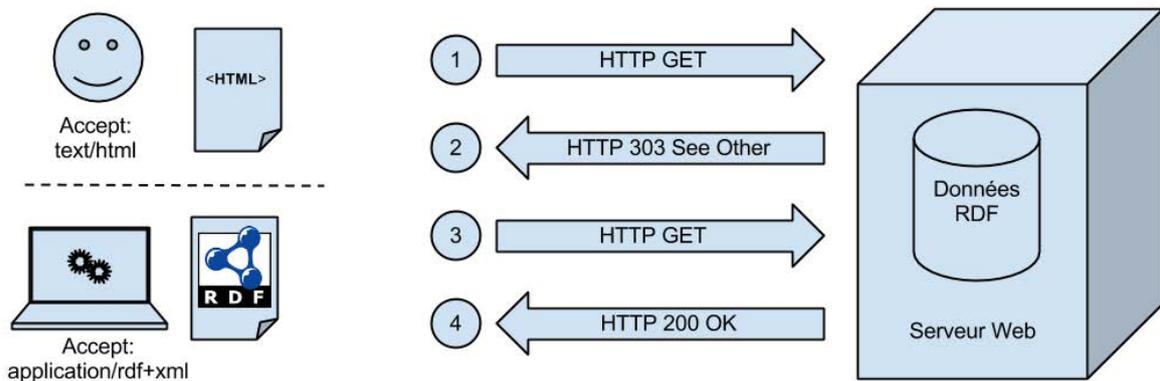


Figure 3-21. Négociation du contenu RDF à travers le HTTP

Il est important de faire la distinction entre l'IRI qui identifie une ressource abstraite et l'IRI qui identifie sa page HTML, par exemple. Sur le DBpedia, toutes les *ressources* ont le préfixe `http://dbpedia.org/resource/`. L'extraction des informations sur la ressource abstraite *Venus* se réalise en interrogeant le serveur à l'adresse `http://dbpedia.org/resource/Venus`.

```
ioan$ curl -I -H "Accept: text/html" http://dbpedia.org/resource/Venus
HTTP/1.1 303 See Other
Date: Thu, 22 Aug 2013 05:40:15 GMT
Content-Type: text/html; charset=UTF-8
Content-Length: 0
Connection: keep-alive
Server: Virtuoso/07.00.3203 (Linux) i686-generic-linux-glibc212-64 VDB
Accept-Ranges: bytes
Location: http://dbpedia.org/page/Venus
```

Extrait de code 3-1. Négociation du contenu http sur le serveur DBpedia pour le format text/html

En spécifiant le format du contenu accepté, le serveur répond en fournissant l'adresse `http://dbpedia.org/page/Venus` (v. Extrait de code 3-1) pour une représentation de la ressource en *HTML* et l'adresse `http://dbpedia.org/data/Venus.xml` (v. Extrait de code 3-2) pour la représentation sous forme RDF enregistrée en format XML.

```

ioan$ curl -I -H "Accept: application/rdf+xml"
http://dbpedia.org/resource/Venus
HTTP/1.1 303 See Other
Date: Thu, 22 Aug 2013 05:44:21 GMT
Content-Type: application/rdf+xml; qs=0.95
Content-Length: 0
Connection: keep-alive
Server: Virtuoso/07.00.3203 (Linux) i686-generic-linux-glibc212-64 VDB
Accept-Ranges: bytes
TCN: choice
Vary: negotiate,accept
Link:
<http://mementoarchive.lanl.gov/dbpedia/timegate/http://dbpedia.org/resourc
e/Venus>; rel="timegate"
Location: http://dbpedia.org/data/Venus.xml

```

Extrait de code 3-2. Négociation du contenu http sur le serveur DBpedia pour le format application/xml+rdf

Dans le cas où la représentation est en format HTML, la réponse est interprétée et présentée à l'utilisateur humain via le navigateur Web.

La spécification des préfixes utilisés dans la déclaration des IRI est réalisée par le propriétaire du nom de domaine (e.g. dans le cas de DBpedia.org `http://dbpedia.org/resource/` est utilisée pour les identifiants abstraits des ressources, `http://dbpedia.org/page/` pour les représentations en format HTML et `http://dbpedia.org/ontology/` pour les termes déclarés dans l'ontologie). En fait, ces préfixes qui simplifient l'écriture des données RDF ont davantage de fonctionnalités dans la pratique. Appelés *espaces de nommage* (*namespaces* en anglais), ils permettent de spécifier des conteneurs abstraits pour les ensembles d'identifiants déclarés. L'espace de nommage permet d'éviter les problèmes de conflit entre les noms des éléments et des propriétés. Dans le tableau Tableau 3-2 nous présentons les principaux espaces de nommage conteneurs des termes utilisés dans l'écriture des ontologies et des fichiers RDF.

Tableau 3-2. Espaces de nommage des vocabulaires de base du Web sémantique

Préfixe ²¹⁶	IRI de l'espace de nommage
rdf	http://www.w3.org/1999/02/22-rdf-syntax-ns#
rdfs	http://www.w3.org/2000/01/rdf-schema#
owl	http://www.w3.org/2002/07/owl#
xsd	http://www.w3.org/2001/XMLSchema#

²¹⁶ Pour la suite, nous utilisons ces préfixes pour représenter les IRI correspondants sans nécessairement invoquer l'IRI entier.

Dans les sections suivantes, nous présentons les éléments et les propriétés déclarés dans les vocabulaires correspondants à ces espaces de noms.

3.3.2. Modèle RDF (Resource Description Framework)

Basé sur le langage XML²¹⁷ (*eXtended Markup Language*) pour l'interopérabilité et situé sur la couche d'identification des ressources sur le Web (p. 145), le *modèle de description des ressources (RDF - Resource Description Framework)* fournit des outils pour la *description* et la *représentation* des ressources sur le Web. Selon le Consortium W3, RDF est un « modèle standard pour l'échange des données sur le Web » (W3.org, 2004). Centré sur la *description des ressources*, le cadre RDF « standardise les modèles, langages et syntaxes de ces descriptions » (Gandon *et al.*, 2012, p. 28).

Le modèle formel RDF est basé sur les concepts suivants :

1. *Ressource* (*Resource*) : entité physique (e.g. page web, personne, livre) ou abstraite (e.g. couleur, qualité). Une ressource peut être une composante d'une autre ressource. Toute ressource possède une IRI.
2. *Propriété* (*Property*) : un aspect spécifique, une caractéristique utilisée pour décrire une autre ressource. Une propriété est elle-même une ressource.
3. *Littéral* (*Literal*) : une chaîne de caractères non interprétée par RDF. Elle ne peut pas avoir de caractéristiques.
4. *Assertion* (*Statement*) : un triplet de la forme {sujet, prédicat, objet} dans lequel *sujet* est une *ressource*, *prédicat* est une *propriété*, et *objet* est soit une *ressource*, soit un *littéral*.

Ainsi, un triplet RDF permet de relier la *ressource* à décrire (*sujet*) à une *valeur* (*objet*) par l'intermédiaire d'une *propriété* (*prédicat*). Représenté sous la forme *nœud - arc -*

²¹⁷ Le langage XML est « un sous-ensemble du langage SGML (*Standard Generalized Markup Language*) [...], ayant pour objectif de permettre la transmission, la réception ou le traitement des documents SGML génériques sur le Web de la même façon qu'avec les documents HTML [...], conçu pour faciliter l'implémentation et l'interopérabilité avec les langages SGML et HTML » (Bray *et al.*, 2008). N'ayant pas pour but de *décrire des ressources*, le modèle principal (natif) de l'XML est un *modèle hiérarchique* selon une *structure arborescente*. Néanmoins, la représentation des données selon le modèle en graphe est toujours possible. L'interrelation des éléments dans la structure de l'XML peut se réaliser selon le modèle du HTML (e.g. l'attribut *href* de l'élément A peut pointer vers l'attribut *id* de l'élément B) (Bos, 2005). Toutefois, ces modèles ne spécifient pas d'une façon formelle le sens ou la signification des relations ou des éléments, sens ou signification qui, dans la structure arborescente, peuvent être influencés par la position dans la structure de l'arbre (le nom des éléments-parents).

nœud, le triplet {sujet, prédicat, objet} peut être modélisé par un graphe composé de deux *nœuds* et un *arc orienté* (v. Figure 3-22).



Figure 3-22. Un triplet RDF

Un ensemble de triplets RDF peut être modélisé par un graphe²¹⁸ RDF (Cyganiak et Wood, 2013) dans lequel les sommets sont les *ressources* et les *littéraux*. Les arcs sont qualifiés par les *prédicats*. Un *graphe RDF* est : « (1) un *multi-graphe*, c'est-à-dire un graphe qui peut contenir plusieurs arcs et même des boucles entre deux mêmes sommets ; (2) un *graphe orienté*, car chaque arc est orienté, allant du sommet représentant le sujet au sommet représentant l'objet ; et (3) un *graphe étiqueté*, car RDF attribue une étiquette à chaque arc et à chaque sommet du graphe » (Gandon et al., 2012, p. 30).

Si nous considérons l'énoncé « Sandro Botticelli est l'auteur de La Naissance de Vénus », sa représentation en graphe RDF prend la forme illustrée dans la Figure 3-23.



Figure 3-23. Exemple d'un énoncé sous forme de graphe

Afin de permettre aux machines de reconnaître et de manipuler ces entités, les ressources et les relations doivent être identifiés par des IRI. En fait, les composantes d'un triplet RDF peuvent être de trois types (*IRI*, *littéraux* ou *nœuds blancs*) de la façon suivante :

- le *sujet* peut être un *IRI* ou un *nœud blanc* (v. page 156) ;
- le *prédicat* peut être un *IRI* ;
- l'*objet* peut être un *IRI* ou un *littéral* ou un *nœud blanc*.

Si nous attribuons un IRI pour chaque composant de l'énoncé illustré dans la Figure 3-23, nous obtenons le graphe présenté dans la Figure 3-24.

²¹⁸ « Le *graphe* est la structure "abstraite" la plus commune rencontrée dans la science de l'informatique (computer science) » (Leeuwen, 1990a, p. 527).

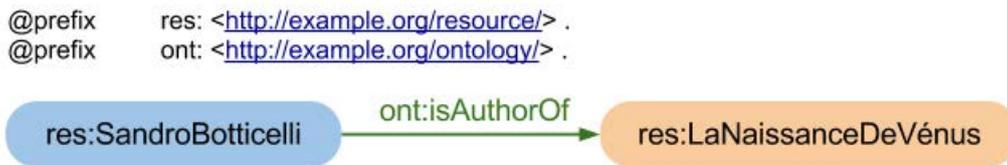


Figure 3-24. L'attribution des IRI dans un graphe RDF

Dans l'exemple illustré nous définissons l'espace de nom `http://example.org/resource/` pour contenir les identifiants attribués aux *ressources* représentées. L'espace de nom `http://example.org/ontology/` permet de définir les termes du *vocabulaire* utilisé pour la description des ressources. Au sujet et à l'objet, nous attribuons un IRI défini dans l'espace de nom des ressources. Au prédicat, nous attribuons un IRI appartenant à l'espace de nom du *vocabulaire* utilisé. L'IRI attribué au prédicat permet de spécifier le sens du celui-ci soit, dans ce cas, le fait que la ressource représentée par l'IRI du sujet *est auteur de* la ressource représentée par l'IRI de l'objet.

À ce point de notre exemple, les ressources sont identifiées par des IRI et, désormais, le graphe est interprétable par la machine. Pourtant, n'utilisant que des IRI, le graphe ne permet pas l'association des *valeurs* sous forme de *chaînes de caractères* aux ressources identifiées par les IRI. Le nœud *objet* doit donc accepter un autre type de ressource qui permettra d'enrichir le triplet RDF avec des valeurs sous forme de chaînes de caractères. Ce type de nœud est appelé *littéral* et consiste dans des chaînes de caractères arbitraires typées ou non typées représentant des valeurs telles que *du texte, des chiffres* ou *des dates*.

Un *littéral* peut être composé de deux ou trois éléments. Dans le premier cas, il consiste en *une forme lexicale* et *un IRI* spécifiant le type des données contenu dans la forme lexicale. Cet IRI détermine comment extraire la *valeur* du littéral de sa *forme lexicale* (e.g. `<"true", xsd:boolean>` et `<"1", xsd:boolean>` ont la valeur *vrai*). « Si et seulement si » l'IRI d'un littéral est `rdf:langString`, la valeur du littéral est une chaîne de caractères tagguée avec une étiquette spécifiant la langue dans laquelle la chaîne de caractères doit être interprétée. Les types de données utilisées sont ceux définis dans la spécification *XML Schema Definition Language (XSD)* (Peterson *et al.*, 2012). Si aucun

avons groupé les tableaux dans un *conteneur* (`rdf:Bag`) spécifié dans le *vocabulaire RDF*.

Le *vocabulaire RDF* (v. ANNEXE S) est une collection des références IRI déclarées dans l'espace de nom `http://www.w3.org/1999/02/22-rdf-syntax-ns#`. Ces références sont destinées à être utilisées dans les graphes RDF et leur *interprétation* est spécifiée dans la recommandation du Consortium W3 sur la sémantique du RDF (Hayes et McBride, 2004). Avançant dans la découverte des niveaux de l'architecture des standards du Web sémantique, nous étudierons les spécifications, les fonctions et les possibilités (e.g. expressivité) offertes par le *schéma RDF*.

3.3.3. Le schéma RDF. L'inférence dans RDF(S)

Le schéma RDF (RDFS), *extension sémantique* du RDF, fournit les *spécifications des termes* du vocabulaire utilisé dans un document RDF. Ce schéma introduit des *constructeurs de langage génériques* permettant la définition des vocabulaires spécialisés dans un domaine de discours (la partie du monde qui doit être décrite). Concrètement, il « fournit des moyens pour spécifier la *connaissance terminologique* sous forme des *hiérarchies de classes* et de *propriétés* ainsi que leur *interprétation sémantique* » (Hitzler *et al.*, 2009, p. 67).

La fonctionnalité de base de tout système de représentation de connaissances est la possibilité de spécifier le *type* d'une ressource afin de grouper les ressources en *classes* selon des critères spécifiques. En RDF, cette fonctionnalité est réalisée avec le prédicat `rdf:type`. Ce prédicat permet d'indiquer qu'une ressource est une *instance* d'une *classe*. Dans l'exemple illustré par la Figure 3-26, la ressource identifiée par l'IRI `res:SandroBotticelli` est une instance de la classe `ont:Person`. A contrario, une *classe* est un ensemble de ressources qui partagent les mêmes caractéristiques (la classe `ont:Person` représente l'ensemble des instances de ce type). RDFS fournit des moyens d'indiquer explicitement qu'une ressource est de type classe et qu'elle est donc une instance de la *classe de toutes les classes*. Cette « méta-classe » est prédéfinie dans les spécifications du vocabulaire RDFS (v. ANNEXE S) et elle est identifiée par l'IRI `rdfs:Class`.

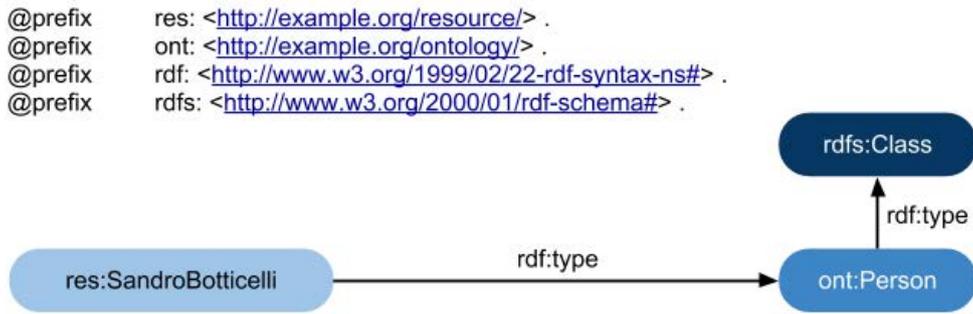


Figure 3-26. Le prédicat `rdf:type` et la métaclasse `rdfs:Class`

Déjà rencontrée sous le nom de *prédicat* dans le triplet RDF, une propriété est une relation entre deux ressources (sujet et objet). La classe de toutes les propriétés est présente dans l'espace de nom `rdf` et est identifiée par l'IRI `rdf:Property`.

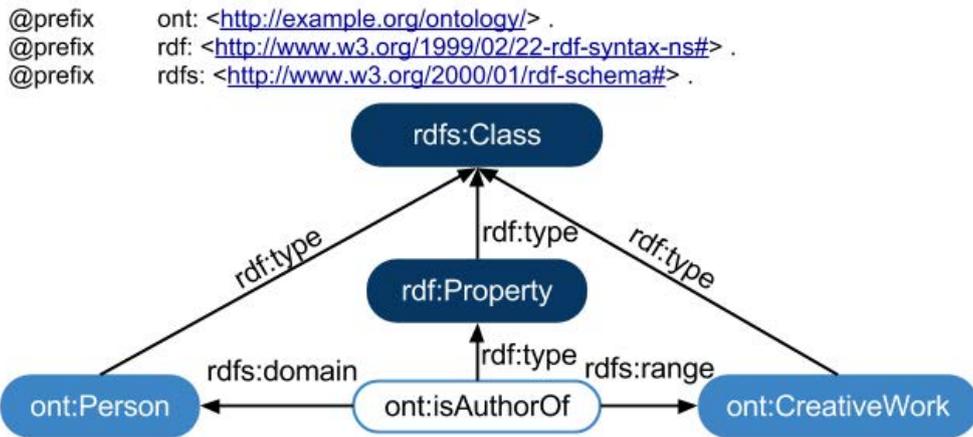


Figure 3-27. La définition d'une propriété RDF

Toujours fournies dans l'espace de nom du RDFS, les IRI `rdfs:domain` et `rdfs:range` sont des instances de la classe `rdf:Property` et permettent de spécifier le domaine et la portée d'une propriété (v. Figure 3-27). Les propriétés `rdfs:subClassOf` et `rdfs:subPropertyOf` permettent de construire des hiérarchies de classes et de propriétés. Toutes les constructions décrites en RDF sont appelées *ressources* et elles sont toutes des *instances* de la classe `rdfs:Resource`. La classe `rdfs:Resource` est une instance de la classe `rdfs:Class`. Toutes les classes déclarées en RDF sont des sous-classes de `rdfs:Resource`.

```

@prefix res: <http://example.org/resource/> .
@prefix ont: <http://example.org/ontology/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .

```

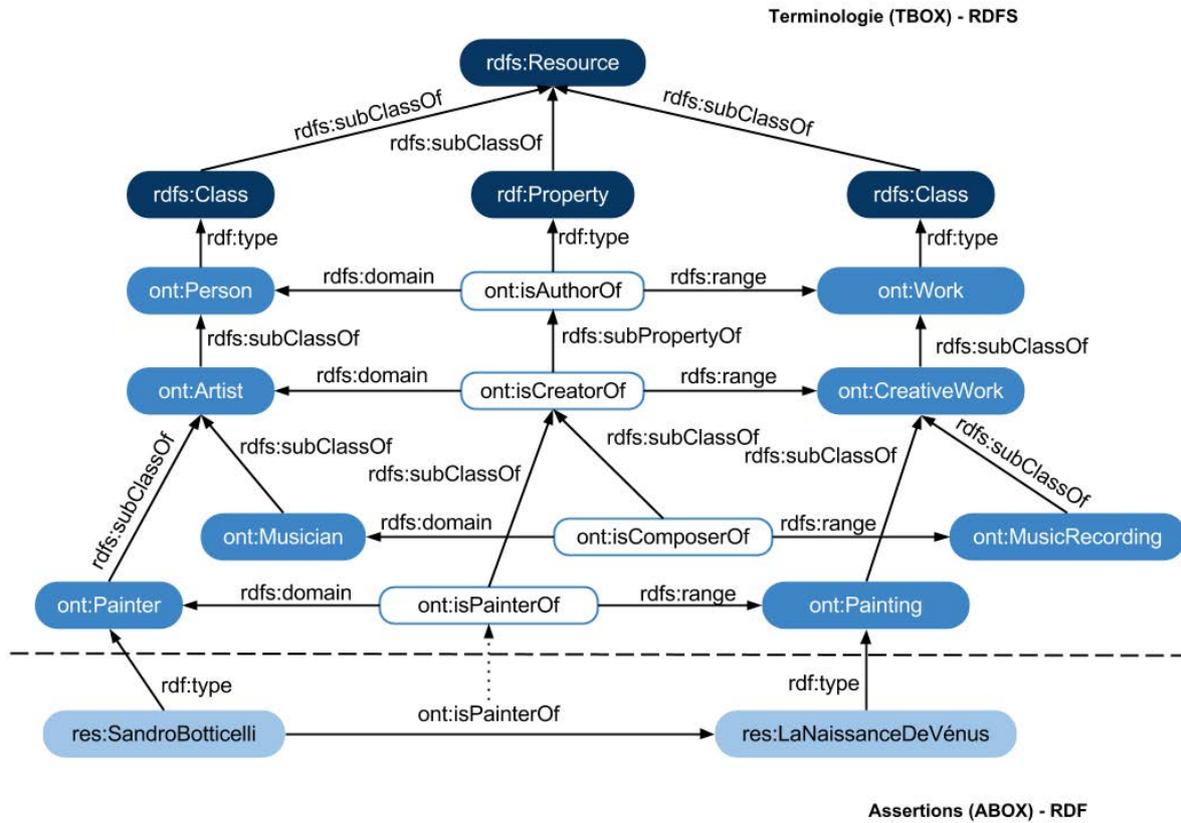


Figure 3-28. Exemple d'un graphe RDF avec la séparation de la connaissance terminologique

L'exemple présenté dans la Figure 3-28 illustre la séparation de la connaissance terminologique déclarée dans le *schéma RDF* et contenue dans la « *boîte terminologique* » (*Terminological Box - TBox*), de la connaissance représentée dans les assertions RDF et contenue dans la « *boîte des assertions* » (*Assertion Box - ABox*). Une *base de connaissances (KB)* est composée de l'ensemble de ces deux types de connaissances : $KB = \{TBox, ABox\}$.

Les références IRI déclarées dans le vocabulaire RDF(S) portent avec elles la signification décrite dans la sémantique formelle de RDF(S) (Hayes et McBride, 2004). Ainsi, selon la richesse et la complexité de la connaissance terminologique, l'association des deux ressources par une propriété permet à un agent intelligent d'inférer plus d'information sur la nature de ces ressources. Suite à la *sémantique formelle* exprimée dans les vocabulaires RDF et RDFS (voir l'ANNEXE F pour la liste des *triplets axiomatiques de RDF et RDFS*), les règles de déduction présentées dans le Tableau 3-3 (pour le V_{RDF}) et le Tableau 3-4 (pour le V_{RDFS}) s'appliquent aux graphes RDF.

Tableau 3-3. Règles de déduction dans le vocabulaire RDF (source : [http://www.w3.org/TR/rdf-
mt/](http://www.w3.org/TR/rdf-
mt/))

Nom de la règle	Si le graph contient	Alors on ajoute :
rdf1	uuu aaa yyy .	aaa rdf:type rdf:Property .
rdf2	uuu aaa lll .	_:nnn rdf:type rdf:XMLLiteral .
	ou lll est un littéral XML bien défini	ou _:nnn identifie un noeud blanc

Tableau 3-4. Règles de déduction dans le vocabulaire RDFS (source : [http://www.w3.org/TR/rdf-
mt/](http://www.w3.org/TR/rdf-
mt/))

Nom de la règle	Si le graph contient	Alors on ajoute :
rdfs1	uuu aaa lll.	_:nnn rdf:type rdfs:Literal .
	ou lll est un littéral simple (sans ou avec une étiquette de langue).	ou _: nnn identifie un noeud blanc.
rdfs2	aaa rdfs:domain xxx . uuu aaa yyy .	uuu rdf:type xxx .
rdfs3	aaa rdfs:range xxx . uuu aaa vvv .	vvv rdf:type xxx .
rdfs4a	uuu aaa xxx .	uuu rdf:type rdfs:Resource .
rdfs4b	uuu aaa vvv.	vvv rdf:type rdfs:Resource .
rdfs5	uuu rdfs:subPropertyOf vvv . vvv rdfs:subPropertyOf xxx .	uuu rdfs:subPropertyOf xxx .
rdfs6	uuu rdf:type rdf:Property .	uuu rdfs:subPropertyOf uuu .
rdfs7	aaa rdfs:subPropertyOf bbb . uuu aaa yyy .	uuu bbb yyy .
rdfs8	uuu rdf:type rdfs:Class .	uuu rdfs:subClassOf rdfs:Resource .
rdfs9	uuu rdfs:subClassOf xxx . vvv rdf:type uuu .	vvv rdf:type xxx .
rdfs10	uuu rdf:type rdfs:Class .	uuu rdfs:subClassOf uuu .
rdfs11	uuu rdfs:subClassOf vvv . vvv rdfs:subClassOf xxx .	uuu rdfs:subClassOf xxx .
rdfs12	uuu rdf:type rdfs:ContainerMembershipProperty .	uuu rdfs:subPropertyOf rdfs:member .
rdfs13	uuu rdf:type rdfs:Datatype .	uuu rdfs:subClassOf rdfs:Literal .

En exploitant ces règles de déduction, les agents intelligents peuvent expliciter des relations implicites existant dans un graphe RDF. Dans l'exemple illustré par la Figure 3-28, la propriété `ont:isPainterOf` est associée à la ressource `res:SandroBotticelli`, dans le domaine de cette propriété: donc les règles de déduction `rdfs2` et `rdfs9` permettent d'inférer que la ressource `res:SandroBotticelli` est à la fois de type (`rdf:type`) `ont:Painter`, `ont:Artist` et `ont:Person`. Le fait que chaque entité décrite dans un graphe RDF est de type `rdfs:Resource` est exprimé dans les règles `rdfs4a` et `rdfs4b`. Conformément à la règle `rdfs4a`, `res:SandroBotticelli` est aussi de type `rdfs:Resource`.

Ces types de déductions sont particulièrement intéressants dans l'implémentation des systèmes intelligents, et il sera avantageux de les exploiter dans le cadre d'un système actif d'apprentissage. Ces inférences permettent de décharger sur la machine (l'agent intelligent) une partie du travail de structuration (e.g. indexation, recherche) des ressources numériques d'apprentissage ou d'adaptation des ces ressources aux apprenants selon des caractéristiques spécifiées (e.g. prérequis, style d'apprentissage).

La principale limitation de l'expressivité du vocabulaire RDF(S) est l'impossibilité d'exprimer des négations. Cela conduit à l'impossibilité de vérifier l'inconsistance de la base de connaissances. Par exemple, en RDFS, on n'a pas la possibilité de définir des classes disjointes (on ne peut pas interdire à une instance d'appartenir simultanément à deux ou plusieurs classes). Si nous ajoutons dans notre graphe RDF un triplet associant la ressource `res:LaNaissanceDeVénus` à la ressource `ont:MusicRecording` par la propriété `rdf:type` (en tenant compte que la classe `ont:MusicalRecord` représente des enregistrements musicaux), l'énoncé de cette association sera l'équivalent, en français, de « *La Naissance de Vénus est un enregistrement musical* », alors que nous savons qu'il s'agit d'une peinture car la portée de la propriété `ont:isPainterOf` est la classe `ont:Painting` qui représente les peintures.

Dans l'architecture en couches des standards du Web sémantique (v. Figure 3-16), on trouve, au-dessus du RDFS, le langage des ontologies pour le Web (Web Ontology Language – OWL). Construit sur la base offerte par RDF(S), OWL surpasse les limitations du RDF(S) et fournit des moyens pour la construction des *ontologies riches* (avec

expressivité élevée). Cependant les *schémas RDF*, appelées aussi *ontologies légères*, fournissent des outils très puissants à la description des ressources pour le Web de données. Dans le chapitre Chapitre 4. Outils sémantiques pour la modélisation, nous présenterons le langage OWL avec ses différents profils et niveaux d'expressivité. Dans le présent chapitre, nous continuons avec la présentation du langage SPARQL qui permet de formuler des requêtes dans un ensemble RDF.

3.4. Langage d'interrogation et stockage du RDF

Plusieurs modes et formats existent pour enregistrer les graphes RDF dans un format exploitable par la machine. Nous allons présenter ensuite les formats de base ainsi que le langage d'interrogation pour le RDF.

3.4.1. Sérialisation du RDF : RDF/XML, Turtle

Décrites comme formats de sérialisation plusieurs syntaxes permettent de concrétiser le modèle de RDF afin d'être traitable par l'ordinateur. Étant donnée la popularité du XML dans les échanges de données, la sérialisation RDF/XML était la seule syntaxe standardisée en 2004²¹⁹ pour exprimer le graphe RDF. Toutefois, couramment les parseurs peuvent interpréter les autres syntaxes aussi. Depuis février 2013 la syntaxe *Turtle (Terse RDF Triple Language)* est aussi une recommandation de W3C²²⁰. Dans la suite nous présentons un exemple de sérialisation (v. Extrait de code 3-3), ainsi que les principales éléments d'un document RDF/XML.

²¹⁹ <http://www.w3.org/TR/2004/REC-rdf-syntax-grammar-20040210/>.

²²⁰ <http://www.w3.org/TR/turtle/>.

```

<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:res="http://ex.org/res/" xmlns:ont="http://ex.org/ont/"
xmlns:xsd="http://www.w3.org/2001/XMLSchema#">
  <rdf:Description rdf:about="http://ex.org/res/SandroBoticelli">
    <ont:birthDate
rdf:datatype="http://www.w3.org/2001/XMLSchema#gYear">1445</ont:birthDate>
    <ont:deathDate
rdf:datatype="http://www.w3.org/2001/XMLSchema#gYear">1510</ont:deathDate>
    <ont:isPainterOf>
      <ont:Painting
rdf:about="http://ex.org/res/LaNaissanceDeVenus" /></ont:isPainterOf>
    <ont:isPainterOf
rdf:resource="http://ex.org/res/LaNaissanceDeVenus" />
    <ont:isPainterOf
rdf:resource="http://ex.org/res/LePrintemps" />
    <ont:isPainterOf
rdf:resource="http://ex.org/res/VenusEtMars" />
  </rdf:Description>
</rdf:RDF>

```

Extrait de code 3-3. La syntaxe RDF/XML

L'élément `rdf:RDF` constitue la racine d'un document RDF/XML. Les attributs `xmlns` de cet élément déclarent les espaces de noms utilisés dans le document. Le nœud `rdf:Description` avec l'attribut `rdf:about` identifie la ressource à décrire. L'attribut `rdf:resource` permet de spécifier un identifiant URI. Les éléments `ont:birthDate`, `ont:deathDate` ou `ont:isPainterOf` représentent des propriétés RDF. L'élément `ont:Painting` spécifie le type de la ressource indiquée par l'attribut `rdf:about` correspondant à cet élément. L'attribut `rdf:datatype` indique les types de valeurs littérales. Les spécifications de la syntaxe RDF/XML sont décrites à l'adresse <http://www.w3.org/TR/REC-rdf-syntax/>.

Les langages *N-triples*, *Turtle* et *Notation3 (N3)* présentent des similarités dans le mode d'écriture des triplets RDF. En fait, le langage N-Triples est un sous-ensemble du langage Turtle, qui est un sous-ensemble du langage N3. Dans l'Extrait de code 3-4 nous présentons la syntaxe N-Triples. La plus simple de ces trois, N-Triples permet d'exprimer un triplet par ligne, présentant dans l'ordre d'apparence le sujet, le prédicat et l'objet, séparés entre eux par un espace. Chaque ligne se termine par un point « . ». Notons la redondance au niveau des espaces de noms qui sont réécrits chaque fois pour les termes utilisés, ou encore les sujet ou les propriétés des triplets.

```

<http://ex.org/res/SandroBoticelli> <http://ex.org/ont/birthDate>
"1445"^^<http://www.w3.org/2001/XMLSchema#gYear>.
<http://ex.org/res/SandroBoticelli> <http://ex.org/ont/deathDate>
"1510"^^<http://www.w3.org/2001/XMLSchema#gYear>.
<http://ex.org/res/SandroBoticelli> <http://ex.org/ont/isPainterOf>
<http://ex.org/res/LaNaissanceDeVenus>.
<http://ex.org/res/SandroBoticelli> <http://ex.org/ont/isPainterOf>
<http://ex.org/res/LePrintemps>.
<http://ex.org/res/SandroBoticelli> <http://ex.org/ont/isPainterOf>
<http://ex.org/res/VenusEtMars>.
<http://ex.org/res/LaNaissanceDeVenus> <http://www.w3.org/1999/02/22-rdf-
syntax-ns#type> <http://ex.org/ont/Painting>.

```

Extrait de code 3-4. La syntaxe N-Triples

La syntaxe *Turtle* permet d'éliminer cette redondance à travers la déclaration des *préfixes* et la réutilisation des sujets et des propriétés (v. Extrait de code 3-5). La déclaration des préfixes est réalisé en utilisant l'élément de syntaxe *@prefix*. La réutilisation des sujets est réalisé en utilisation le point-virgule « ; », alors que la réutilisation des sujets et prédicats est réalisé en utilisant la virgule « , ». Toujours dans l'idée d'améliorer la visibilité, Turtle permet l'utilisation de l'article anglais « a » à la place de la propriété `rdf:type`.

```

@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix ont: <http://ex.org/ont/> .
@prefix res: <http://ex.org/res/> .

res:LaNaissanceDeVenus      a          ont:Painting .
res:SandroBoticelli  ont:birthDate  "1445"^^xsd:gYear ;
                        ont:deathDate "1510"^^xsd:gYear ;
                        ont:isPainterOf res:LaNaissanceDeVenus ,
                                         res:LePrintemps ,
                                         res:VenusEtMars .

```

Extrait de code 3-5. La syntaxe Turtle

Le langage N3 intègre des éléments de syntaxe qui ne sont pas présents en Turtle (e.g. « {...} », « @forAll », « @forSome »). Les principales différences entre ces langages sont présentées sur <http://www.w3.org/TeamSubmission/turtle/>. Les exemples présentés dans cette thèse utilisent la syntaxe Turtle.

3.4.2. Le langage d'interrogation SPARQL

Une recommandation de W3C depuis janvier 2008, le langage SPARQL²²¹ (SPARQL Protocol And RDF Query Language) est conçu comme un langage d'interrogation pour le Web sémantique. Depuis mars 2013 W3C recommande un ensemble des spécifications sous la version SPARQL 1.1²²². Parmi ces spécifications nous rappelons *SPARQL 1.1 Query Language*²²³ pour l'interrogation des graphes RDF et *SPARQL 1.1 Update Language*²²⁴ pour la modification et la mise à jour des graphes RDF. Les spécifications incluent aussi le *SPARQL 1.1 Protocol*²²⁵ pour la description du protocole de transport des requêtes via HTTP. Dans l'Extrait de code 3-6 nous présentons l'anatomie d'une requête de type SPARQL Query.

Prologue	<pre>PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#> PREFIX ex: <http://example.org/></pre>
En-tête	<pre>ASK SELECT [DISTINCT] CONSTRUCT DESCRIBE [FROM FROM NAMED <dataset-uri>]</pre>
Corps de la requête	<pre>WHERE { ?s ?p ?o . [GRAPH <graph-uri> { ?s ?p ?o . }] [SERVICE [SILENT] <sparql-end-point> {?s ?p ?o }] [OPTIONAL { ?s ?p ?o . }] [FILTER expression] ... } [ORDER BY expression LIMIT integer > 0 OFFSET integer >0]</pre>

Extrait de code 3-6. Anatomie d'une requête SPARQL Query

Ainsi, nous pouvons structurer une requête de type *SPARQL Query* en trois parties distinctes. La première partie spécifie les espaces de noms utilisés dans la requête à

²²¹ Actuellement un acronyme récursif. Au départ *Simple Protocol And RDF Query Language* : <http://lists.w3.org/Archives/Public/public-rdf-dawg/2004JulSep/0500.html>, <http://www.w3.org/TR/rdf-sparql-query/>

²²² <http://www.w3.org/TR/sparql11-overview/>

²²³ <http://www.w3.org/TR/sparql11-query/>

²²⁴ <http://www.w3.org/TR/sparql11-update/>

²²⁵ <http://www.w3.org/TR/sparql11-protocol/>

travers le mot clé `PREFIX`, de manière similaire à la syntaxe Turtle. Dans la deuxième partie de la requête les mots clés `ASK`, `SELECT`, `CONSTRUCT` ou `DESCRIBE` déterminent le format général du résultat de la requête. Ainsi, le résultat d'une requête de type `ASK` peut être `vrai` ou `faux`, le résultat des requêtes de type `CONSTRUCT` ou `DESCRIBE` est un modèle RDF et le résultat de la requête de type `SELECT` est constitué d'un tableau contenant les données retrouvées dans le graph RDF interrogé, toute en respectant les contraintes (e.g. relations) déclarés dans la troisième partie de la requête. La troisième partie de la requête, initié par le mot clé `WHERE` spécifie donc, à travers des *triplets modèles* (e.g. templates, patterns) spécifiques, les relations contraignantes des variables cherchées avec d'autres données. Ces relations sont exprimées en utilisant la syntaxe Turtle. Les triplets séparés par le « . » au niveau du corps de la requête sont reliés entre eux via des *conjunctions*. Cela signifie que pour satisfaire l'ensemble du corps de la requête, tous les triples doivent être satisfaits. Le mot clé `OPTIONAL` permet d'ajouter des triplets facultatifs au corps de la requête.

Le résultat d'une requête SPARQL peut être modifié en utilisant des traitements pour limiter le nombre de réponses (e.g. `LIMIT`) ou pour trier les réponses selon un ou plusieurs critères (e.g. `ORDER BY`). Dans l'Extrait de code 3-7 nous présentons un exemple de requête SPARQL Query pour obtenir la liste des artistes nés avant 1500, ainsi que leur peintures²²⁶.

```
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX ont: <http://ex.org/ont/>
PREFIX res: <http://ex.org/res/>

SELECT ?artist ?painting
WHERE {
    ?artist          ont:isPainterOf          ?painting ;
                    ont:birthDate           ?bdate .

    FILTER (?bdate < "1500"^^xsd:gYear)
}
ORDER BY DESC(?artist)
```

Extrait de code 3-7. Exemple de requête SPARQL Query

²²⁶ Pour cette requête nous considérons une modélisation conforme à celle présenté dans l'Extrait de code 3-5.

La structure des requêtes de type *SPARQL Update* est analogue avec celle de *SPARQL Query*, les changements les plus évidents intervenants au niveau de l'en-tête de la requête. Les mots clés dans ce cas sont `INSERT`, `DELETE`, `INSERT DATA` ou `DELETE DATA`. Une requête de type *update* est réalisée via l'ensemble de requêtes `DELETE` et `INSERT`.

3.4.3. Le stockage des données RDF : le Triplestore

Le stockage des triplets RDF est réalisé en utilisant des systèmes de management de base de données spécialisés. Si dans les systèmes de base de données relationnelles (RDBMS) l'interrogation de données est réalisée à travers les langages d'interrogation de type SQL (Structured Query Language), l'interrogation des triplestores est réalisé à travers des requêtes SPARQL.

En général dans un *triplestore* les données RDF sont structurées de la façon suivante : *un triplestore* peut contenir un ou plusieurs *datasets*. Un *dataset* peut contenir un ou plusieurs *graphes* RDF (un graphe implicite et plusieurs graphes nommés). Enfin, le graphe RDF contient les triplets RDF effectives. Souvent les *triplestores* permettent la fonctionnalité de graphe nommé (Named Graph), et dans ce cas, on les appelle des *Quad Stores*. Un des éléments clés des triplestores est leur capacité d'intégrer et d'effectuer des inférences sur l'ensemble des données RDF. Quelques exemples de triplestores RDF natives²²⁷ sont : AllegroGraph²²⁸, BigData RDF Database²²⁹, Jena TDB²³⁰, Sesame²³¹, Stardog²³² ou OWLIM²³³. Sans tenir compte des fonctionnalités offerts par les moteurs SPARQL ou d'inférence, dans la vision du Web sémantique, nous pouvons considérer le Web même le triplestore ultime.

²²⁷ D'autre triplestores comme Jena SDB (<http://jena.apache.org/documentation/sdb/index.html> - son développement est arrêté) ou Virtuoso (<http://virtuoso.openlinksw.com/>) fournissent une couche qui permet l'utilisation des systèmes de bases de données relationnelles pour stocker les données RDF.

²²⁸ <http://www.franz.com/agraph/allegrograph/>

²²⁹ <http://www.systap.com/bigdata.htm>

²³⁰ <http://jena.apache.org/documentation/tdb/index.html>

²³¹ <http://www.openrdf.org/>

²³² <http://stardog.com/>

²³³ <http://www.ontotext.com/owlim>

Chapitre 4. Outils sémantiques pour la modélisation des connaissances

« *La philosophie est écrite dans ce livre gigantesque qui est continuellement ouvert à nos yeux (je parle de l'Univers), mais on ne peut le comprendre si d'abord on n'apprend pas à comprendre la langue et à connaître les caractères dans lesquels il est écrit. Il est écrit en langage mathématique, et les caractères sont des triangles, des cercles, et d'autres figures géométriques, sans lesquelles il est impossible d'y comprendre un mot. Dépourvu de ces moyens, on erre vainement dans un labyrinthe obscur.* »
Galileo Galilei, *Il Saggiatore*; traduction française de Christiane Chauviré, *L'Essayeur*, Les Belles-Lettres, Paris, 1980.

De façon générale, un *modèle* est une description simplifiée, une image réduite d'un certain aspect du monde réel. Il peut être utilisé pour « la structuration, la compréhension, la communication, l'explication, ou pour la prédiction d'une partie du monde réel » (Hitzler *et al.*, 2009, p. 2). Le modèle est « un ensemble d'assertions à propos d'un système étudié » (Seidewitz, 2003, p. 27), « une simplification d'un système construit avec un objectif spécifique » (Bézivin et Gerbé, 2001, p. 2). Le modèle donne de la « *signification* aux constructions syntaxiques » (Leeuwen, 1990b, p. 260).

La théorie des modèles est « une branche de la *logique mathématique* concernant les relations entre les structures mathématiques et les langages formels » (Manzano, 1999, p. 1), ou « une *théorie sémantique formelle* qui rapporte les expressions à des *interprétations*²³⁴ » (Hayes et McBride, 2004). Dans un sens plus large, la théorie des modèles a pour étude « l'interprétation d'une langue, formelle ou naturelle, au moyen

²³⁴ *L'interprétation* est définie comme « une description formelle minimale des aspects d'un monde qui est juste suffisante pour établir la vérité ou la fausseté d'une expression logique » [...] une interprétation satisfaisante est appelée « modèle » (Hayes et McBride, 2004). Un langage est considéré comme *pleinement interprétable*, si chacun de ses énoncés (assertions) possède une sémantique qui le fait soit *vrai*, soit *faux* (Hodges, 2013).

des ensembles structurés, avec la définition de la vérité d'Alfred Tarski²³⁵ (Tarski, Woodger, et Corcoran, 1956) comme paradigme » (Hodges, 2013).

Selon (Frigg et Hartmann, 2012), *les modèles* peuvent avoir « deux fonctions de représentation fondamentalement différentes », non-exclusives mutuellement : (1) une fonction de représentation d'une partie du monde (e.g. système ciblé) ou (2) une fonction de représentation d'une théorie, dans le sens où elle interprète les lois et les axiomes de cette théorie. Dans le premier cas, en fonction de la nature du système ciblé, les modèles peuvent être « des modèles de *phénomènes* ou des modèles de *données* » (ibid.).

La modélisation à laquelle nous nous intéressons consiste essentiellement en des données et des informations qui représentent les entités du système d'apprentissage. Selon (Lee, 1999, p. 315) un *modèle d'information* est « une représentation des *concepts*, des *relations*, *contraintes*, des *règles* et des *opérations pour spécifier la sémantique* des données pour un domaine de discours choisi ». La modélisation doit répondre au compromis entre la simplicité et le fait d'assurer une représentation correcte du sous-ensemble considéré du monde réel.

Les modèles peuvent être décrits « soit en utilisant un langage spécifique au domaine, soit en utilisant un langage de modélisation indépendant du domaine (UML par exemple) » (Teguiak, 2012, p. 33). Un langage de modélisation de l'information est « une *syntaxe formelle* qui permet aux utilisateurs de capturer la sémantique des données et des contraintes relatives à ces données » (Lee, 1999, p. 315). Les méthodes de modélisation à travers les technologies du Web sémantique se rapprochent de celles utilisées dans le cadre de l'Ingénierie Dirigée par les Modèles (IDM) ou *Model-Driven Engineering* ²³⁶ (MDE) (Guarino et Welty, 2000). Les approches utilisées pour la

²³⁵ Alfred Tarski (1901-1983), logicien, mathématicien et philosophe, formule une théorie qui permet d'assurer la consistance des prédicats exprimés dans un langage (e.g. surmonter le paradoxe du menteur : « Je mens », « Cette phrase est fausse »). Pour un *langage objet* (*L*) dans lequel sont exprimés les faits, la sémantique des prédicats exprimés dans ce langage devrait être définie dans un *métalangage* (*ML*), la théorie de la vérité de Tarski étant une théorie récursive : « "P" est vrai si et seulement si P » (e.g. "La neige est blanche" si et seulement si la neige est blanche).

²³⁶ *Model-Driven Engineering* (MDE) consiste en des méthodes de développement de systèmes logiciels basées sur l'utilisation de *modèles*.

modélisation dans le cadre du Web sémantique présentent des ressemblances et des éléments communs avec l'Architecture Dirigée par les Modèles (ADM) ou *Model-Driven Architecture*²³⁷ (MDA). L'ADM représente une des approches les plus utilisées de type IDM utilisée dans le développement des systèmes logiciels (Bézivin et Gerbé, 2001; Frankel, 2003; Frankel *et al.*, 2004; Staab *et al.*, 2010; Álvarez *et al.*, 2010). Découlant de pratiques existantes dans l'ingénierie des logiciels, l'objectif des modèles (dans le contexte de l'ADM) est de simplifier la description des logiciels. Le langage le plus utilisé pour la description de modèles est le langage *Unified Modeling Language (UML)*, qui permet de spécifier la sémantique d'un modèle à travers l'utilisation de *méta-modèles* (Atkinson, 2006).

Les ontologies, en tant que langages de modélisation ont pour but de capturer avec précision la connaissance du domaine de discours, en utilisant le formalisme développé dans le cadre des logiques formelles. Elles permettent la découverte de nouvelles connaissances par le biais de raisonnements automatiques et de moteurs d'inférence. Écrites le plus souvent dans le langage OWL (*Web Ontology Language*), les ontologies permettent de spécifier la sémantique d'une représentation en utilisant la *logique de description* (*ibid.*).

La hiérarchie des modèles (v. Figure 4-1) trouve des équivalences dans les technologies du Web sémantique de la manière suivante : le langage RDF permet de décrire des modèles, alors que les langages RDFS et OWL permettent de décrire des méta-modèles.

²³⁷ *Model-Driven Architecture* est un paradigme de développement de systèmes logiciels basé sur la construction et l'utilisation des *modèles* proposé par le *Object Management Group (OMG : <http://www.omg.org>)* en 2001. Ce groupe a aussi proposé le langage de modélisation *Unified Modeling Language (UML)*.

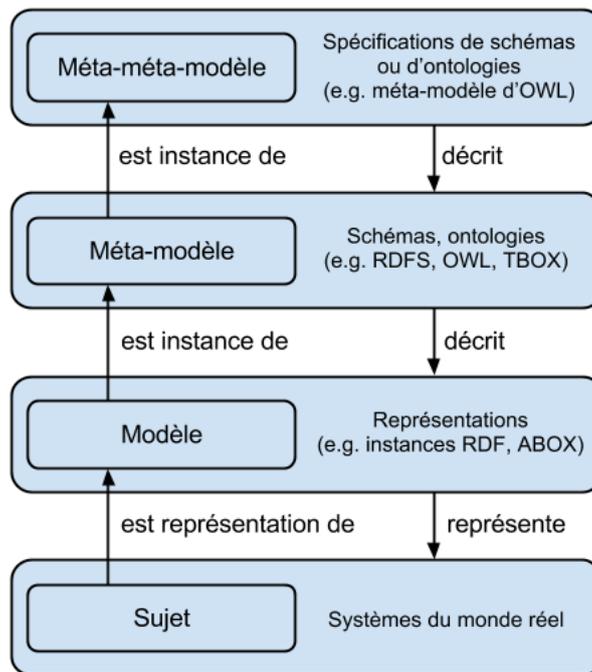


Figure 4-1. Hiérarchie des modèles (adaptée selon : Staab *et al.*, 2010; Omg.org, 2009)

Utilisant les technologies du Web, les langages RDF(S) et OWL sont particulièrement attirants pour la représentation de l'information dans un environnement distribué (e.g. le Web), pour l'exploitation de bases de connaissances partagées.

Comme on l'a vu précédemment (voir p. 158, Figure 3-26) deux types de connaissances rentrent dans la composition d'une base de connaissances :

- La *connaissance terminologique* (*boîte terminologique, TBox – Terminological Box*) comprend la déclaration des classes, des types de relations entre ces classes, ainsi que des contraintes (e.g. cardinalité, disjonction). Cette forme de connaissance correspond au méta-modèle. Cette notion de *méta-modèle* est « fortement associée à la notion *d'ontologie* » (Bézivin et Gerbé, 2001, p. 2). Ces *ontologies* peuvent être décrites en utilisant les langages RDFS ou OWL. Ce méta-modèle définit la *sémantique* du modèle : « le sens d'un modèle est donné par son interprétation » (Seidewitz, 2003, p. 27).
- La *connaissance assertionnelle* (*boîte assertionnelle, ABox – Assertional Box*) comprend les faits, les instances ou les assertions. Ce type de connaissance correspond au modèle. La modélisation à ce niveau consiste en l'identification des objets, via l'attribution d'un identifiant IRI, la spécification de leurs

caractéristiques et l'établissement des relations entre eux. Les assertions sont représentées en utilisant les termes définis dans la connaissance terminologique.

4.1. Ontologies et représentation des connaissances

Nous avons énuméré dans les paragraphes 1.3.1 et 1.3.2 les principales caractéristiques de la connaissance et de la représentation des connaissances. La Figure 4-2 montre les étapes et les éléments de construction d'une base de connaissances.

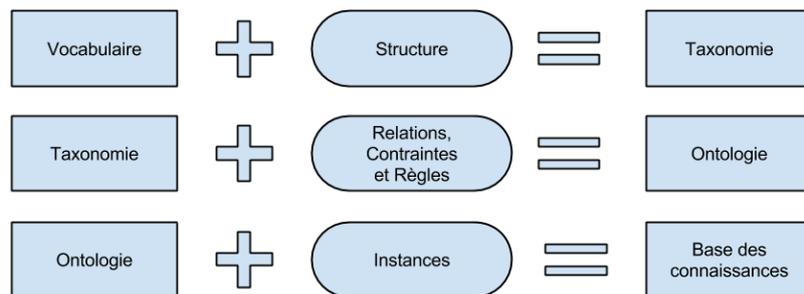


Figure 4-2. Les éléments d'une base de connaissances (source : Breslin *et al.*, 2009, p. 58)

Le vocabulaire, c'est-à-dire une collection de termes utilisés dans le domaine de discours considéré, permet d'identifier les types d'entités (e.g. classes) ainsi que leurs caractéristiques (e.g. propriétés) à modéliser. Ce vocabulaire peut être structuré (e.g. hiérarchique) pour former une taxonomie²³⁸. L'enrichissement de la taxonomie par des relations, des contraintes et des règles permet de construire l'ontologie. Enfin, l'ajout d'instances à l'ontologie forme la base de connaissances. L'élément central dans la composition d'une base de connaissances est donc l'ontologie. La notion d'ontologie et ses idées connexes ont une origine ancienne et bénéficient des travaux anciens et récents de plusieurs disciplines (e.g. philosophie, logique). Nous présentons ci-dessous les principales étapes de l'évolution du concept d'ontologie puis son intégration dans l'informatique.

²³⁸ Taxonomie du grec *τάξις (taxis)* « ordre » et *νομός (nomos)* « loi, science » : « étude théorique des bases, lois, règles, principes d'une classification », « science qui décrit, identifie et classe les êtres vivants en taxons selon leurs ressemblances » (Le Robert *et al.*, 2010, p. 2515). Dans son *Systema Naturæ*, Carl Linnaeus (ou Carl Von Linné) (1707-1778) pose les bases de la « classification biologique moderne, comme moyen de classier toutes les formes de vie ». « L'abondance des classifications taxonomiques suggère que les structures hiérarchiques constituent la base et, probablement, la structure la plus révélatrice dans la modélisation » (Hitzler, Krötzsch, et Rudolph, 2009, p. 3-4). Ici le terme se réfère au résultat du processus de classification.

4.1.1. Ontologie : de la philosophie à l'informatique

En philosophie, l'ontologie est une « partie de la métaphysique qui s'attache à l'étude ou à la théorie de l'être dans son essence, indépendamment de ses déterminations particulières » (Le Robert *et al.*, 2010, p. 1744). Emprunté du latin « *ontologia* », le terme a ses origines dans le grec ancien *ὄν* (*ôn, ontos*), signifiant « étant, ce qui est » et le suffixe *-logia*²³⁹ (« ontologie - Wiktionnaire », s. d.). Le mot « ontologie » est employé pour la première fois comme synonyme de « métaphysique » en 1606 par Jacob Lorhard, dans son livre « *Ogdoas Scholastica* ». En 1613 le mot apparaît dans « *Lexicon philosophicum* » de Rudolf Göckel, mentionné comme « ontologie, la philosophie de l'être » (« *ontologia, philosophia de ente* ») (« History of Ontology from Suárez to Kant (1597-1781) », s. d.).

En tant que science, l'Ontologie s'intéresse à la notion d'être, d'existence, aux catégories fondamentales de l'existant et étudie les propriétés les plus générales de l'être. Considérons, selon deux centrages possibles, quelques exemples de « *questions ontologiques* » témoignant de l'intérêt de cette discipline :

- Centrée sur *être* ou *exister*, l'ontologie formule des questions du type : Qu'est-ce qu'exister ? Qu'est-ce qu'une substance ? Qu'est-ce que l'identité ?
- Centrée sur *l'existence des catégories*, sur la classification des entités et leurs propriétés, elle conduit à des questions telles que : Qu'est-ce qu'un tout ? Qu'est-ce qu'une relation ? Qu'est-ce que la dépendance ? Qu'est-ce que la causalité ? Qu'est-ce qu'une propriété ? Qu'est-ce qu'un état ? Qu'est-ce qu'un type ? (Mulligan, 2000).

Les préoccupations concernant l'ontologie remontent à l'antiquité grecque et parcourent toute l'histoire de la philosophie. Bien qu'il trouve son origine lointaine dans la "science" aristotélicienne de « l'être en tant qu'être »²⁴⁰, le concept d'*ontologie* n'est repris en informatique que par une commode analogie, car l'*ontologie informatique* se constitue dans une optique très modestement pragmatique. Cependant, cette *ontologie* au sens informatique du terme n'en constitue pas moins l'élément central de la représentation des connaissances dans le cadre du Web sémantique, conformément à

²³⁹ Tiré du grec ancien *λόγος*, *logos*, signifiant « discours », « traité », « parole », « fable », « bruit », « lettres ».

²⁴⁰ Aristote, *Métaphysique*, Livre Γ, éd. Vrin, tome I (http://philia.online.fr/txt/arist_050.php).

l'assomption suivante : « pour les systèmes à base de connaissances ce qui "existe" est exactement ce qui peut être représenté » (Gruber, 1993, p. 199).

Pourtant, différentes idées et caractéristiques issues de l'ontologie "philosophique" sont transférées et utilisées en informatique : classifications, catégories, propriétés, langage et terminologie partagés, identification pertinente et sans ambiguïté, capacités de calcul logique. Le terme *ontologie* est un « terme technique désignant un artefact qui est conçu pour un objectif, à savoir permettre la modélisation des connaissances dans un domaine quelconque, réel ou imaginaire » (ibid.).

Donnons quelques définitions d'*ontologie* dans ce contexte :

- « spécification explicite et formelle d'une conceptualisation partagée » (Borst, 1997, p. 12) ;
- « compréhension partagée d'un certain domaine d'intérêt » (Uschold et Gruninger, 1996, p. 97) ;
- « modèle computationnel d'une certaine partie du monde » (Huhns et Singh, 1997, p. 81).

Les caractéristiques principales de l'ontologie, d'après ces définitions, sont :

- spécification : concepts, propriétés, relations, fonctions, contraintes, axiomes ;
- conceptualisation : modèle abstrait d'une certaine partie du monde que l'on veut modéliser ;
- explicite : le type des concepts et les contraintes relatives à leur utilisation sont explicitement définis ;
- formelle : interprétable (traitable, calculable, manipulable) par la machine ;
- partagée : l'ontologie capture la connaissance consensuelle, qui n'est pas propre à un individu mais validée par un groupe.

En informatique, l'ontologie est donc une spécification de la conceptualisation de la partie du monde à modéliser. Elle permet le partage et la réutilisation des connaissances. Elle permet l'expression d'*engagements ontologiques*, c'est-à-dire des *accords* d'utilisation de vocabulaires spécifiques et des accords sur l'interprétation de la sémantique associée à ces vocabulaires. Ces engagements ontologiques permettent aux utilisateurs de l'ontologie (e.g. agents intelligents) d'exploiter, partager et interpréter le

contenu de l'ontologie. L'ontologie permet de surmonter les différences de terminologie et de réaliser la correspondance entre les différentes entités appartenant aux différents mondes décrits à travers différentes ontologies. Dans le cadre du Web sémantique, *l'ontologie* est l'élément central de la représentation des connaissances et de la construction des bases de connaissances. *L'ontologie* est une notion abstraite définie en termes de *spécification structurelle*, indépendante de syntaxes concrètes utilisées pour écrire les ontologies particulières.

4.1.2. Le langage d'ontologie pour le Web (OWL)

L'ontologie est une notion abstraite définie en termes de *spécifications structurelles*, indépendantes des syntaxes concrètes d'échange pour les ontologies. Chaque ontologie est associée à un *document d'ontologie*, qui contient physiquement l'ontologie enregistrée au format texte utilisant une des syntaxes disponibles pour les langages d'ontologie pour le Web (« OWL 2 Web Ontology Language Structural Specification and Functional-Style Syntax (Second Edition) », 2012). La syntaxe d'échange de base pour le langage OWL est la *syntaxe RDF/XML*. Outre cette syntaxe, obligatoire pour les logiciels conformes, d'autres syntaxes²⁴¹ peuvent être utilisées (« OWL 2 Web Ontology Language Document Overview (Second Edition) », s. d.). Dans ce qui suit, nous employons la *syntaxe Turtle* pour illustrer les éléments de base composants d'une ontologie.

Le langage d'ontologie pour le Web (Web Ontology Language – OWL²⁴²) permet de surmonter les limites du RDF(S) et fournit des moyens qui permettent plus *d'expressivité*²⁴³ dans la représentation des modèles. Toutefois l'augmentation en

²⁴¹ La *syntaxe OWL/XML* (facilite le traitement avec des outils XML), la *syntaxe fonctionnelle* (plus facile pour voir la structure formelle des ontologies), la *syntaxe Manchester* (facile pour lire et écrire des ontologies dans le langage OWL DL), la *syntaxe Turtle* (facile pour lire et écrire des triplets RDF).

²⁴² La variation de l'acronyme (OWL au lieu de WOL) était proposée par Tim Finin dans un email de 27 Decembre 2001 destiné au groupe de travail du W3C pour les avantages : « (1) la prononciation évidente qui est plus facile pour l'oreille, (2) elle ouvre de grandes opportunités pour les logos, (3) les hiboux (en anglais *owls*) sont associés à la sagesse; (4), il a une histoire de retour intéressant » (Finin, 2001).

²⁴³ L'expressivité d'un langage représente sa capacité de distinguer entre les modèles. La comparaison de l'expressivité des deux langages peut être faite si ils sont interprétés sur les mêmes classes de modèles. Le langage L_1 est plus expressif que le langage L_2 si L_1 peut distinguer deux modèles qui sont indiscernables par le langage L_2 (<http://iaoa.org/isc2012/docs/expressivity.pdf>).

expressivité vient aux dépens de l'efficacité du processus de raisonnement. Les classes de complexité supérieure augmentent la complexité des calculs. Pour concilier ces deux aspects, le compromis entre l'expressivité du langage et la complexité des calculs impose la structuration du langage OWL en trois sous-classes, appelées espèces, selon leur niveau d'expressivité (v. Figure 4-3).



Figure 4-3. Sous-langages pour le langage OWL

Ces sous-langages ont une relation d'inclusion (v. Figure 4-3), OWL Lite étant le langage le plus restrictif, inclus dans OWL DL. Fondé sur les logiques de description et augmentant son niveau d'expressivité, OWL DL est inclus dans OWL Full, langage permettant le niveau maximal d'expressivité. Le langage OWL Lite est destiné aux utilisateurs ayant principalement besoin d'une hiérarchie de classifications et de contraintes simples (Smith, Welty, et McGuinness, 2004).

OWL DL offre le meilleur compromis, en permettant une grande expressivité tout en assurant la complétude des calculs (il garantit que toutes les inférences seront calculées) et la décidabilité (tous les calculs s'achèveront dans un intervalle de temps fini). OWL DL et OWL Full contiennent les mêmes éléments de construction, la différence résidant dans les restrictions sur l'utilisation de certaines caractéristiques et sur l'utilisation des fonctionnalités RDF. De plus, OWL DL est structuré en trois sous-profils permettant d'alléger la complexité de calcul pour les profils spécifiés. Ces profils sont présentés dans le § 4.1.3.

Le langage *OWL Full* offre le niveau maximum d'expressivité, sans garanties calculatoires. OWL Full permet la libre utilisation et le mélange d'OWL avec RDFS. A la

manière de RDFS, OWL n'impose pas une séparation stricte entre les classes, les propriétés, les individus ou encore les valeurs des données.

La première spécification du langage OWL est proposée par W3C en février 2004. Le standard actuel, une extension de la version précédente est le langage OWL 2, proposé en décembre 2012. OWL 2 inclut le langage OWL et assure qu'une ontologie écrite en OWL 1 reste valide en OWL 2²⁴⁴. Nous présentons, dans ce qui suit, les principaux éléments composant un document d'ontologie.

L'en-tête du document d'ontologie contient des informations sur *les espaces de nom* utilisés dans le document, la *version* et quelques *annotations* sur l'ontologie (v. Extrait de code 4-1).

```
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix xml: <http://www.w3.org/XML/1998/namespace> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@base <http://semlearn.pu-pm.univ-fcomte.fr/ontologies/asls> .

<http://semlearn.pu-pm.univ-fcomte.fr/ontologies/asls> rdf:type owl:Ontology ;
                                                    rdfs:label "Active Semantic
Learning System Ontology" ;
                                                    rdfs:label "Active Semantic
Learnni" ;
                                                    owl:versionInfo "v0.0.5" ;
                                                    owl:versionIRI
<http://semlearn.pu-pm.univ-fcomte.fr/ontologies/asls/0.0.5> .
```

Extrait de code 4-1. L'en-tête d'un document d'ontologie

Notons le préfixe `http://www.w3.org/2002/07/owl#` présenté dans l'Extrait de code 4-1. Ce préfixe est utilisé pour définir les éléments de construction déclarés dans le langage OWL. Les annotations doivent être déclarées comme étant de type `owl:AnnotationProperty`. Sous certaines conditions, en OWL DL, des annotations peuvent être déclarées sur les classes, les propriétés, les individus et sur les en-têtes de l'ontologie. Les propriétés d'annotation en OWL peuvent être déclarées en utilisant les termes suivants : `owl:versionInfo`, `rdfs:label`, `rdfs:comment`, `rdfs:seeAlso`, `rdfs:isDefinedBy`.

²⁴⁴ Sauf spécification explicite, nous considérons la version OWL 2 du langage.

Les composants de base pour la construction d'une ontologie sont *les classes et les propriétés* (appelés aussi *rôles*). Les classes représentent des *ensembles d'entités, d'individus (instances de classe)*. Dans les paragraphes suivants, nous allons présenter les principales constructions d'OWL utilisées pour la déclaration de *classes*, de *propriétés* et d'*individus*.

Les classes sont définies en utilisant le terme `owl:Class`²⁴⁵. Deux classes prédéfinies existent : (1) `owl:Thing`, la classe la plus générale, comprend tous les individus comme instances de cette classe ; et (2) `owl:Nothing`, la négation de `owl:Thing`, qui par définition ne permet pas d'avoir des *instances*. Un individu peut être *assigné à une classe* en utilisant le prédicat `rdf:type`. Dans ce cas, l'individu est une *instance* de cette classe. La hiérarchie des classes peut être déclarée à travers la propriété `rdfs:subClassOf`. OWL élargit les possibilités de déclaration de classes à travers les propriétés suivantes :

- `owl:equivalentClass` – définit des classes équivalentes, c'est-à-dire contenant les mêmes ensembles d'individus ;
- `owl:disjointWith` – définit des classes disjointes, c'est-à-dire que les individus ne peuvent pas être des instances simultanées des classes déclarées disjointes ;
- `owl:oneOf` – permet la déclaration d'une classe en énumérant tous ses individus.

De plus, en OWL, les classes peuvent être déclarées en utilisant des restrictions sur les propriétés. Une restriction, indiquée par le terme `owl:Restriction`, spécifie une classe dont tous les individus satisfont la restriction. La restriction s'applique à la propriété indiquée par le terme `owl:onProperty` et peut spécifier la *valeur* que la propriété peut prendre ou sa *cardinalité*, c'est-à-dire le nombre de valeurs que la propriété peut prendre.

Les termes suivants permettent de définir une contrainte sur *les valeurs d'une propriété* associées à une classe :

- `owl:allValuesFrom` – spécifie une classe d'individus pour lesquels toutes les valeurs de la propriété sur laquelle est appliqué la contrainte sont des instances de la classe ou des valeurs du type de données indiqué par cette contrainte ; cette restriction est analogue au *quantificateur universel* (\forall) de la logique des

²⁴⁵ `owl:Class` est un sous-classe de la classe `rdfs:Class`.

prédicats ; cette propriété a pour valeur une description de classe ou un type de donnée ;

- `owl:someValuesFrom` – spécifie une classe dont au moins une des valeurs de la propriété sur laquelle est appliquée la restriction est instance de la classe ou des valeurs du type indiqué par cette contrainte ; cette restriction est analogue au *quantificateur existentiel* (\exists) de la logique des prédicats ; cette propriété a pour valeur une description de classe ou un type de donnée ;
- `owl:hasValue` – permet de spécifier une classe dont les membres sont spécifiés dans la valeur de la propriété sur laquelle est appliquée la restriction ; cette propriété a pour valeur un individu ou une valeur littérale ;

Les termes pour définir une contrainte sur *la cardinalité d'une propriété* sont les suivants²⁴⁶ :

- `owl:maxCardinality` – décrit une classe dont tous les individus ont *au plus le nombre spécifié* de valeurs distinctes pour la propriété sur laquelle la restriction s'applique ;
- `owl:minCardinality` – décrit une classe dont tous les individus ont *au moins le nombre spécifié* de valeurs distinctes pour la propriété sur laquelle la restriction s'applique ;
- `owl:cardinality` – décrit une classe dont tous les individus ont *un nombre spécifié de valeurs distinctes* pour la propriété sur laquelle la restriction s'applique.

En plus des restrictions sur les propriétés, *l'intersection*, *l'union* et *le complément* sont des opérateurs qui peuvent être aussi utilisés pour la composition des classes en OWL. Cela est réalisé à travers les termes suivants :

- `owl:intersectionOf` – décrit une classe composée des individus appartenant à l'intersection des classes spécifiées ; analogue à la *conjonction logique*, à l'opérateur ET (\wedge) ;

²⁴⁶ OWL Lite inclut l'utilisation de ces trois types de contraintes de cardinalité, mais seulement quand il est utilisé avec les valeurs « 0 » ou « 1 ».

- `owl:unionOf` – décrit une classe composée des individus appartenant au moins à une classe parmi les classes spécifiées ; analogue à la *disjonction logique*, à l'opérateur OU (\vee);
- `owl:complementOf` – décrit une classe composée des individus qui n'appartiennent pas à la classe spécifiée ; analogue à la négation logique, à l'opérateur NEGATION (\neg).

Le langage OWL comprend deux types de propriétés : (1) *propriétés objet (rôles abstraits)* déclarées en utilisant le terme `owl:ObjectProperty` ou (2) *propriétés données (rôles concrets)* déclarées en utilisant le terme `owl:DatatypeProperty`. Les propriétés objet permettent de connecter les individus entre eux, pendant que les propriétés concrètes permettent de connecter un individu avec *une donnée* en spécifiant aussi le type de données de sa valeur. OWL permet de définir des hiérarchies des propriétés, notamment en utilisant le prédicat `rdfs:subPropertyOf`. Les concepts `owl:ObjectProperty` et `owl:DatatypeProperty` sont des *sous-concepts* de la classe `rdf:Property`. Le domaine et la portée d'une propriété peuvent être déclarés de la même manière qu'en RDFS, en utilisant les prédicats `rdfs:domain` et `rdfs:range`.

D'autres relations sur les propriétés peuvent être exprimées avec les termes suivants :

- `owl:equivalentProperty` – permet de définir l'équivalence²⁴⁷ entre deux propriétés ;
- `owl:inverseOf` – permet de définir une nouvelle propriété à partir d'une propriété initiale en commutant le domaine et la portée de la propriété initiale.

La cardinalité globale sur les propriétés peut être définie avec les termes suivants :

- `owl:FunctionalProperty` – permet de spécifier que la propriété en question admet une seule (et unique) valeur dans la portée de la propriété pour chaque individu spécifié dans le domaine de la propriété (analogue à une clé dans une table de base de données relationnelle).

²⁴⁷ Il faut faire la distinction entre *l'équivalence* et *l'égalité*. L'équivalence des propriétés définit les mêmes extensions de la propriété. L'égalité peut se déclarer à travers le terme `owl:sameAs`. L'utilisation du terme `owl:sameAs` pour les *propriétés* est permise uniquement en OWL Full.

- `owl:InverseFunctionalProperty` – permet de spécifier que la propriété en question admet une seule (et unique) valeur dans le domaine de la propriété pour chaque individu spécifié dans la portée de la propriété.

OWL permet la déclaration de relations *transitives* et *symétriques* (sous-propriétés de `owl:ObjectProperty`) en utilisant les termes `owl:TransitiveProperty` et `owl:SymmetricProperty`.

OWL spécifie quelques termes pour manipuler les individus déclarés dans une ontologie. Le mécanisme d'identification sur le Web assure qu'une IRI identifie une entité (e.g. personne, document) au niveau global, mais il n'assure pas qu'il n'existe pas d'autre IRI identifiant la même entité. En fait, le mécanisme d'identification des entités, des individus sur le Web, n'assume pas que différents noms (e.g. IRI) puissent désigner différents individus. A l'image du Web, et à la différence des autres langages de représentation de connaissances, l'OWL n'impose pas *l'assomption de nom unique (Unique Name Assumption)*. *L'assomption de nom unique* « suppose » que pour différents individus soient utilisés différents noms.

Prenons comme exemple les assertions suivantes, dont la première déclare que la propriété `ont:hasBirthPlace` (a lieu de naissance) est *fonctionnelle* (admet une seule valeur). Autrement dit, quelqu'un peut avoir un seul et unique lieu de naissance. Chacune des deux assertions suivantes spécifie que `res:Botticelli` a pour lieu de naissance `res:Florence`, et respectivement `res:Florence_Italy`.

```
ont:hasBirthPlace    rdf:type          owl:FunctionalProperty .
```

```
res:Botticelli      ont:hasBirthPlace  res:Florence .
```

```
res:Botticelli      ont:hasBirthPlace  res:Florence_Italy .
```

Le fait d'avoir déclaré la propriété `ont:hasBirthPlace` comme *fonctionnelle* permet aux moteurs d'inférences de conclure que `res:Florence` et `res:Florence_Italy` référencient *la même entité*, donc deux *noms (e.g. IRI)* pour le même individu. Pour expliciter l'égalité ou la distinction des individus, en OWL on utilise les termes suivants :

- `owl:sameAs` – permet de spécifier que deux IRI référencient la même entité, donc d’avoir plusieurs identifiants pour la même entité ;
- `owl:differentFrom` – permet de spécifier que deux IRI référencient deux entités différentes.

Pour spécifier la distinction des entités en OWL, `owl:AllDifferent` et `owl:distinctMembers` permettent de définir une liste dont les membres sont tous distincts.

Présentant des garanties de calculabilité et décidabilité, la construction des ontologies est particulièrement attirante, notamment en utilisant le langage OWL DL. Dans le paragraphe suivant, nous allons nous concentrer sur l’utilisation et les constructions du langage OWL DL et de ses profils.

4.1.3. Logiques de description. Profils d’OWL

L’aspect *formel* de l’ontologie implique sa représentation dans une forme interprétable par la machine. Issus de la *logique formelle*, les langages formels et la *sémantique associée* permettent d’exprimer les modèles clairement et sans ambiguïté. Nous rappelons les principaux points dans l’évolution de la *logique*²⁴⁸ au cours des années.

George Boole (1815 – 1864) fonde la logique symbolique moderne et soutient que « la logique doit être attachée aux mathématiques et non à la philosophie » (Pire, s. d.). En 1854, il publie « *Une exploration des lois de la pensée* » (An investigation of the Laws of Thought) où il introduit la logique propositionnelle²⁴⁹ (ou booléenne). Introduisant le principe de quantification, Gottlob Frege (1848 – 1925) pose les fondations de la logique de premier et deuxième ordre. Il publie en 1879 le *Begriffsschrift (Idéographie)*, une œuvre dans laquelle il propose un langage formel pour « exprimer un contenu à travers des signes écrits d’une manière plus précise et plus claire que cela est possible à travers les mots » (Frege, 1879, p. 2). L’avènement des formalismes d’expressions logiques conduit aux convictions que toute pensée rationnelle, et donc même les mathématiques, pourrait être basée sur la logique. C’est le *logicisme*, qui n’est pas seulement une sous-

²⁴⁸ Logique : du grec *logikê*, dérivé de *logos*.

²⁴⁹ « Il réduit la logique à un calcul propositionnel, l’algèbre booléenne, dont les lois sont calquées sur celles de l’algèbre classique » (Pire, s. d.).

discipline des mathématiques, mais qui permet de déduire n'importe quelle vérité mathématique à partir d'axiomes en vertu de déductions purement logiques. Dans la même idée de défense du *logicisme*, Alfred N. Whitehead (1861 – 1947) et Bertrand Russell (1872 – 1970) écrivent « *Principia Mathematica* » (1910-1913) en trois volumes, qui demeure aujourd'hui l'ouvrage de référence pour logique. Kurt Gödel (1906 – 1978) et Alan Turing (1912 – 1954) explorent les limites de la logique et introduisent des notions centrales comme la *décidabilité* et la *calculabilité*. (Hitzler *et al.*, 2009, p. 6). Ces notions sont centrales dans l'élaboration et la construction des bases de connaissances et des ontologies. Pouvant être identifiés comme un *fragment décidable* de la *logique de premier ordre*, le langage OWL DL bénéficie de ces évolutions et travaux au fil des années.

D'autre part, « la logique de description peut être reliée aux *réseaux sémantiques*²⁵⁰ » (ibid., p. 159). Les *logiques de description* « descendent des *réseaux de structures d'héritage*, qui ont été introduits pour répondre à l'ambiguïté des premiers *cadres* et *réseaux sémantiques* » (Baader, 2003, p. 44). Le besoin d'avoir une sémantique des termes déterminée et définie a conduit au développement des *logiques de description*. Les *logiques de description* forment une famille de langages de *représentation de connaissances* utilisée pour modéliser la connaissance d'un domaine d'application d'une façon structurée et formelle.

Les descriptions élémentaires sont les *concepts atomiques* (e.g. l'équivalent des classes d'OWL) et les *rôles atomiques* (e.g. l'équivalent des propriétés d'OWL). Les descriptions complexes se construisent en utilisant des *constructeurs de concepts* sur les descriptions élémentaires. Les langages de description se différencient entre eux en fonction des *constructeurs* qu'ils fournissent. Ainsi, l'expressivité du langage est déterminée par ses constructeurs. Toutefois, l'augmentation de l'expressivité du langage s'accompagne d'une augmentation de la complexité du modèle. Permettant un choix de la classe de complexité du modèle, l'OWL DL admet trois groupes de constructeurs, appelés *profils OWL*. Dans ce qui suit, nous présentons les constructeurs fournis par le langage OWL DL et ses différents sous-profils. Dans les notations suivantes, *A* est utilisé pour représenter

²⁵⁰ Les réseaux sémantiques permettent la modélisation de relations simples entre des entités et des classes à travers des propriétés, de manière comparable à RDFS (Hitzler *et al.*, 2009, p. 159).

les *concepts atomiques*, R pour les *rôles atomiques*, C et D pour les *descriptions de concepts*.

Le langage minimal d'intérêt pratique est le *langage attributif*, noté \mathcal{AL} (*Attributive Language*). Les éléments et les constructeurs du langage \mathcal{AL} sont les suivants :

- concept atomique : A ;
- concept universel : \top ;
- concept vide (*bottom concept*) : \perp ;
- négation atomique : $\neg A$;
- intersection : $C \sqcap D$;
- restriction sur la valeur : $\forall R. C$;
- quantification existentielle *limitée* : $\exists R. C$.

Notons les constructeurs suivants :

- l'union, indiquée par $\mathcal{U} : C \sqcup D$;
- la quantification existentielle complète, indiquée par $\mathcal{E} : \exists R. C$;
- la négation des concepts, indiquée par $\mathcal{C} : \neg C$.

En ajoutant les trois constructeurs cités ci-dessus au langage \mathcal{AL} , nous obtenons le langage $\mathcal{ALU\mathcal{E}\mathcal{C}}$. Toutefois les équivalences $C \sqcup D \equiv \neg(\neg C \sqcap \neg D)$ et $\exists R. C \equiv \neg \forall R. \neg C$ montrent que l'union et la quantification existentielle peuvent être exprimées en utilisant la négation. Cela conduit à la notation \mathcal{ALC} (*Attributive Language with Complement*). Le langage \mathcal{ALC} est considéré comme « le plus restreint des langages des logiques descriptives comprenant toutes les connectives booléennes et fournissant un algorithme de subsomption *cohérent (sound) et complet* » (Tobies, 2001, p. 11). Le \mathcal{ALC} DL permet la construction de concepts complexes à partir des constructeurs de la logique propositionnelle²⁵¹ : \sqcap (et, intersection des classes), \sqcup (ou, union des classes) et \neg (non, complémentation des classes). Les constructeurs du langage \mathcal{ALC} sont les suivants :

$$C, D ::= A \mid \top \mid \perp \mid \neg C \mid C \sqcap D \mid C \sqcup D \mid \forall R. C \mid \exists R. C \quad \text{Équation 4-1}$$

²⁵¹ En logique propositionnelle, les arguments sont des *phrases*. La syntaxe de la logique propositionnelle comprend : (1) un ensemble de *variables propositionnelles* (e.g. $\{p, q, r\}$) ; (2) les connecteurs $\wedge, \vee, \rightarrow, \neg$ correspondant respectivement à la conjonction (et), à la disjonction (ou), à l'implication (si...alors...) et à la négation (non) ; et (3) les parenthèses « (» et «) ».

Dans le Tableau 4-1, nous présentons quelques associations des termes du langage OWL aux constructeurs logiques.

Tableau 4-1. Représentation en OWL des constructeurs des logiques de description et leur syntaxe en FOL (First-Order Logics)

Constructeur OWL	Syntaxe DL	Exemple	Syntaxe FOL
owl:intersectionOf	$C_1 \sqcap \dots \sqcap C_n$	Human \sqcap Artist	$C_1(x) \wedge \dots \wedge C_n(x)$
owl:unionOf	$C_1 \sqcup \dots \sqcup C_n$	Musician \sqcup Painter	$C_1(x) \vee \dots \vee C_n(x)$
owl:complementOf	$\neg C$	\neg Musician	$\neg C(x)$
owl:oneOf	$\{x_1\} \sqcup \dots \sqcup \{x_n\}$	{NaissanceDeVénus} \sqcup {LeJeuneHommeEtLesArts}	$x = x_1 \vee \dots \vee x = x_n$
owl:allValuesFrom	$\forall P.C$	\forall hasPainted.Paint	$\forall y. P(x, y) \rightarrow C(y)$
owl:someValuesFrom	$\exists P.C$	\exists hasPainted.Paint	$\exists y. P(x, y) \wedge C(y)$
owl:maxCardinality	$\leq nP$	≤ 1 isPaintedBy	$\exists^{\leq n} y. P(x, y)$
owl:minCardinality	$\geq nP$	≥ 2 hasPainted	$\exists^{\geq n} y. P(x, y)$

En OWL, le concept universel $\top \equiv C \sqcap \neg C$ est représenté à l'aide du concept owl:Thing. Le concept vide $\perp \equiv C \sqcup \neg C$ (bottom concept) est représenté à l'aide du terme owl:Nothing. La disjonction, correspondant en OWL au terme owl:disjointWith peut être représentée en utilisant $C \sqcap D \sqsubseteq \perp$, ou par $C \sqsubseteq \neg D$. Le domaine d'un rôle, rdfs:domain peut être représenté par $\exists R. \top \sqsubseteq C$, alors que owl:range permet de spécifier la portée d'un rôle et peut être exprimée avec $\top \sqsubseteq \forall R. C$, C représentant le domaine et R le rôle.

Dans les sections suivantes, nous rappelons quelques constructeurs, ainsi que leurs notations, permettant l'extension des logiques de description \mathcal{ALC} :

- \mathcal{S} = \mathcal{ALC} + transitivité des rôles (e.g. $Tr(hasAncestor)$) ;
- \mathcal{R} – inclusion des rôles (e.g. $hasParent \circ hasBrother \sqsubseteq hasUncle$) ;
- \mathcal{H} – hiérarchies des rôles (e.g. $isPainterOf \sqsubseteq isCreatorOf$) ;
- \mathcal{O} – nominaux, énumération d'individus (e.g. $\{NaissanceDeVénus\}$) ;
- \mathcal{I} – rôles inverses (e.g. $isCreatorOf \equiv hasCreated^-$) ;
- \mathcal{N} – restrictions de cardinalité (e.g. $\geq 2hasCreated, \geq nR$ et $\leq nR$) ;
- \mathcal{Q} – restrictions qualifiées de cardinalité (e.g. $\geq 2hasCreated. Artist$) ;
- \mathcal{F} – rôles fonctionnels (e.g. $\leq 1isAuthorOf$) ;

- \mathcal{D} – types de données

L'ajout des différents constructeurs à l'ensemble \mathcal{ALC} permet la construction des différentes familles de logique (v. Figure 4-4), correspondant aux différents niveaux de complexité²⁵².

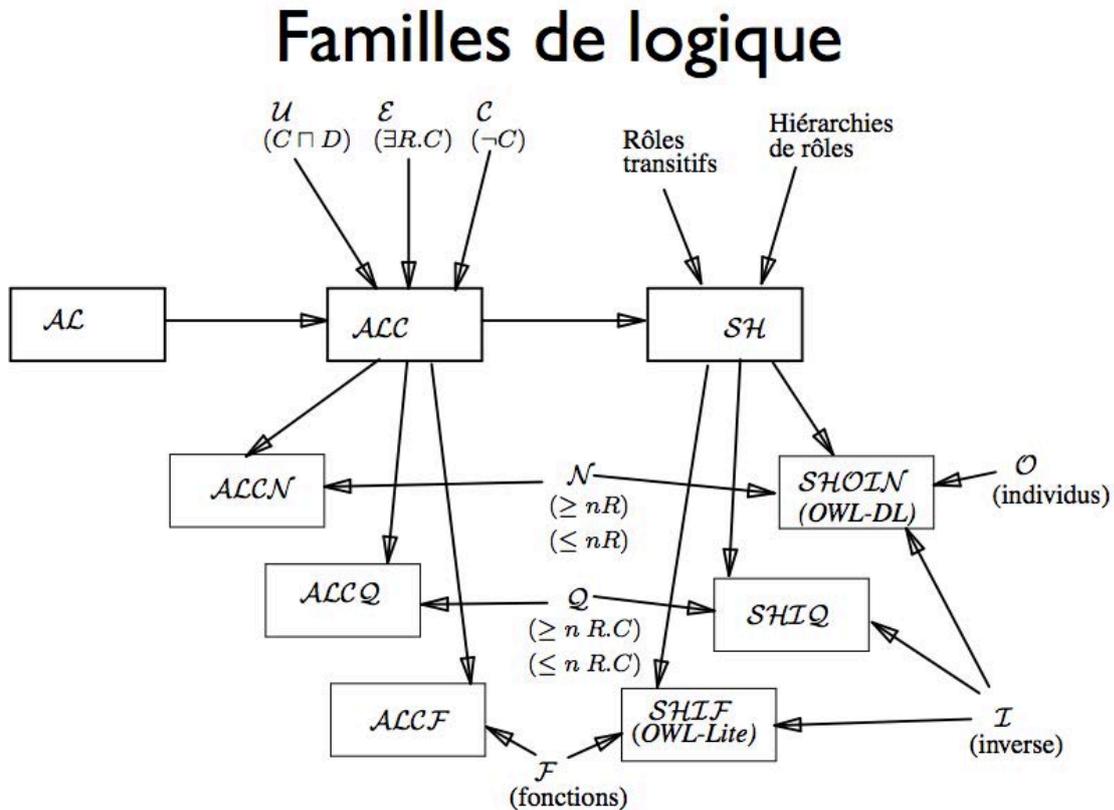


Figure tirée de : Michel Gagnon, *Logique descriptive et OWL*, École polytechnique, Montréal, 2012.

Figure 4-4. Familles de logique

D'après les remarques précédentes, le langage OWL Lite correspond à l'ensemble $\mathcal{SHIF}(\mathcal{D})$. Il offre une expressivité limitée, résultant en une complexité et un temps de calcul réduits. Le langage OWL DL intègre l'ensemble des éléments d'OWL Lite, tout en augmentant l'expressivité du langage et en assurant la décidabilité sur l'ensemble des connaissances ainsi modélisées. Le langage OWL DL correspond à l'ensemble $\mathcal{SHOIN}(\mathcal{D})$. OWL Full ne faisant pas partie des logiques de description, il n'est pas concerné par cette notation. En fait, OWL Full comprend les mêmes éléments que la langage OWL DL mais sans les restrictions imposées dans OWL DL²⁵³.

²⁵² <http://www.cs.man.ac.uk/~ezolin/dl/> présente une application qui permet de calculer la complexité de raisonnement dans les logiques de description.

²⁵³ Par exemple, en OWL DL, une classe ne peut pas être un individu ou une propriété.

OWL 2 DL correspond quant à lui à l'ensemble $\mathcal{SROIQ}(\mathcal{D})$ et inclut les structures du langage OWL DL (Hitzler *et al.*, 2009, p. 166). Le langage OWL 2 DL définit trois profils, chacun offrant des avantages différents dans des scénarios d'application spécifiques. Il s'agit de trois sous-langages (sous-ensembles syntaxiques) permettant l'abaissement de leur *classe de complexité* (v. l'ANNEXE G) par adoption de l'ensemble des constructeurs, les restrictions d'utilisation de ces constructeurs résultant de l'adaptation des algorithmes²⁵⁴ de résolution. L'évaluation de la complexité des logiques de description est considérée selon quatre dimensions :

1. la complexité de données : mesurée en termes de nombre d'assertions (taille de l'*ABox*) ;
2. la complexité taxonomique : mesurée en termes de nombre d'axiomes (taille de *TBox*) ;
3. la complexité des requêtes : mesurée en termes de nombre d'éléments d'une *requête conjonctive*²⁵⁵, de la taille de la requête ;
4. la complexité combinée : mesurée en fonction des complexités énumérées ci-dessus.

La Figure 4-5 montre les classes de complexité correspondant aux sous-langages d'OWL 2 DL.

²⁵⁴ On distingue deux types d'algorithmes dans la définition des classes de complexité : déterministe et non-déterministe. Différent de l'algorithme déterministe, dont toutes les étapes sont déterminées et connues à l'avance, l'algorithme non-déterministe est celui dans lequel, pour une instance d'entrée donnée, chaque étape intermédiaire présente une ou plusieurs possibilités. Cela signifie qu'il peut y avoir plus d'une voie et que l'algorithme peut en choisir *arbitrairement* une. Tous les chemins ne conduisent pas au succès, c'est-à-dire au résultat souhaité. L'algorithme non déterministe fonctionne de façon qu'il peut choisir toujours un chemin qui mène au succès, donc au résultat correct. (Bhowmick, P., 2008, *Lecture 29: Nondeterministic Algorithms*. <http://intinno.iitkgp.ernet.in/courses/91/wfiles/43255>).

²⁵⁵ Une requête conjonctive est construite à partir des *conjonctions* (" \wedge ", « et ») entre les éléments composants et le quantificateur existentielle (" \exists ").

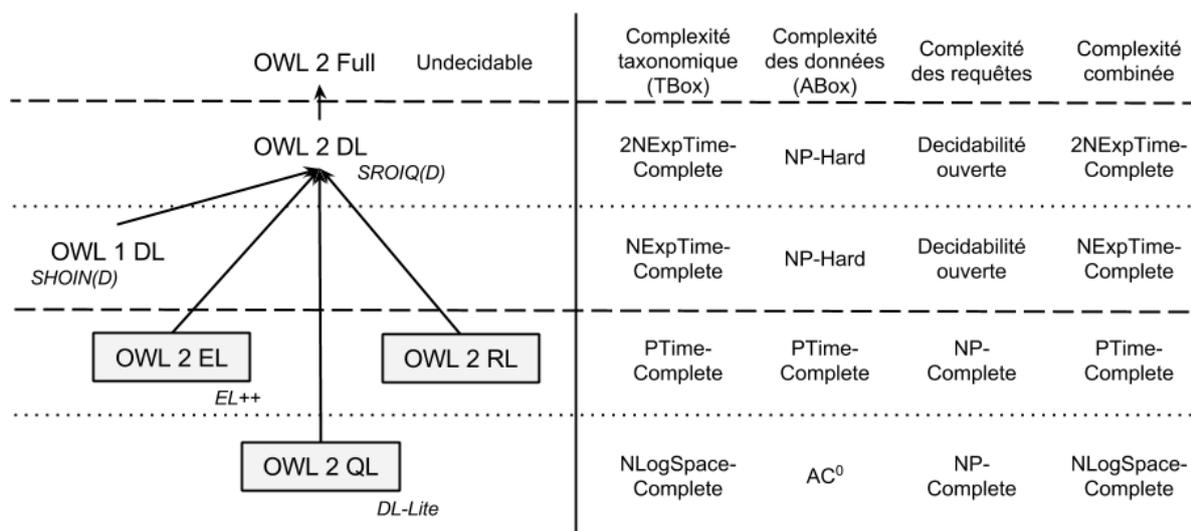


Figure 4-5. Les profils d'OWL 2 DL et leurs classes de complexité (selon : <http://www.w3.org/TR/owl-profiles/>)

L'abaissement d'une classe de complexité est réalisé en sacrifiant le niveau d'expressivité du langage. Cela permet de répondre efficacement aux spécificités de l'application pour laquelle l'ontologie est utilisée, ainsi que de diminuer les ressources requises par les agents de raisonnement. Les trois profils du langage OWL 2 DL sont l'OWL 2 EL, OWL 2 QL et OWL 2 RL. Les algorithmes de résolution d'OWL 2 DL font partie de la classe de complexité $2NExpTime$, alors que les algorithmes de résolution pour les sous-langages d'OWL existent dans la classe de complexité $PTime$ et AC^0 (v. la Figure 4-5). La combinaison de n'importe laquelle de deux de ces sous-langages conduit à l'augmentation de la classe de complexité du langage résultant.

OWL 2 EL admet une complexité combinée dans la classe $PTime$. Pour maintenir les caractéristiques computationnelles, l'utilisation des certains constructeurs est limitée. Par exemple, les constructeurs `owl:allValuesFrom`, `owl:unionOf`, `owl:complementOf`, `owl:disjointWith` ainsi que les restrictions de cardinalité ne sont pas permises. Ce sous-langage est adapté aux ontologies comprenant un grand nombre des classes et propriétés. OWL 2 EL est basé sur les logiques de description \mathcal{EL}^{++} . Ces logiques de description étendent la liste des constructeurs de l'ensemble \mathcal{EL} , avec des constructeurs comme le concept vide ou les nominaux. Les logiques de description \mathcal{EL} fournissent des constructions pour exprimer le concept universel, la conjonction et la restriction existentielle. (Baader, Brandt, et Lutz, 2008).

OWL 2 QL admet une complexité dans la classe *NLogSpace*. Il impose des restrictions sur les parties *sujet* et *objet* dans l'utilisation du terme `rdfs:subClassOf`. La transitivité, `owl:propertyChainAxiom`, `owl:sameAs` et l'attribution de rôles négatifs (compléments) ne sont pas permises. OWL 2 QL est adapté pour faciliter l'implémentation en utilisant des systèmes de bases de données relationnelles classiques. Il est efficace pour répondre rapidement à des requêtes sur de grands ensembles de données.

OWL 2 RL admet une complexité dans la classe *PTime*. Il impose des restrictions sur les parties *sujet* et *objet* dans l'utilisation du terme `rdfs:subClassOf`. Les rôles universel et vide sont exclus, de même que la réflexivité entre restrictions. OWL 2 RL est adapté à l'exploitation des *outils d'inférences* à base des *règles*. Il fournit un équilibre entre le raisonnement extensible et le niveau d'expressivité.

On trouvera plus d'informations sur les classes d'expressivité, les restrictions et les constructeurs utilisés dans chaque profil d'OWL à l'adresse : <http://www.w3.org/TR/owl-profiles/>.

4.1.4. La sémantique du modèle théorique OWL

La syntaxe s'intéresse à la définition des règles pour la construction des *expressions*, des séquences de mots et de symboles (chaînes de caractères ou de signes) valides et *bien formées* (en anglais *well-formed*) (« SYNTAXE : Définition de SYNTAXE », s. d.). Par opposition à la syntaxe, *la sémantique* vise à donner une *signification* à la syntaxe. Elle définit les règles qui spécifient comment la signification des expressions, des séquences complexes de symboles, peut être dérivée des séquences atomiques des symboles. La *sémantique formelle*, définie dans le cadre des langages logiques, vise à exprimer la signification des séquences de symboles (programmes informatiques) de telle manière que les affirmations sur les séquences de symboles (programmes) puissent être prouvées par l'application de règles de déduction.

La sémantique du modèle théorique²⁵⁶ permet de spécifier *le modèle* à travers *l'interprétation*. L'interprétation du domaine et des constructeurs OWL est définie dans le cadre de la *sémantique directe*²⁵⁷ du langage OWL. *L'interprétation*, donnée par la relation $I = (\Delta^I, \cdot^I)$ contient le *domaine d'interprétation*, (Δ^I) et une *fonction d'interprétation* (\cdot^I) . La fonction d'interprétation permet de spécifier comment les éléments qui composent une ontologie, *les individus (I), les classes (C) et les rôles (R)*, devraient être interprétés.

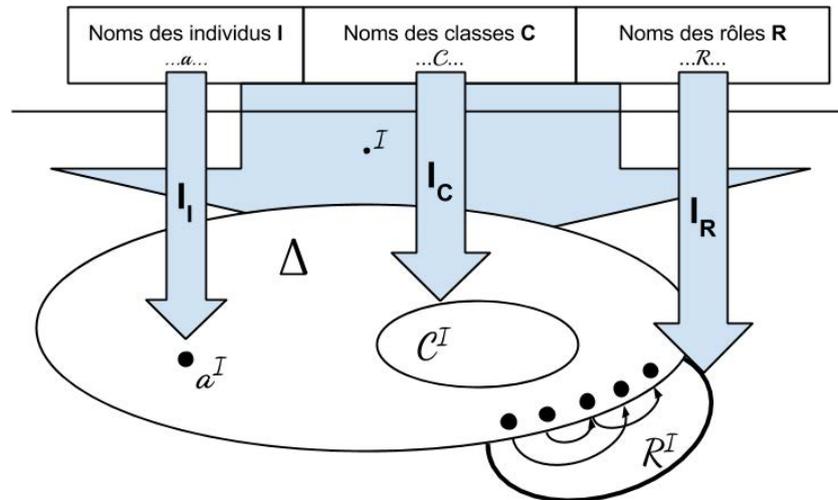


Figure 4-6. Représentation d'une interprétation DL (source: Hitzler *et al.*, 2009, p. 173)

La fonction I_I associe des individus aux éléments du domaine : $I_I: I \rightarrow \Delta$; la fonction I_C associe les noms de classes aux ensembles du domaine : $I_C: C \rightarrow 2^\Delta$ (extension des classes) ; et la fonction I_R permet d'associer les rôles aux relations binaires dans le domaine : $I_R: R \rightarrow 2^{\Delta \times \Delta}$ (extension des propriétés). La fonction *d'interprétation pour le OWL 2* est donnée par $I = (\Delta_I, \Delta_D, \cdot^C, \cdot^{OP}, \cdot^{DP}, \cdot^I, \cdot^{DT}, \cdot^{LT}, \cdot^{FA}, NAMED)$ (Motik, Patel-Schneider, et Grau, 2012), ou :

- Δ_I est un ensemble non-vide appelé *domaine de objets* ;
- Δ_D est un ensemble non-vide disjoint de Δ_I appelé *domaine de données* de telle sorte que $(DT)^{DT} \subseteq \Delta_D$ pour chaque *datatype* $DT \in V_{DT}$ (V_{DT} étant l'ensemble des tous les *datatypes*) ;

²⁵⁶ La sémantique du modèle théorique de la logique propositionnelle spécifie la sémantique des connecteurs logiques à l'aide des tableaux de vérité.

²⁵⁷ <http://www.w3.org/TR/owl2-direct-semantics/>

- \cdot^C est la fonction d'interprétation des classes qui assigne à chaque classe $C \in V_C$ (V_C étant l'ensemble des classes définies dans les spécifications de OWL 2) un sous-ensemble $(C)^C \subseteq \Delta_I$ de telle sorte que
 - $(\text{owl:Thing})^C = \Delta_I$ et
 - $(\text{owl:Nothing})^C = \emptyset$;
- \cdot^{OP} est la fonction d'interprétation des propriétés d'objets qui assigne à chaque propriété $OP \in V_{OP}$ (V_{OP} étant l'ensemble des *propriétés d'objets* définies dans les spécifications de OWL 2) un sous-ensemble $(OP)^{OP} \subseteq \Delta_I \times \Delta_I$ de telle sorte que
 - $(\text{owl:topObjectProperty})^{OP} = \Delta_I \times \Delta_I$ et
 - $(\text{owl:bottomObjectProperty})^{OP} = \emptyset$;
- \cdot^{DP} est la fonction d'interprétation des *propriétés de données* qui assigne à chaque propriété de donnée $DP \in V_{DP}$ (V_{DP} étant l'ensemble des propriétés de données définies dans les spécifications de OWL 2) un sous-ensemble $(DP)^{DP} \subseteq \Delta_I \times \Delta_D$ de telle sorte que
 - $(\text{owl:topDataProperty})^{DP} = \Delta_I \times \Delta_D$ et
 - $(\text{owl:bottomDataProperty})^{DP} = \emptyset$;
- \cdot^I est la fonction d'interprétation de l'individu, qui assigne à chaque individu $a \in V_I$ (V_I étant l'ensemble des individus définis dans les spécifications de OWL 2) un élément $(a)^I \in \Delta_I$;
- \cdot^{DT} est la fonction d'interprétation de type de donnée qui assigne à chaque type de donnée $DT \in V_{DT}$ (V_{DT} étant l'ensemble des types de données contenant `rdfs:Literal`, ainsi que d'autres types de données) un sous-ensemble $(DT)^{DT} \subseteq \Delta_D$ de telle sorte que
 - \cdot^{DT} is the same as in D for each datatype $DT \in N_{DT}$ (N_{DT} étant un ensemble de types de données, sauf `rdfs:Literal`), et
 - $(\text{rdfs:Literal})^{DT} = \Delta_D$;
- \cdot^{LT} est la fonction d'interprétation du *littéral*, définie comme $(lt)^{LT} = (LV, DT)^{LS}$ pour chaque $lt \in V_{LT}$ (V_{LT} étant l'ensemble des littéraux $LV \wedge DT$, LV étant la forme lexicale de lt et DT le type de donnée de lt ;

- \cdot^{FA} est la fonction d'interprétation des facettes, définie comme $(F, It)^{FA} = (F, (It)^{LT})^{FS}$ pour chaque $(F, It) \in V_{FA}$ (V_{FA} étant l'ensemble (F, It) pour chaque facette contraignante F , type de donnée DT et littéral It);
- $NAMED$ est un sous-ensemble de Δ_I de telle sorte que $(a)^I \in NAMED$ pour chaque individu nommé $a \in V_I$.

On trouvera plus d'informations sur la sémantique de OWL et sur les fonctions d'interprétation à l'adresse <http://www.w3.org/TR/2012/REC-owl2-direct-semantics-20121211/>.

4.1.5. Règles et inférences dans le Web sémantique

Dans le cadre du Web sémantique, les moyens de représentation de connaissances se complètent avec les *règles*. Dans le sens général, *une règle* est une « déclaration qui affirme qu'une certaine *conclusion* doit être valide lorsqu'une certaine *prémisse* est satisfaite » (Hitzler *et al.*, 2009, p. 213). Une règle peut être toute affirmation de type « si *prémisse* alors *conclusion* ». Elle permet d'étendre une base de connaissances avec des nouvelles affirmations déduites en conformité avec ses prémisses et conclusions. Un exemple de règle écrite en syntaxe de la logique de premier ordre est la suivante :

$$Person(x) \wedge isPainterOf(x, y) \wedge Painting(y) \rightarrow Painter(x)$$

Cela spécifié que « si x est une *Person* et x *isPainterOf* y et y est un *Painting* alors x est un *Painter* »²⁵⁸. En syntaxe DL cette règle peut être écrite de la façon suivante²⁵⁹ :

$$Person \sqcap \exists isPainterOf. Painting \sqsubseteq Painter$$

Pour permettre aux systèmes utilisant des règles, sur le Web en général et dans le cadre du Web sémantique en particulier, d'échanger leurs règles entre eux, le W3C propose le standard *Rule Interchange Format*²⁶⁰ (*RIF*), qui devient une recommandation depuis juin 2010. RIF met l'accent sur l'échange plutôt que sur la définition d'un seul langage pour

²⁵⁸ Dans un langage courant la règle pourrait s'écrire sous la forme suivante : « Si une personne a peint un tableau, alors cette personne est un peintre ».

²⁵⁹ La même règle en utilisant des termes RDF en syntaxe des règles Jena peut être écrite de la façon suivante :

```
@prefix rdf:      <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix ont:      <http://example.org/ontologie#> .
```

```
[rule_painter:(?x rdf:type ont:Person), (?x ont:isPainterOf ?y), (?y
rdf:type ont:Painting) -> (?x rdf:type ont:Painter)]
```

²⁶⁰ RIF n'est pas seulement un format, mais aussi un medium d'échange entre les différents systèmes de règles et un format d'échange de données XML (Kifer et Boley, 2013).

tout type de règle. « Le RIF se présente sous la forme d'un ensemble extensible de *dialectes* de règles dont trois sont définis dans la recommandation » (Gandon *et al.*, 2012, p. 134) : *RIF Core Dialect (RIF-Core)*, *RIF Basic Logic Dialect (RIF-BLD)* et *RIF Production Rule Dialect (RIF-PRD)* (Kifer et Boley, 2013). *RIF Core* est le langage élémentaire des règles, le noyau de primitives commun aux dialectes RIF. *RIF-BLD* correspond à la *logique de Horn*²⁶¹ avec égalité. *RIF-BLD* permet de représenter les *règles déclaratives*, « des programmes logiques, c'est-à-dire des suites de règles d'inférence, sur des faits positifs (donc sans négation) ». *RIF-PRD* permet de représenter les *règles de production*, « des règles dont l'application déclenche des actions d'ajout, de modification ou de suppression de faits dans la base, dont le résultat dépend de l'ordre d'application des règles ». (Gandon *et al.*, 2012, p. 134). Formatée pour intégration avec *OWL-RIF*, la règle énoncée ci-dessus peut être représentée de la façon suivante :

```
Forall ?x ?y (
  ?y[rdf:type ont:Painter] :-
  And (?x[ont:isPainterOf ?y] ?y[rdf:type ont:Painting])
)
```

Les règles, la sémantique formelle ou les logiques de description utilisées dans le cadre du Web sémantique permettent le développement des systèmes « intelligents ». L'intelligence de ces systèmes à base de connaissances consiste dans la capacité de *moteurs d'inférences* à trouver les conséquences implicites de connaissances explicitement représentées. Dans la *logique propositionnelle*²⁶², sur laquelle se base la logique de premier ordre et donc les logiques de description, la règle primitive d'inférence est l'*implication-élimination*, appelée *modus ponens*²⁶³.

²⁶¹ La *logique* ou la *clause Horn* est « un cas particulier où l'inférence peut se réaliser beaucoup plus ciblée que dans le cas de la résolution générale ». Une *clause Horn* est une clause comportant au maximum un *littéral positif* (e.g. un sous-ensemble des formes normales disjonctives, $(\neg p_1 \vee \neg p_2 \vee \dots \vee \neg p_n) \vee h$) (« Horn clause logic », s. d.).

²⁶² La *logique propositionnelle* ou *calcul des propositions* considère les *phrases* comme des éléments de base qui peuvent prendre une valeur booléenne. Les phrases peuvent être ensuite reliées par des opérateurs logiques (conjonction, disjonction, négation, implication conditionnelle, implication biconditionnelle).

²⁶³ *Modus Ponens*, du latin *Modus Ponendo Ponens*, signifie *mode qui affirme par l'affirmation* et consiste à affirmer une implication (e.g. A implique B). La contraposition est le *Modus Tollens*, du latin *Modus Tollendo Tollens* signifiant *le mode qui nie par la négation* (e.g. non-B implique non-A).

« Les systèmes DL utilisent les *logiques de description* pour représenter la connaissance d'un domaine et offre des services d'*inférence* basés sur la *sémantique formelle* des logiques de description pour dériver la connaissance *implicite* de faits représentés *explicitement* » (Tobies, 2001, p. 2). Sur le Web, les faits sont enregistrés de façon distribuée. Cela conduit à la répartition des connaissances sur le Web et donc à la distribution des systèmes d'inférences (voir les *TriplesStore*). Cela impose aux systèmes d'inférences dans le cadre du Web sémantique de respecter *l'assomption de monde ouvert*²⁶⁴ (*Open World Assumption – OWA*). Cette assomption affirme que la *valeur de la vérité* est indépendante de ce qui est connu par un système spécifique. Autrement dit, les faits qui ne sont pas présents dans la base de connaissances ne sont pas considérés comme faux. Toutefois, pour être utile dans les applications, un système DL doit assurer que les inférences sont (ibid., p. 3) :

- *cohérentes (sound)* – toutes les conclusions doivent être correctes, valides ;
- *complètes* – toutes les conclusions doivent être rendues.

Les inférences typiques pour une base de connaissances en OWL peuvent consister en calculs concernant : la subsomption (e.g. $C \sqsubseteq D$), l'équivalence des classes (e.g. $C \equiv D$), la disjonction des classes (e.g. $C \sqcap D \sqsubseteq \perp$), la cohérence des classes (e.g. $C \sqsubseteq \perp$), la vérification des instances (e.g. si l'instance a appartient à une classe C , $C(a) ?$) ou la récupération des instances (e.g. trouver toutes les instances appartenant à une classe C , tous les a ayant $C(a)$) et la *cohérence globale* de la base de connaissances (Hitzler *et al.*, 2009, p. 183). La *cohérence globale* de la base de connaissances est donnée par l'absence des contradictions. Ces calculs sont effectués par des *moteurs d'inférences*, appelés encore *raisonneurs*, ou *raisonneurs sémantiques*, en appliquant des *algorithmes de preuve* sur la base de connaissances. Un de ces algorithmes utilisés pour déterminer si une base de connaissances est cohérente est *l'algorithme tableau*. L'algorithme tableau construit un arbre en appliquant des règles d'expansion résultant de la simplification de la formule logique. Deux types de règles s'appliquent dans la construction de l'arbre : les α -règles dans le cas des *conjonctions* et les β -règles dans le cas de *disjonctions*. Selon la règle appliquée, l'arbre fait pousser ses branches (α -règles) ou donne naissance à une

²⁶⁴ À l'opposé de *l'assomption de monde ouvert* se trouve *l'assomption de monde fermé* (*Closed World Assumption – CWA*), qui affirme que toute affirmation qui n'est pas connue pour être vraie, est fautive (e.g. si une base de connaissances contient seulement l'affirmation « Sandro Botticelli est un peintre », à la question « Léonard de Vinci est-il un peintre ? » les systèmes d'inférences basés sur le OWA répondent *inconnu*, alors que les systèmes CWA répondent *non*).

autre branche (*β -règles*). Les règles sont appliquées jusqu'à la simplification des formules logiques en ensembles d'axiomes. Les nœuds sont des ensembles de formules et les feuilles sont des axiomes. L'existence d'une contradiction sur une branche rend la base de connaissances incohérente. L'algorithme tableau peut être vu comme une « spécialisation du calcul du tableau pour la logique du prédicat de premier ordre » (Baader et Sattler, 2001, p. 7).

4.2. Ingénierie ontologique

Définir une ontologie pour la représentation des connaissances c'est « définir, pour un domaine et un problème donnés, la signature fonctionnelle et relationnelle d'un langage formel de représentation et la sémantique associée » (Bachimont, 2000). Bachimont distingue trois niveaux de l'ontologie :

1. L'ontologie régionale représente le niveau sémantique ou interprétatif. Les *concepts sémantiques* construisant l'arbre de l'ontologie sont caractérisés par des libellés linguistiques. Ces libellés sont empruntés à la langue du domaine. Le libellé doit avoir « un sens univoque et non contextuel » pour être utilisé comme « primitive de représentation ».
2. L'ontologie référentielle représente « le niveau formel ou référentiel », elle peut être vue comme un « treillis de concepts formels ». Chacun de ces concepts formels est caractérisé par un libellé dont la sémantique est définie par une « extension d'objets » : « chaque concept formel est défini par un engagement ontologique qui spécifie quels objets doivent exister dans le domaine pour utiliser le concept conformément à sa signification formelle.
3. L'ontologie computationnelle représente le niveau « opérationnel ou computationnel », elle est un « treillis de concepts computationnels ». Ces concepts sont caractérisés par les « opérations qu'il est possible de leur appliquer : ces opérations leur confèrent une sémantique dans le système effectif construit. »

L'ingénierie ontologique fait référence à « l'ensemble des activités qui concernent le processus de développement d'ontologies, le cycle de vie d'une ontologie, les méthodes et les méthodologies pour la construction d'ontologies, ainsi que les suites d'outils et des langages qui les soutiennent » (Gomez-Pérez, Fernandez-Lopez, et Corcho, 2003).

Dans l'article de Fernandez-Lopez *et al.* (Fernández-López, Gómez-Pérez, et Juristo, 1997) les activités d'élaboration d'une ontologie sont structurées en :

1. activités de gestion des ontologies (e.g. calendrier, contrôle, assurance de la qualité) ;
2. activités orientées vers le développement des ontologies (e.g. études d'environnement, spécification, conceptualisation, formalisation, implémentation, maintenance, réutilisation) ;
3. activités d'assistance (e.g. évaluation, intégration, alignement, orchestration, documentation).

Dans la suite, nous allons présenter les étapes principales de l'élaboration d'une ontologie.

4.2.1. Conception et construction d'ontologies

La conception et la construction des ontologies s'inscrivent dans les pratiques de conception des bases de connaissances, de modélisation de la connaissance. En même temps, comme dans la conception des logiciels informatiques, la construction des ontologies est censée intégrer des méthodes et techniques utilisées dans l'ingénierie des logiciels²⁶⁵. N. Noy et D. McGuinness présentent une méthode de développement des ontologies en 7 étapes. Rappelant qu'« il n'y a pas une seule manière ou méthodologie *correcte* pour développer une ontologie », les auteurs précisent que le développement des ontologies est *nécessairement* un processus *itératif* (Noy et McGuinness, 2001). Les sept étapes de construction d'une ontologie sont les suivantes :

1. la définition du domaine et de la portée de l'ontologie (e.g. but d'ontologie, question de compétences d'ontologie) ;
2. la considération d'une éventuelle réutilisation des ontologies existantes ;
3. l'énumération des termes importants dans l'ontologie (e.g. concepts, propriétés) ;
4. la définition des classes et de la hiérarchie des classes (e.g. approches de haut en bas – les plus générales en premier ; de bas en haut – les plus spécifiques en premier ; combiné – les plus proéminents ou importants en premier) ;
5. la définition des propriétés des classes et des attributs ;

²⁶⁵ Gestion de cycle de vie, maintenance, versionnement.

6. la définition des contraintes des propriétés (e.g. restrictions de cardinalité) ;
7. la création d'instances.

Les termes de l'ontologie devraient être *proches* des objets et des relations existant dans le domaine d'intérêt : *substantifs* pour nommer les classes ou les objets, et *verbes* pour nommer les relations.

Le problème de réutilisabilité des ontologies, l'étape 2 dans la méthode présentée, est formulé dans (Bachimont, 2000). L'hypothèse proposée est la suivante : « une ontologie peut être générique pour un ensemble de tâches données si elle possède un niveau de finesse de description permettant aux point de vue de chaque tâche de se représenter. Mais il n'existe pas de niveau de description permettant de décrire toutes les tâches du domaine. Il n'y a donc pas de généralité universelle ou absolue, mais de généralité relative à des tâches ».

Les problèmes de correspondance des ontologies sont relevés aussi dans (Calabretto, 2011) dans le cadre de l'alignement des ontologies utilisées dans le domaine urbain. Ainsi les problèmes d'alignement des ontologies renvoient à l'hétérogénéité des ontologies, qui peut être :

- syntaxique – les ontologies ne sont pas exprimées dans le même langage ;
- terminologique – l'utilisation des synonymes, variations des noms attribués ;
- conceptuelle (hétérogénéité sémantique, inadéquation logique) – différence dans la modélisation du même domaine de discours :
 - o différences au niveau de la couverture de l'ontologie ;
 - o différences au niveau de la granularité des concepts ;
 - o différences en perspective ;
- sémiotique (hétérogénéité pragmatique) – comment les entités sont interprétées par les humains.

Dans la construction des ontologies du SASA nous considérons des équivalences avec des termes définis dans d'autres ontologies sur le Web. Dans le § 4.2.4 nous présentons quelques exemples d'ontologies de *domaine* utilisées sur le Web.

4.2.2. Types d'ontologies

Selon leur contenu et l'utilisation des ontologies dans les bases de connaissances, plusieurs types se distinguent (Guarino, 1997) :

4. ontologies de haut niveau : elles représentent des concepts très généraux, indépendamment du contexte d'utilisation (e.g. temps, espace, état, événement) ;
5. ontologies de domaine – elles comprennent les concepts fondamentaux spécifiques au domaine modélisé ou au discours (e.g. électronique, médical) ;
6. ontologies de tâche – comprennent les termes spécifiques à une activité ou à un tâche ;
7. ontologie d'application – concentrée sur une tâche et un domaine spécifique.

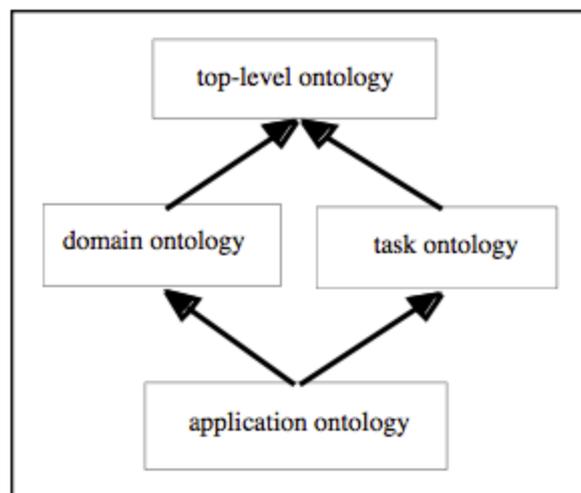


Figure 4-7. Types d'ontologies (source : ibid.)

D'autres classifications (Fensel, 2001) intègrent les *ontologies métadonnées* (e.g. Dublin Core), ou encore les *ontologies représentationnelles* (e.g. Frame Ontology²⁶⁶). Les ontologies métadonnées fournissent un *vocabulaire* pour décrire le contenu des ressources d'information en ligne, alors que les ontologies représentationnelles proposent les termes généraux nécessaires à la représentation des entités, sans spécifier l'objet de la représentation. O. Lassila et D. McGuinness (Lassila et McGuinness, 2001) présentent une classification en fonction du niveau d'expressivité des ontologies (v. Figure 4-8).

²⁶⁶ Le *Frame Ontology* définit des termes afin de fournir des conventions utilisées dans les systèmes de représentation de connaissances centrés-objet (<http://www.aii.ed.ac.uk/project/enterprise/enterprise/ontology-code/frame-ontology/index.html>)

Un *vocabulaire contrôlé* consiste en une liste finie de termes (catalogue), alors que le *glossaire* ajoute une définition informelle au vocabulaire. Le thésaurus (traduction latine du grec *thesauros*, *trésor*) est un vocabulaire contrôlé, les concepts étant connectés à travers des relations d'équivalence (synonymes), de hiérarchie (sous-classe) ou d'autres associations (concepts similaires, sujets connexes).

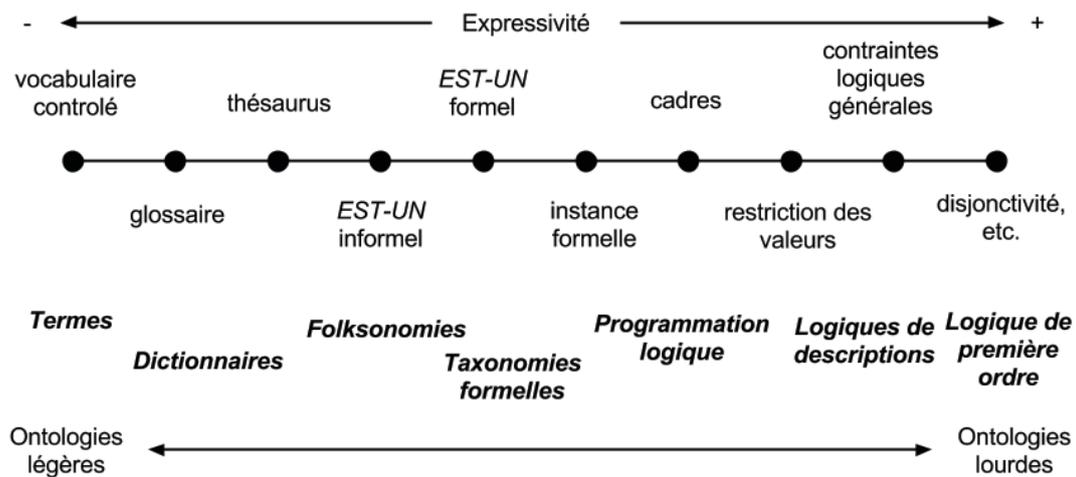


Figure 4-8. Spectre ontologique (source : *ibid.*; Harald Sack, 2013)

Les hiérarchies formelles (hiérarchie explicite, relations strictes de classes) ou informelles (hiérarchie explicite, relations non-strictes de classes) et les instances formelles (hiérarchie explicite, relations d'intanciement) permettent de définir des taxonomies et folksonomies. Ensuite, le niveau d'expressivité de l'ontologie est associé au langage de logique sur lequel s'appuie l'ontologie.

4.2.3. Approche modulaire

Dans la pratique, une base de connaissances contient des informations sur des sujets appartenant à plusieurs domaines d'intérêt. Une des décisions concernant l'intégration des ontologies dans les bases de connaissances concerne l'aspect de *modularisation des ontologies*. Il s'agit alors d'étudier comment les grandes ontologies peuvent être structurées en parties plus petites, appelées *modules*. La modularisation des ontologies est une « stratégie souhaitable » (Hitzler *et al.*, 2009, p. 327) pour plusieurs raisons. En effet, la facilitation des tâches de maintenance, la possibilité d'utiliser le module dans d'autres contextes ou l'efficacité des moteurs d'inférences quand ils travaillent sur une taille et complexité réduite des données, sont autant d'incitations à aborder une

approche modulaire dans la structuration des ontologies. OWL et OWL 2 permet l'inclusion d'autres ontologies — qui peuvent être situées ailleurs sur le Web — en utilisant le terme `owl:import`.

Le système d'apprentissage que nous proposons présente *six modules* correspondant aux six ontologies groupées selon leur contenu. Plus de détails sont présentés dans le Chapitre 5.

4.2.4. Exemples d'ontologies et vocabulaires utilisés sur le Web

Plusieurs vocabulaires et ontologies existent pour décrire la variété des ressources existant sur le Web. En juin 2011 une initiative commune de Bing, Google et Yahoo met en place le *Schema.org*²⁶⁷, dans le but de fournir un ensemble commun de schémas pour l'annotation des données sur le Web. L'objectif est d'améliorer la visibilité des contenus des pages web dans les moteurs de recherche. L'utilisation de ces vocabulaires dans les pages Web permet une annotation spécifique résultant dans une identification plus pertinente dans les recherches. Ainsi les moteurs de recherche peuvent afficher des *extraits riches*²⁶⁸ d'information sur du contenu ciblé. L'intégration des termes dans les pages Web est réalisée en utilisant *Microdata* ou *RDFa*²⁶⁹. Schema.org regroupe plusieurs vocabulaires définis dans des ontologies populaires sur le Web.

Nous présentons ensuite les principales ontologies utilisées pour décrire les ressources les plus souvent représentées sur le Web²⁷⁰. Les ontologies les plus populaires sur le Web sont celles qui décrivent les entités (e.g. personnes, documents, lieux, événements) les plus présentes sur le Web. Notons l'ontologie *Friend Of A Friend (FOAF)*²⁷¹ pour

²⁶⁷ <http://schema.org>

²⁶⁸ Un des exemples les plus pertinent pour l'intégration des annotations sémantiques dans les documents Web est le projet SearchMonkey de Yahoo ! (<http://developer.yahoo.com/searchmonkey/siteowner.html>).

²⁶⁹ RDFa – Resource Description Framework in Attributes est une recommandation de W3C qui permet d'ajouter une extension d'attribues au langage HTML afin d'intégrer des métadonnées dans le documents Web.

²⁷⁰ <http://lov.okfn.org/dataset/lov/>

²⁷¹ <http://xmlns.com/foaf/spec/>

décrire les personnes, leurs activités et leurs relations avec d'autres personnes²⁷² ou objets. L'ontologie vCard²⁷³ fournit les termes nécessaires pour décrire les personnes et les organisations. DublinCore est le vocabulaire le plus utilisé pour décrire des documents (e.g. vidéos, images, pages Web, livres) sur le Web. L'ontologie SKOS²⁷⁴ (*Simple Knowledge Organization System*) fournit des termes utilisés dans des systèmes d'organisation de la connaissance (e.g. thésaurus, taxonomies). SIOC²⁷⁵ (*Semantically Interlinked Online Communities*) est une ontologie qui permet de décrire les communautés en ligne (e.g. blogs, wiki, forums).

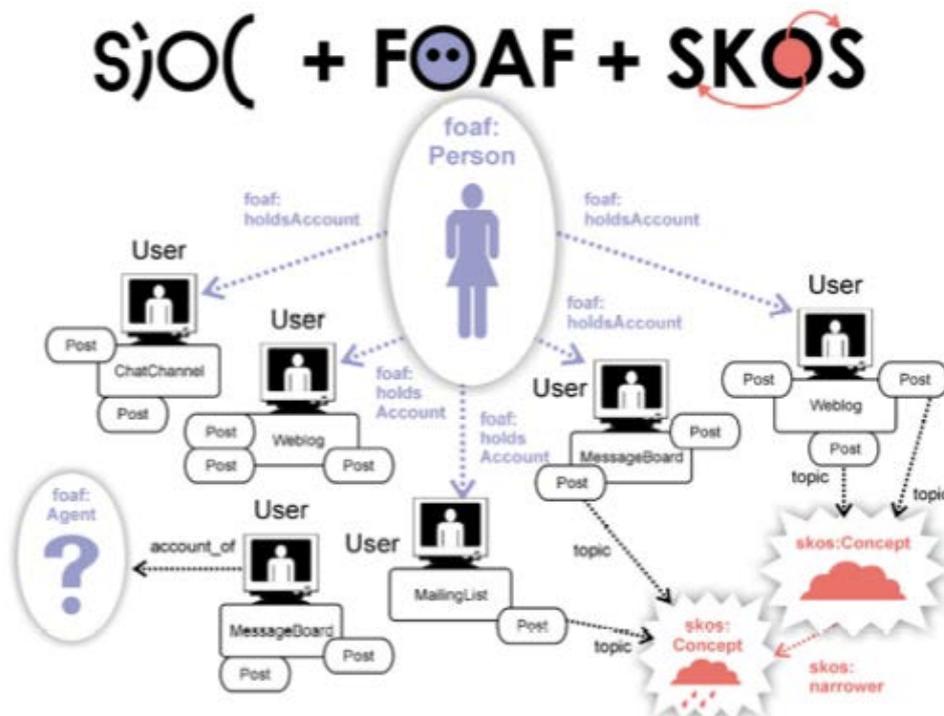


Figure 4-9. L'articulation entre les ontologies FOAF, SKOS et SIOC (source : Breslin *et al.*, 2009, p. 199)

Cette ontologie permet de décrire la structure et les activités des communautés en ligne. Articulée avec l'ontologie FOAF et l'ontologie SKOS, cette ontologie permet de réduire l'écart entre les discussions et les sujets appartenant aux différentes communautés

²⁷² FOAF spécifie les relations entre les personnes en utilisant la propriété `foaf:knows`. Cette relation étant souvent générale, l'ontologie *Relationship* (<http://vocab.org/relationship/html>) fournit des propriétés plus spécifiques pour décrire les relations entre les personnes.

²⁷³ <http://www.w3.org/TR/vcard-rdf/>

²⁷⁴ <http://www.w3.org/2004/02/skos/>

²⁷⁵ <http://sioc-project.org/>

(ibid.). Dans ce contexte s'inscrivent les ontologies décrivant *le tag*, comme le MOAT (Meaning Of A Tag) ou le NiceTag²⁷⁶.

Ces ontologies permettent l'intégration du contenu généré par l'utilisateur (Web social) dans des applications sémantiques. La description du contenu généré par les utilisateurs permet d'améliorer l'identification du contenu pertinent dans les scénarios d'apprentissage collaboratif.

4.3. Le Web sémantique pour les systèmes d'apprentissage

Les possibilités de représentation de l'information et de modélisation de la connaissance dans le cadre du Web, suscitent discussions et intérêt dans la communauté pour l'intégration de technologies du Web sémantique dans plusieurs domaines d'application. C. Nicolle et C. Cruz constatent la prolifération des parties impliquées dans les processus de conception, de construction et d'entretien des bâtiments, ainsi que l'hétérogénéité des outils utilisées pour la modélisation de ces processus (Nicolle et Cruz, 2011). Ils proposent une approche basée sur les technologies du Web sémantique pour l'administration des bâtiments. Utilisant cette approche ils résolvent les problèmes de l'évolution du modèle, de mappage de données, de la gestion de données temporelles et de *l'adaptation des données en fonction de l'utilisation et de l'utilisateur*.

D'autant plus, dans les systèmes d'apprentissage, les effets de l'intégration des technologies et outils du Web sémantique devraient se retrouver dans une adaptation et personnalisation du processus d'apprentissage.

D. Sampson *et al.* présentent une synthèse des travaux qui présentent les problématiques et les principales questions posées (D. G. Sampson, Lytras, et Diaz, 2004). Deux éléments clés se distinguent dans la construction des systèmes d'apprentissage en utilisant les technologies sémantiques. Le premier d'entre eux concerne *la modélisation* ou la représentation des contenus d'apprentissage, généralement réalisée à travers des « *taxonomies formelles* », construites en utilisant RDFS et OWL. Le deuxième point concerne la construction dynamique et personnalisée

²⁷⁶ <http://ns.inria.fr/nicetag/2010/09/09/voc.html>

des contenus d'apprentissage. Cette construction dynamique est possible grâce aux moteurs d'inférence et aux langages de règles du Web sémantique. Reprenant les deux contributions clés identifiées plus haut, nous nous intéressons donc à l'application des technologies du Web sémantique pour : (1) la modélisation des entités participant à l'apprentissage et (2) la construction et l'adaptation dynamique du parcours d'apprentissage.

Pour (Aroyo et Dicheva, 2004), une des questions clés est l'interopérabilité entre les systèmes d'apprentissage, la conceptualisation sémantique et les ontologies, ainsi que le support automatisé, structuré et unifié pour la création des contenus d'apprentissage. S. Yang *et al.* (Yang, Chen, et Shao, 2004) proposent deux modèles de métadonnées : un modèle pour le contenu d'apprentissage et un autre pour les annotations sur le contenu. Ces annotations sur le contenu permettent ensuite la personnalisation du contenu d'apprentissage.

Henze *et al.* (Henze, Dolog, et Nejdl, 2004) proposent un cadre pour l'apprentissage en ligne personnalisé dans le Web sémantique. Ils proposent une approche basée sur l'utilisation des règles logiques et des moteurs d'inférences pour la personnalisation du contenu d'apprentissage.

Alila *et al.* (Soualah-Alila, Mendes, et Nicolle, 2013) présentent un système adaptatif pour les dispositifs portables basé sur la modélisation sémantique du contenu et du contexte d'apprentissage. Ils modélisent quatre dimensions du contexte d'apprentissage : (1) la dimension temporelle, (2) la dimension spatiale, (3) la dimension du dispositif et (4) la dimension de l'apprenant. L'adaptation du contenu d'apprentissage tenant compte du temps, sa location, son style d'apprentissage et le cours qu'il étudie, le système lui fournit une liste des suggestions adaptées. Ils utilisent le langage SWRL pour déclarer les règles d'adaptation.

Grandbastien *et al.* présentent l'ontologie *s-lom* (Semantic LOM), fondée sur LOM, utilisée dans le cadre du projet LUISA. Ce projet intègre quatre ontologies : (1) l'ontologie *s-lom* (fondée sur LOM) ; (2) l'ontologie de compétences GCS (General

Competency Schema) ; (3) l'ontologie de disciplines ; et (3) l'ontologie de logiciels (Grandbastien, Huynh-Kim-Bang, et Monceaux, 2009).

Toujours pour la description des objets pédagogiques, au sein de notre équipe, R. Balog-Crişan et I. Roxin proposent l'outil *RDF4LOM* pour la représentation RDF des métadonnées LOM (Radu Balog-Crişan et Roxin, 2007).

S. Dehors exploite les technologies du Web sémantique et du management des connaissances pour une analyse sémantique des parcours d'apprentissage (Dehors, 2007). Il définit le QBLS (Question Based Learning Strategy) afin de guider l'apprenant dans son parcours d'apprentissage. Il utilise les technologies sémantiques pour améliorer la navigation de l'apprenant dans une *réseau de concepts*.

Vu la taille du Web et les technologies éprouvées déjà existantes, l'adoption de nouveaux standards n'est pas évidente. Les technologies du Web sémantique s'accompagnent de coûts dans le développement et la construction de modélisations et d'ontologies. Le développement d'une ontologie demande un niveau élevé d'expertise de la part de créateur et l'acceptation par la communauté.

A.I. Cristea (Cristea, 2004) analyse comment le Web sémantiques peut améliorer *LAOS*²⁷⁷ (Layered AHS Authoring-Model and Operators) (Cristea et de Mooij, 2003). Cristea propose l'écriture des modèles utilisés en MOT (My Online Teacher), un système hypermédia adaptatif basé sur le cadre fournit par LAOS, en utilisant les technologies du Web sémantique. Elle identifie l'inertie dans l'adoption de ces technologies comme la limite principale de développement des systèmes utilisant les technologies du Web sémantique. Elle conclut que les technologies du Web sémantique peuvent améliorer les systèmes d'apprentissage, mais compte tenu du travail supplémentaire nécessaire pour l'intégration de ces technologies, les décisions sont à prendre au cas par cas.

Sur la base des approches présentées, nous avons retenu les technologies du Web sémantique comme pouvant répondre aux problématiques liées au développement de

²⁷⁷ LAOS est un modèle générique d'un système hypermédia adaptatif. Il contient cinq niveaux : modèle de domaine, modèle des buts et des contraintes, modèle d'utilisateur, modèle d'adaptation et modèle de présentation.

notre prototype de système d'apprentissage. Dans un premier temps, en basant la représentation des entités sur les technologies sémantiques nous permettons à notre système une identification pertinente des ressources d'apprentissage. Dans un second temps, la définition, l'implémentation et l'utilisation de règles logiques par le système nous permet de rendre notre système intelligent. Le système doit être capable d'accompagner et de faciliter le processus d'apprentissage pour l'apprenant tout en améliorant le travail de l'enseignant.

Chapitre 5. Modélisation du Système Actif et Sémantique d'Apprentissage (SASA)

« Il faut essayer de connaître la connaissance, si l'on veut connaître les sources de nos erreurs ou illusions. Or la connaissance est l'objet le plus incertain de la connaissance philosophique et l'objet le moins connu de la connaissance scientifique. Qu'est-ce qu'un cerveau qui peut produire un esprit qui le connaît ? Qu'est-ce qu'un esprit qui peut concevoir un cerveau qui le produit ? Qu'est-ce qu'un esprit/cerveau qui ne saurait penser sans un langage et une culture ? Qu'est-ce qu'une connaissance qui croit refléter la nature des choses alors qu'elle est traduction et construction ? Comment connaître ce qui connaît ? »
E. Morin

La contribution majeure de cette thèse est la conception et la réalisation d'un prototype pour un Système Actif et Sémantique d'Apprentissage (SASA). Un *système d'apprentissage* est « le produit d'un processus d'ingénierie pédagogique et il sert de soutien à la diffusion des apprentissages » (Paquette, 2002a, p. 67). Dans l'approche systémique du processus d'apprentissage, la méthodologie spécifique utilisée est « orientée vers *l'acquisition des objectifs spécifiques* » et elle se « base sur le travail de recherche dans les domaines de l'apprentissage et de la communication humaine » (C. Clark, 2008, p. 87).

Notre SASA peut être vu dans un premier temps comme un *système conceptuel*. Dans un système conceptuel (composé des *concepts*), l'accent est mis sur « la description des objets qui composent » le système et sur « les liens qui unissent » ces objets (Paquette, 2002b, p. 96). Ainsi, *l'ensemble des éléments* du SASA est décrit en utilisant les technologies du Web sémantique. Notre objectif est de concevoir un système *actif d'apprentissage*. Comme évoqué dans le § 2.3.1, le terme de *système actif* est proposé dans le rapport avec *l'apprenant*. La pro-activité (anticipation du parcours d'apprentissage) ou la réactivité du système (adaptation du parcours d'apprentissage)

permettent l'organisation dynamique des éléments du système ayant un effet direct sur le processus d'apprentissage de l'apprenant.

Ce système d'apprentissage se concrétise dans une *modélisation* de l'espace d'apprentissage : l'apprenant, son environnement d'apprentissage (e.g. objectif d'apprentissage, objet pédagogique) et ses interactions dans son environnement d'apprentissage. La modélisation du SASA est implémentée et exploitée dans le cadre d'un système informatique basé sur les technologies du Web sémantique.

Dans la première partie de ce chapitre nous allons identifier et modéliser les entités qui composent l'espace du SASA (e.g. l'apprenant, l'objectif d'apprentissage, l'objet pédagogique, la compétence), leurs caractéristiques et relations, et ensuite nous allons procéder à la construction des ontologies correspondantes.

Afin d'être interprétables et exploitables par le système informatique, les modélisations et les relations entre entités sont enregistrées en utilisant les technologies du Web sémantique. Dans la deuxième partie de ce chapitre nous allons présenter les ontologies résultant de cette modélisation et leur harmonisation dans le cadre du système d'apprentissage. Cet ensemble d'ontologies constitue la *spécification de la conceptualisation* du SASA.

5.1. Identification des objets dans le cadre du système d'apprentissage

Dans le domaine de la psychopédagogie le concept *d'objet* a un sens large (Raynal et Rieunier, 2010, p. 322). Ainsi nous identifions trois catégories d'objets :

1. les objets réels inanimés, qui ont le sens commun d'objet (e.g. le livre, l'ordinateur, la machine) ;
2. les objets réels vivants, les être animés (e.g. l'apprenant, le professeur) ;
3. les objets symboliques (e.g. des dessins, des schémas, des mots, des concepts, des théories).

Dans le système d'apprentissage, les objets du monde réel sont représentés par des objets virtuels. Ces objets virtuels portent le contenu, les caractéristiques et les propriétés d'objets réels qu'ils représentent.

En partant du triangle pédagogique de J.Houssay, nous identifions trois objets élémentaires de l'apprentissage : l'apprenant, l'objet pédagogique et l'enseignant. Dans un regard plus proche, autour de ces objets il existe d'autres objets qui permettent à ces trois objets de se rapporter sur différentes dimensions (v. Figure 5-1). Prenons l'*objectif d'apprentissage*, qui peut caractériser à la fois l'*apprenant* et l'*objet pédagogique*.

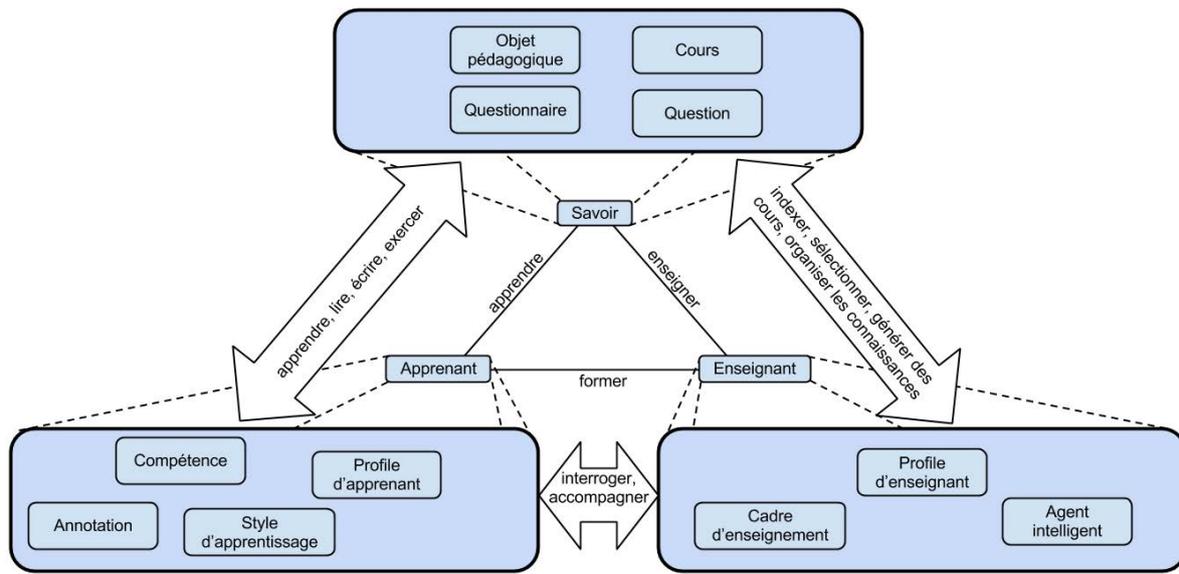


Figure 5-1. Objets dans le système d'apprentissage

Selon l'approche systémique, l'organisation des éléments se réalise autour de l'*objectif du système*. L'objectif du *système d'apprentissage* est hérité de l'objectif de l'apprentissage en général : produire des *effets d'apprentissage* (e.g. *changement du comportement, acquisition de savoir, acquisition de compétences*) chez l'apprenant. Les éléments du système sont organisés autour de cet objectif. Selon C. Clark (2008, p. 87), les « exigences minimales pour un système d'apprentissage sont : (1) l'*apprenant*, (2) l'*objectif d'apprentissage* et (3) les *instructions* pour atteindre l'objectif d'apprentissage ». Le premier objet à modéliser, présent aussi dans l'objectif du système, est donc l'*apprenant*.

Les *objectifs d'apprentissage* sont définis en relation avec l'apprenant. L'apprenant suit des objectifs spécifiques d'apprentissage qui dirigent son processus et parcours d'apprentissage. Ces objectifs d'apprentissage consistent généralement en l'acquisition

des *compétences, des connaissances, des savoir*. Un deuxième objet à modéliser est donc constitué par *l'objectif d'apprentissage* de l'apprenant.

Ensuite, toujours en relation avec l'apprenant et l'objectif d'apprentissage, *l'objet pédagogique* est l'élément qui fournit le contenu et les instructions d'apprentissage. L'objet pédagogique est défini par sa relation avec l'objectif d'apprentissage.

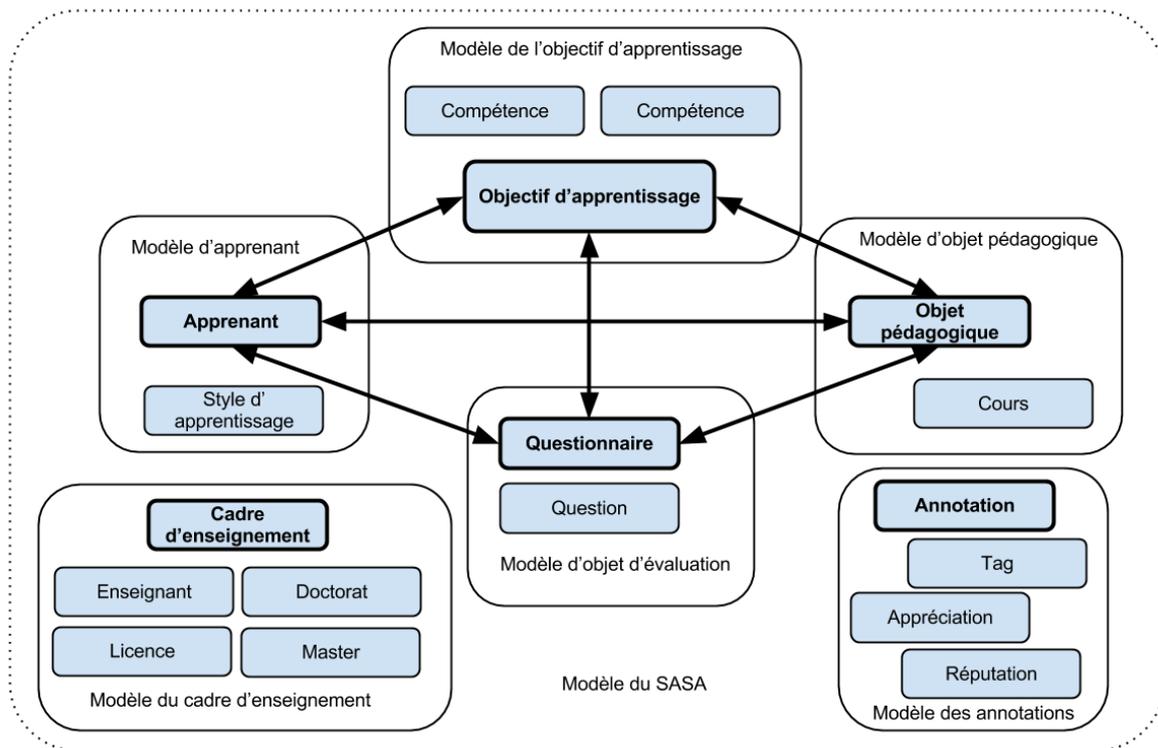


Figure 5-2. Objets dans le modèle du SASA

En relation avec ces trois objets, il en existe un quatrième qui renvoie aux mécanismes d'autorégulation des systèmes et rend possible la mesure du progrès de l'apprenant. C'est *l'objet d'évaluation* qui est en relation avec l'apprenant (évaluer QUI ?), avec l'objectif d'apprentissage (évaluer QUOI ?) et avec l'objet pédagogique (relation entre les contenus de ces objets). L'illustration de ces relations est présentée dans la Figure 5-2.

A ces quatre éléments s'ajoute une autre catégorie d'objets qui fournit le cadre d'enseignement et permet l'administration et la classification des ressources d'apprentissage selon les règles utilisées dans la vie réel (e.g. système licence – master – doctorat, spécialisation). Cette catégorie regroupe les *objets d'annotation*, qui permettent l'association de *tags*, de *notes* ou d'*appréciations* aux objets participant à

l'apprentissage (e.g. objet pédagogique, objet d'évaluation, enseignant), et ce pour améliorer leur visibilité dans le cadre du système.

Nous distinguons donc *six objets* centraux dans notre modélisation : *l'apprenant, l'objet pédagogique, l'objectif d'apprentissage, l'objet d'évaluation, l'objet d'annotation et le cadre d'enseignement*. D'autres objets en relation avec ces objets sont décrits dans les sous-paragraphes correspondants.

5.2. Modélisation des objets participant à l'apprentissage

Après l'identification des objets et des domaines à modéliser, nous regardons maintenant, plus en détail, chaque domaine considéré, les objets spécifiques et les relations existant entre ces objets. Les relations entre les différentes entités sont marquées dans les illustrations présentées pour chaque modélisation. La sémantique des classes et des relations, ainsi que les éventuelles contraintes définies dans ces modèles sont décrites dans l'ontologie correspondant à chaque modèle.

5.2.1. Modélisation de l'apprenant

Le modèle de l'apprenant est une description abstraite qui peut être conçue par l'enseignant ou, dans notre cas, réalisée par un agent logiciel selon une spécification définie au préalable. Dans les systèmes de gestion de l'apprentissage, le concept de *profile d'apprenant* ou de *modèle d'apprenant* est utilisé pour représenter les *caractéristiques* d'un apprenant (Brusilovsky, 2001; Rueda *et al.*, 2003). Le modèle de l'apprenant est une « *représentation* des connaissances de l'apprenant *interprétée* par un agent logiciel » (Holt *et al.*, 1994, p. 4), une « description globale de l'utilisateur maintenue par une application (le système) pour juger le *niveau d'expérience d'utilisateur, la compréhension des contenus du domaine, la maîtrise des tâches, des préférences, des attitudes ou des styles cognitifs* » (Vassileva, McCalla, et Greer, 2003, p. 179).

Dans le domaine de l'Intelligence Artificielle le *modèle d'utilisateur* peut être défini comme « une source de connaissance [...], qui contient des *suppositions explicites* sur tous les aspects de l'utilisateur qui peuvent être pertinents pour le comportement du

système » (Wahlster et Kobsa, 1989). Le *modèle d'apprenant* est un modèle particulier, spécialisé du *modèle d'utilisateur*. Pour permettre l'identification de l'apprenant dans le cadre du système, les premières caractéristiques sont héritées d'un modèle d'utilisateur (v. Figure 5-3). Ces caractéristiques ne dépendent pas de l'évolution du système, elles sont intrinsèques à l'apprenant. À l'exception des champs *identifiant*, *nom d'utilisateur* et *mot de passe*, ces variables peuvent être importés d'un enregistrement de type *hcard*.

L'apprenant doit être décrit dans un *contexte d'apprentissage*, spécifiant les caractéristiques pertinentes à l'apprentissage. Le modèle est un modèle des « connaissances, des difficultés et des malentendus d'un individu » (Bull, 2004), « une représentation qualitative de ses *connaissance cognitives et affectives* [...] » (Woolf, 2010, p. 49).

Les données enregistrées dans le modèle d'apprenant sont « mises à jour au fur et à mesure que l'apprenant assimile le contenu d'apprentissage ciblé, afin de refléter les croyances actuelles de l'utilisateur » (Bull, 2004). Ce modèle tente d'expliquer la performance de l'apprenant et permet l'ajustement du processus d'apprentissage (e.g. contenu d'apprentissage, activités de l'apprenant) selon des règles et des raisonnements logique au niveau d'agent logiciel.

L'agent logiciel doit avoir une représentation « exacte » de l'apprenant du point de vue de l'apprentissage, afin de permettre une représentation fidèle de la réalité dans le cadre du système. L'enrichissement du modèle via des caractéristiques diverses (e.g. styles d'apprentissage différents) permet d'affiner la représentation de l'apprenant dans le cadre du système. Rendre l'apprenant acteur de son propre apprentissage implique son accès ouvert aux informations qui le concernent (e.g. objectifs d'apprentissage, caractéristiques personnels). *L'Open Learner Model* est une approche qui permet de rendre le modèle d'apprenant visible et accessible à l'apprenant (Bull et Pain, 1995; Hansen et McCalla, 2003). Cette approche encourage l'apprenant à prendre conscience et à contrôler son processus d'apprentissage (Bull, 1997). Elle favorise la métacognition (e.g. planification, réflexion), l'autonomie, l'adaptabilité, la correction des informations enregistrées dans le modèle, augmente la motivation de l'apprenant et fournit un cadre

favorable pour l'autorégulation (Kay *et al.*, 1997; Hansen et McCalla, 2003; Bull, 2004; Bull et Kay, 2010).

Un modèle d'apprenant « complet inclurait *tout l'apprentissage concerné préalable, le progrès de l'apprenant dans son programme scolaire (curriculum), le style d'apprentissage préféré, aussi bien que d'autres types d'informations relatives à un apprenant* » (Holt *et al.*, 1994, p. 4).

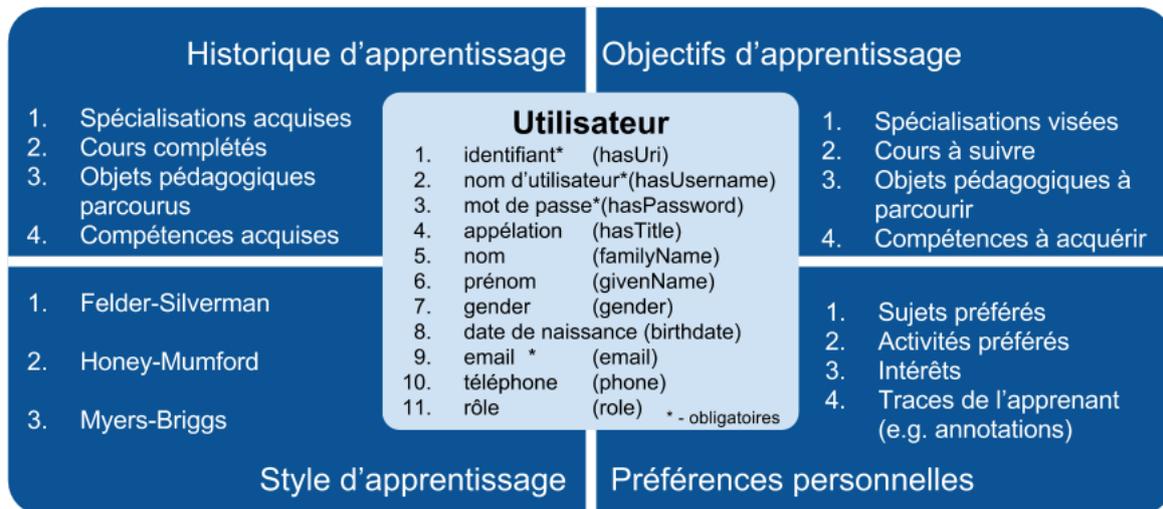


Figure 5-3. Dimensions et caractéristiques de l'apprenant

Nous proposons un modèle d'apprenant décliné sur quatre dimensions (v. Figure 5-3) (Szilagyi et Roxin, 2012). Dans un premier temps, le modèle d'apprenant capte l'évolution d'acquisition des connaissances de l'apprenant et suit son accomplissement des objectifs d'apprentissage. Ces informations sont associées à une *activité* qui encadre un *objet d'activité* (e.g. objet pédagogique, objet d'évaluation, cours). L'activité est spécifiée dans le temps à l'aide des propriétés `up:startedTime` et `up:completedTime`. Ainsi, des activités spécifiques sont modélisées en faisant référence aux dimensions personnelles de l'apprenant. Les deux premières dimensions concernent le *niveau d'acquisition des connaissances* : (1) l'historique d'apprentissage et (2) les objectifs d'apprentissage (v. Figure 5-4).

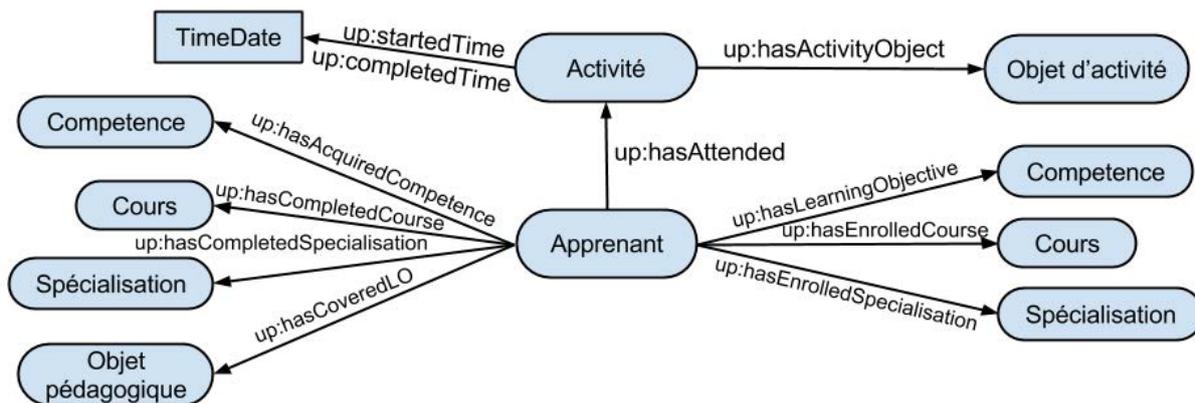


Figure 5-4. Niveau d'acquisition des connaissances

Ces caractéristiques permettent d'encadrer l'apprenant dans son parcours d'apprentissage. Elles sont enregistrées sur trois niveaux : spécialisation, cours, compétences. *L'historique d'apprentissage* est exprimé à travers quatre niveaux : (1) *compétences acquises*, (2) *cours parcourus et complétés*, (3) *spécialisations acquises, obtenues* et (4) *les objets pédagogiques parcourus*. L'état et les capacités actuelles de l'apprenant facilitent la sélection des prochains objectifs d'apprentissage. Connaissant les compétences actuelles de l'apprenant, le système permet à l'apprenant de choisir et fixer son meilleur parcours d'apprentissage afin d'atteindre ses *objectifs d'apprentissage spécifiques*. Les relations entre les *prérequis* et les *compétences* permettent de choisir un parcours optimal. Enregistrant le niveau de connaissances ou de compétences acquises, ces caractéristiques permettent d'influer sur le choix des objets pédagogiques sur la base de leur *contenu* (e.g. relation objet pédagogique – objectif d'apprentissage), plutôt que sur la base du type de présentation du contenu.

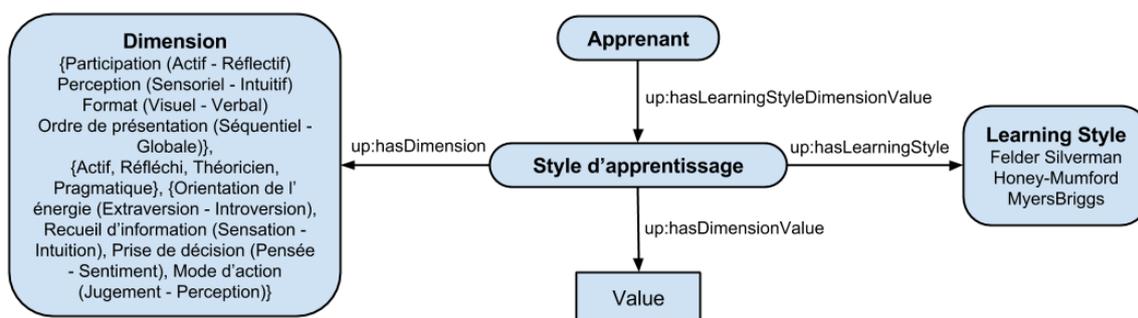


Figure 5-5. Représentation du style d'apprentissage

La troisième dimension porte sur le style d'apprentissage de l'apprenant (v Figure 5-6). Le style d'apprentissage *Felder-Silverman* (Felder et Silverman, 1988) décline sur quatre dimensions les manières possibles d'apprendre. Chaque dimension est exploitée sur ses deux extrémités (v. Figure 5-6).

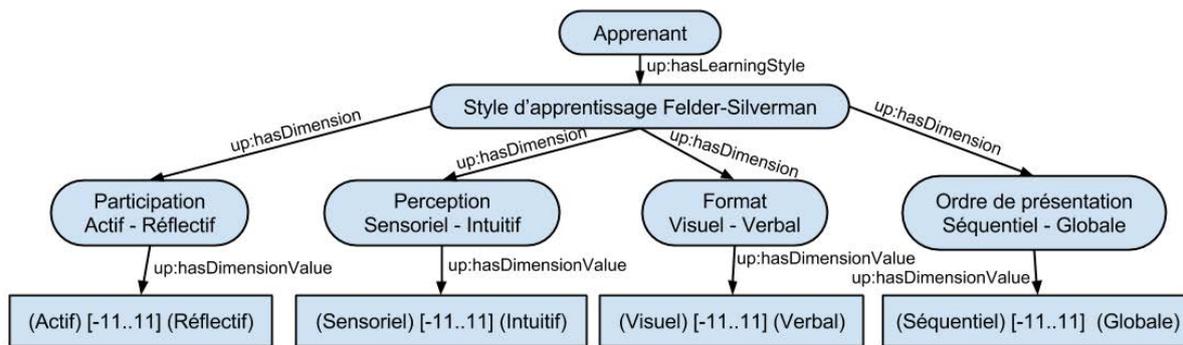


Figure 5-6. Le style d'apprentissage Felder-Felderman

L'Index du Style d'Apprentissage est un questionnaire avec 44 questions, 11 questions pour chaque dimension. Ce questionnaire permet d'établir un score de préférence de l'apprenant pour chaque dimension (Soloman et Felder, 2004). Les préférences sont exprimées en utilisant des valeurs numériques de -11 à +11. D'autres modèles de styles d'apprentissage existent dans la littérature comme celui proposé par D. Kolb (Kolb, 1984) ou P. Honey et A. Mumford (Honey et Mumford, 1992) et peuvent être intégrés dans le modèle de l'apprenant. Nous avons choisi d'utiliser celui proposé par Felder-Silverman, car « approprié pour des cours hypermédia » (Carver, Howard, et Lane, 1999). Pour permettre une représentation plus détaillée du style d'apprentissage, E. Popescu (2008) réalise une correspondance entre les différentes dimensions des styles d'apprentissage et propose le *Unified Learning Style Model* (Popescu, Trigano, et Badica, 2007). Les caractéristiques définies dans le style d'apprentissage permettent d'influer sur le choix des objets pédagogiques en regardant, à la fois, le *format du contenu* et la *forme de présentation*, plutôt que le contenu des objets pédagogiques.

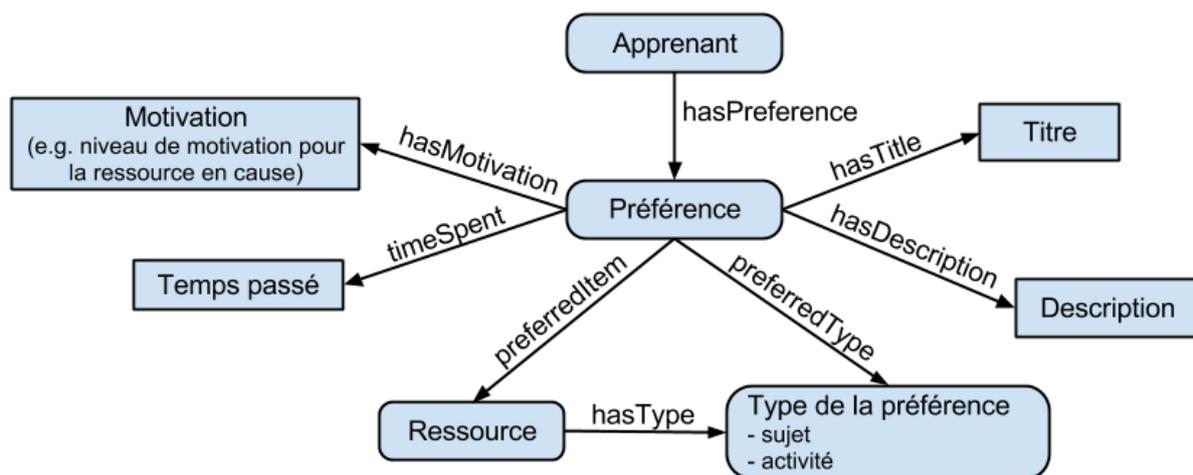


Figure 5-7. Préférences personnelles de l'apprenant

La quatrième dimension du modèle est constituée par les préférences personnelles de l'apprenant (v. Figure 5-7). Cette dimension permet d'enregistrer les particularités de l'apprenant en termes de *sujets* et *d'activités* personnelles préférés. Nous prenons en compte le niveau de la motivation reflété par le temps passé sur ces *sujets* ou en exerçant les *activités* préférés. Ces indicateurs permettent de mesurer *l'activité* et la *motivation* de l'apprenant sur les *sujets* d'intérêt, d'identifier les difficultés rencontrées ou de repérer des objets pédagogiques en connexion avec les sujets d'intérêt.

Le modèle de l'apprenant est un *sous-système* qui constitue l'identité de l'apprenant dans le système d'apprentissage. Il est mis à jour au fur et à mesure de l'interaction de l'apprenant avec le système d'apprentissage. Ces mises à jour représentent les *états du sous-système apprenant* et influencent l'état du système entier (e.g. compétences acquises par l'apprenant). Tenant compte de ces dimensions, la différence entre l'état visé par l'apprenant et l'état actuel de sous-système apprenant permet de construire un parcours d'apprentissage personnalisé (e.g. un parcours d'apprentissage déduit de l'enchaînement des objectifs d'apprentissage et de caractéristiques personnelles).

5.2.2. Modélisation de l'objet pédagogique

La modélisation de l'objet pédagogique utilisée dans notre approche se base sur les métadonnées spécifiées dans le standard LOM. Modélisé en RDF, le format du fichier n'est plus soumis aux contraintes de la structure hiérarchique imposée par XML. Pour la réalisation du modèle d'objet pédagogique nous nous basons sur le travail effectué par

M. Nilsson (2010) et R. Balog-Crişan (2011). Ces deux travaux utilisent le langage RDF pour exprimer les métadonnées d'objets pédagogiques conformes au standard LOM, mais se différencient par l'intégration du vocabulaire *DublinCore* dans les travaux de M. Nilsson. Cette différence se retrouve sur les identifiants associés aux termes déclarés dans la spécification LOM, ainsi que sur l'espace de nom choisi.

Ayant toujours la base offerte par la spécification LOM, nous proposons une harmonisation entre le vocabulaire défini par M. Nilsson, R. Balog-Crişan et les métadonnées déclarées dans le *Learning Resource Metadata Initiative (LRMI)*. Nous présentons ici la modélisation de trois catégories de la spécification LOM : la catégorie *Général*, la catégorie *Pédagogique* et la catégorie *Classification*.

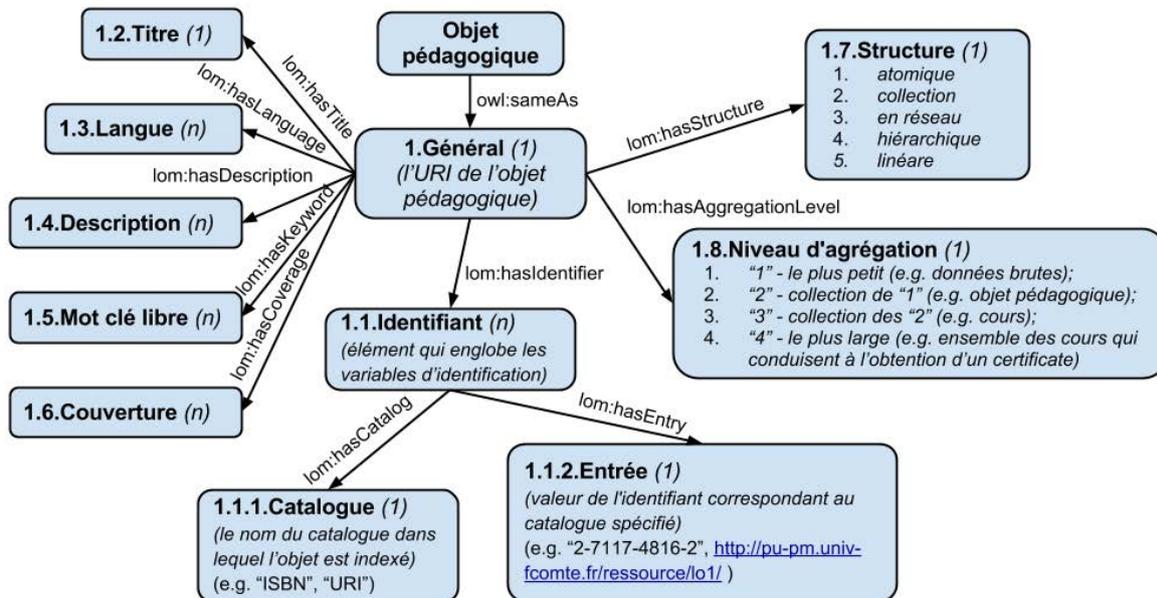


Figure 5-8. Représentation de la catégorie *Général* de LOM

La sous-catégorie *1.1. Identifiant* de la catégorie *1. Général* permet d'identifier l'objet pédagogique. L'identification se réalise à travers de deux variables : *1.1.1. Catalogue*, qui permet de spécifier le catalogue dans lequel l'objet pédagogique est indexé (e.g. « ISBN », « URI »), et *1.1.2. Entrée*, pour pointer (e.g. « 2-7117-4816-2 », « <http://pu-pm.univ-fcomte.fr/ressource/lo1/> ») vers l'objet choisi dans le catalogue spécifié. Le *niveau d'agrégation* permet de distinguer entre *objet pédagogique* et *cours*. La version française

du LOM (LOM-FR) introduit une nouvelle caractéristique appelée *Type documentaire*²⁷⁸. Elle est située dans la catégorie 1. *Général*, ayant le chiffre d'indexation 1.9. Ce type spécifie la *nature ou le genre du contenu* de la ressource d'apprentissage (AFNOR, 2006). Cette caractéristique se distingue de la caractéristique 4.1. *Format* du standard LOM, qui spécifie le *format* de la ressource d'apprentissage d'un point de vue *technique* (e.g. *format du fichier*). Elle se distingue également de la caractéristique 5.2. *Type de la ressource pédagogique*, qui spécifie le type de la ressource d'apprentissage d'un point de vue pédagogique (e.g. *exercice, simulation, cours*).

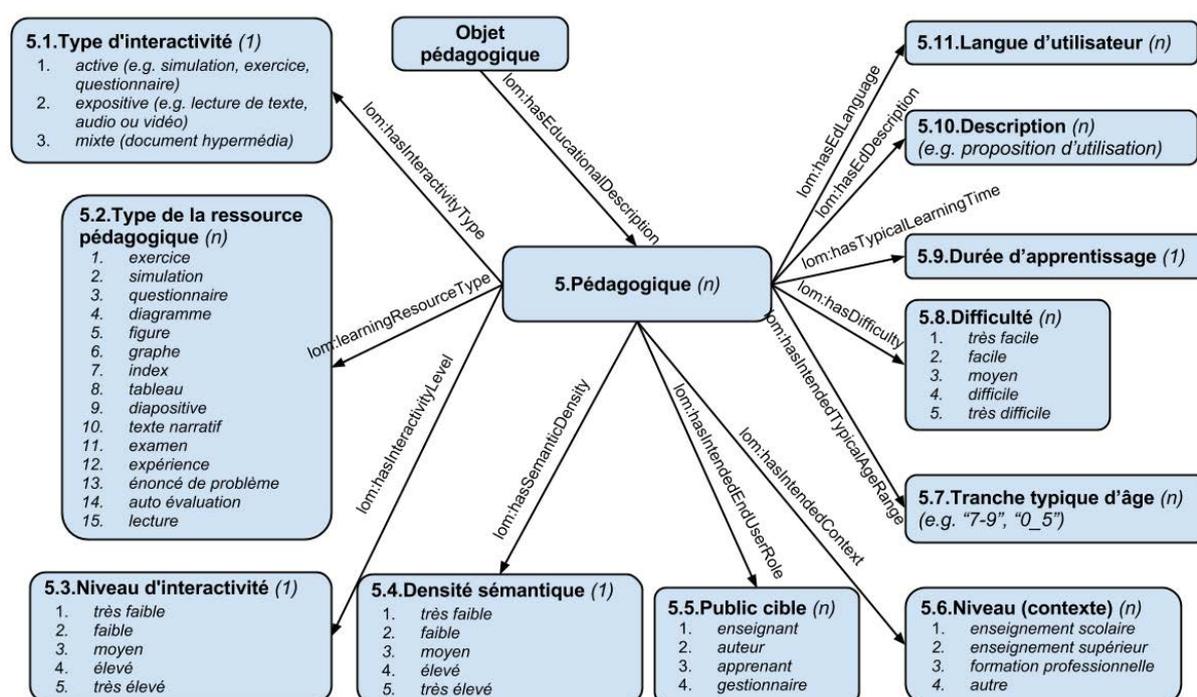


Figure 5-9. Représentation de la catégorie *Pédagogique* de LOM

La catégorie 5. *Pédagogique* (v. Figure 5-9) permet de spécifier les descriptions et les caractéristiques pédagogiques de l'objet. Le *type de la ressource pédagogique*, les *types et niveaux d'interactivité*, le *niveau et contexte d'enseignement*, la *difficulté dans un contexte spécifique* ou le *public ciblé* sont des caractéristiques qui permettent d'encadrer la ressource dans le contexte d'apprentissage.

Les descriptions de cette catégorie augmentent la visibilité de la ressource d'apprentissage pour des recherches avec des critères spécifiques du point de vue

²⁷⁸ http://www.sup.lomfr.fr/index.php?title=1.9_Type_documentaire

pédagogique. L'association d'identifiants IRI aux termes utilisés (e.g. *exercice*, *questionnaire*) dans la description des variables (e.g. *type de la ressource pédagogique*) permet l'harmonisation de l'ensemble des ressources d'apprentissage selon les critères déclarés et un filtrage pertinent selon ces critères.

Les équivalences entre les propriétés déclarées conformes au standard LOM et celles déclarées dans le *LRMI* sont illustrées dans le Tableau 5-1. Compte tenu du fait que le *LRMI* était conçu particulièrement pour décrire les ressources existantes sur le Web, ces équivalences facilitent l'intégration des ressources distribuées sur le Web dans le cadre du système.

Tableau 5-1. Association entre les éléments du LOM et LRMI

<i>Description</i>	<i>LOM</i>	<i>LRMI</i>
le rôle qui décrit le public cible du contenu d'apprentissage	<i>hasIntendedEndUserRole</i>	<i>educationalRole</i>
l'objectif de la ressource dans le contexte d'apprentissage (e.g. devoir, travail en groupe)	<i>hasEdDescription</i>	<i>educationalUse</i> (sous-propriété)
le temps approximatif nécessaire pour travailler avec la ressource d'apprentissage en rapport avec l'audience typique ciblé	<i>hasTypicalLearningTime</i>	<i>timeRequired</i>
l'âge de l'utilisateur pour lequel la ressource d'apprentissage était prévu	<i>hasIntendedTypicalAgeRange</i>	<i>typicalAgeRange</i>
le mode d'apprentissage prédominant soutenu par cette ressource d'apprentissage, les valeurs acceptables sont <i>active</i> , <i>expositive</i> ou <i>mixed</i>	<i>hasInteractivityType</i>	<i>interactivityType</i>
le type prédominant qui caractérise la ressource d'apprentissage (e.g. présentation)	<i>learningResourceType</i>	<i>learningResourceType</i>

Comme souligné avant, la relation d'une ressource d'information avec un *objectif d'apprentissage* définit un objet pédagogique. Ce type de relation est décrit dans la variable 9.1. *Objectif*, sous-catégorie de la catégorie 9. *Classification* (v. Figure 5-10) de la spécification LOM. Cette variable peut prendre 9 valeurs : *discipline*, *notion*, *prérequis*,

objectif pédagogique, restriction d'accessibilité, niveau éducatif, niveau de compétence, niveau de sécurité, compétence. La sous-catégorie 9.2. Chemin du taxon permet d'englober les éléments d'identification du taxon.

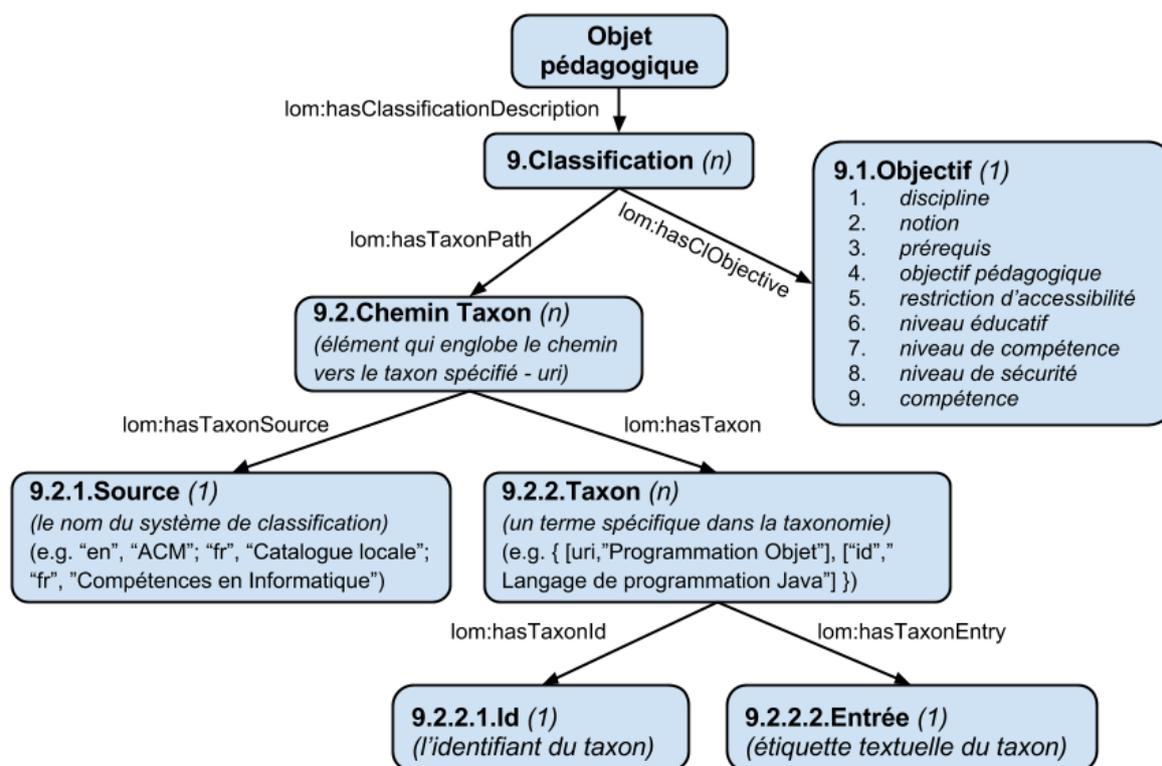


Figure 5-10. Représentation de la catégorie Classification de LOM

Un taxon est un « nœud qui a défini une étiquette et peut avoir associé un identificateur pour une référence standardisé » (IEEE, 2002, p. 35)

L'absence de certaines de ces descriptions peut rendre difficile l'utilisation de la ressource dans le contexte d'apprentissage. Les ressources qui ne sont pas spécialement conçues pour l'apprentissage, mais qui sont reconnues efficaces dans le processus d'apprentissage, notamment les articles wiki, les articles de blog, les discussions de forums ou les commentaires associés peuvent toujours être intégrées dans le parcours d'apprentissage. Sans un objectif d'apprentissage ou une composante d'instruction éducationnelle (e.g. éléments pédagogiques - spécifiques à l'objet pédagogique), ces objets sont considérés comme des ressources d'apprentissage. La classe *Objet pédagogique* est une sous-classe de la classe *Ressource d'apprentissage* (v. Figure 5-11).

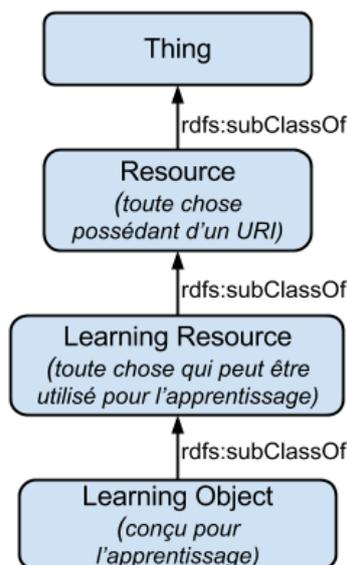


Figure 5-11. Relation *objet pédagogique - ressource d'apprentissage*

Les pratiques sociales sur le Web, comme les *tags*, permettent l'annotation des ressources (e.g. articles wiki, articles de blog) facilitant leur visibilité et indexation. L'association de ces *tags* (propriété *hasTag*) aux *mots-clés* (propriété *hasKeyword*) d'un objet pédagogique offre un lien entre l'objet pédagogique et la ressource d'apprentissage. L'utilisation d'outils et vocabulaires spécifiques (e.g. Schema.org, SIOC) facilite l'intégration de ces ressources dans l'ensemble du SASA.

5.2.3. Modélisation de l'objectif d'apprentissage

Voulant rendre l'enseignement plus efficace, B.F. Skinner « se proposait de définir les *objectifs de l'éducation* en termes de *comportements observables et mesurables* » (Rézeau, 2001, p. 247) : « La pensée humaine doit se définir en termes de *comportements réels*, qui méritent d'être traités pour eux-mêmes comme les *objectifs concrets de l'éducation* » (Skinner, 1968, p. 35). L'accentuation de l'importance des *objectifs d'apprentissage* (ou pédagogiques) dans le processus d'apprentissage conduit à la *pédagogie par objectifs*²⁷⁹ (PPO). Promoteurs de la pédagogie par objectifs, R.W. Tyler et R. Mager proposent une définition de l'objectif d'apprentissage proche du point de vue behaviouriste. Ils « proposent aux enseignants de définir les objectifs pédagogiques en termes de

²⁷⁹ La PPO rappelle du mécanisme de régulation dans le cadre des systèmes. La PPO consiste généralement en trois étapes principales : (1) la définition de l'objectif comportemental (spécifique et observable), (2) l'évaluation du niveau d'atteinte de l'objectif et (3) en fonction du résultat obtenu, la modification des stratégies (Raynal et Rieunier, 2010).

comportements observables » (Raynal et Rieunier, 2010, p. 342). B. Bloom propose un encadrement de l'objectif d'apprentissage sur trois domaines : (1) *le domaine cognitif* (e.g. savoir, connaissances), qui regroupe les connaissances et les capacités régissant des facultés de la mémoire, (2) *le domaine psychomoteur* (e.g. savoir-faire, habiletés), qui englobe les comportements ou les actions corporelles et (3) *le domaine affectif*²⁸⁰ (e.g. savoir-être, attitudes), qui réunit les attitudes et les émotions (ibid., p. 82). Dans le domaine cognitif la *taxonomie des objectifs éducationnels* (connue sous le nom de *taxonomie de Bloom*) est structurée sur six niveaux majeurs : (1) *connaissance*, (2) *compréhension*, (3) *application*, (4) *analyse*, (5) *synthèse* et (6) *évaluation* (Bloom et al., 1956).

Un *objectif d'apprentissage* est un *énoncé* de *ce qu'un individu est censé apprendre* et le niveau acceptable de ce qu'il a appris (Koohang et Harman, 2007, p. 154). Selon R. Tyler, l'objectif d'apprentissage doit décrire : (1) le comportement observable ; (2) les conditions de réalisation et (3) les critères de performance (Tyler, 1949).

Généralement *l'objectif d'apprentissage* encadre (1) un contenu relatif au sujet d'intérêt et (2) une description de ce que doit être fait avec ce contenu. Les *énoncés* des objectifs d'apprentissage sont constitués d'un « *substantif ou phrase substantive* », qui décrit le contenu du sujet, et d'un « *verbe ou phrase verbe* », qui spécifie le processus cognitif (David R. Krathwohl, 2002, p. 213). Si, dans la version originale de la *taxonomie de Bloom*, les termes pouvaient inclure les deux aspects, dans sa version mise à jour, publiée en 2001 (L. W. Anderson et al., 2001), tous les niveaux sont exprimés à travers un *verbe* : (1) *se rappeler* ; (2) *comprendre* ; (3) *appliquer* ; (4) *analyser* ; (5) *évaluer* et (6) *créer*.

L'existence des dimensions de l'objectif pédagogique (e.g. cognitive, psychomoteur, affective) dans la définition de la compétence révèle l'équivalence de la composition de ces deux concepts. L'acquisition des objectifs d'apprentissage vise l'achèvement des

²⁸⁰ Le domaine affectif comprend la manière dont nous traitons les choses émotionnellement (e.g. sentiments, valeurs, appréciations, enthousiasmes, motivations, attitudes). Les cinq grandes catégories (du plus simple à la plus complexe) sont : (1) réception (e.g. sensibilité, demande) ; (2) réponse (e.g. réponde, aide) ; (3) valorisation (e.g. explique, démontre) ; (4) organisation (e.g. adhère, organise) et (5) caractérisation par une valeur ou un système de valeur (e.g. agis, influence) (D. R. Krathwohl, Bloom, et Masia, 1964).

compétences. À son origine, le concept de *compétence* « était développé en psychologie référant la capacité d'un individu à répondre à certaines exigences imposées par l'environnement » (D. Sampson et Fytros, 2008, p. 160). Ces capacités des individus résultent d'un apprentissage, dont l'accomplissement des objectifs d'apprentissage entraîne l'acquisition des *compétences*. En fait, l'IEEE intègre le concept d'objectif d'apprentissage à celui de compétence : la compétence est un objectif d'apprentissage déjà acquis (IEEE, 2006, p. 3).

À cheval entre le domaine de l'éducation et celui des ressources humaines, la définition du concept de *compétence* est une tâche difficile même dans la littérature concernée (Le Deist et Winterton, 2005; D. Sampson et Fytros, 2008). Soulignant les éléments intégrant le concept de *compétence*, nous rappelons ici quelques définitions qui nous permettent de visualiser les caractéristiques essentielles :

- « Une combinaison d'*habiletés*, d'*aptitudes* et de *connaissances* nécessaires pour accomplir une tâche spécifique » (Voorhees, 2001).
- « Une combinaison dynamique de *connaissances*, *compréhensions*, *habiletés* et *aptitudes* » (Sánchez-Ruiz, Edwards, et Ballester-Sarrias, 2006).
- Une compétence est un « énoncé de principe qui détermine une relation entre un *public cible* ou « acteur », une *habileté* et une *connaissance* » (Paquette, 2002b, p. 187).
- Selon IEEE, la *compétence* est « toute forme de *connaissances*, d'*habiletés*, d'*attitudes*, de *capacités*, ou d'*objectifs d'apprentissage* qui peut être décrit dans un contexte d'apprentissage, d'éducation ou de formation » (IEEE, 2006, p. 3).
- Conformément au *Conseil international des normes de formation, de performance et de l'instruction*²⁸¹, une compétence est « un ensemble intégré d'*habiletés*, de *connaissances* et d'*attitudes* qui permettent à une personne d'accomplir efficacement les activités d'une profession ou fonction donnée selon des normes exigées dans une profession ou milieux de travail spécifique » (Richey, Fields, et Foxon, 2001).

²⁸¹ International Board of Standards for Training, Performance and Instruction : <http://www.ibstpi.org>

- L'organisation des *Nations Unis* définit la compétence comme « une combinaison des *habiletés, attributs* et *comportements* qui sont directement liés au succès de performance d'un tâche » (United Nations, s. d.).
- A travers la spécification des compétences clés dans le cadre du programme *Formation tout au long de la vie*, L'Union Européenne voit la compétence comme « un ensemble de *connaissances, d'aptitudes* et *d'attitudes* appropriées au contexte » (Europa.eu, 2007, p. 3).

Les caractéristiques principales de la compétence, résultantes de ces définitions, sont présentées dans la modélisation proposée par D. Sampson et D. Fytros (D. Sampson et Fytros, 2008) :

1. les caractéristiques ou les capacités de l'individu (e.g. ensemble des connaissances, des habiletés, des attitudes) ;
2. le niveau de performance (e.g. le niveau d'exercice des capacités de l'individu selon des classifications spécifiques) ;
3. le contexte (e.g. le cadre dans lequel la compétence est appliquée).

Les dimensions de l'objectif d'apprentissage se retrouvent dans les *caractéristiques de l'individu*. Elles sont caractérisées par un *niveau de performance* appliqué dans un *contexte spécifique*.

Nous proposons une modélisation des objectifs d'apprentissage sur trois niveaux spécifiques (v. Figure 5-12).

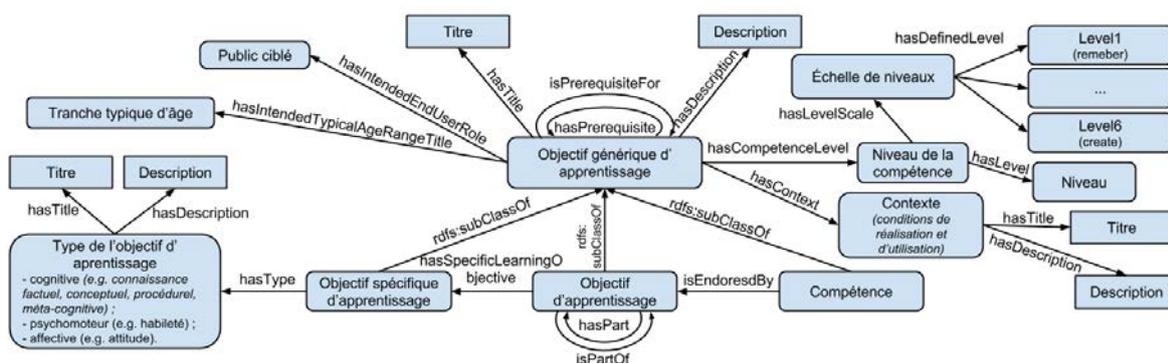


Figure 5-12. Modélisation des objectifs d'apprentissage

L'*objectif spécifique d'apprentissage* permet de définir un type de caractéristique (e.g. cognitive, psychomoteur, affective) que l'apprenant va apprendre. L'acquisition de plusieurs objectifs d'apprentissage *spécifiques* permet la composition de l'*objectif d'apprentissage*. Il permet de regrouper plusieurs objectifs spécifiques ayant des types

de caractéristiques différents. Ce type d'objectif pédagogique permet la *recomposition* des objectifs d'apprentissage (spécifiques ou non) afin de satisfaire une plus grande couverture selon les nécessités dans l'utilisation. Enfin, *l'objectif générique d'apprentissage* encadre tous les types des objectifs d'apprentissage, ainsi que la *compétence* et permet la définition des caractéristiques communes de ces objets. Le niveau de performance de l'objectif, l'échelle correspondante et le contexte (de réalisation ou d'utilisation) sont hérités des caractéristiques spécifiées au niveau de l'objectif d'apprentissage générique.

Les objectifs d'apprentissage acquis (spécifiques ou génériques) permettent de construire *le profil des capacités* ou *des compétences* de l'apprenant. Le choix d'une spécialisation permet de regrouper plusieurs *objectifs génériques d'apprentissage*. Dans le cas général, un apprenant choisit la spécialisation qu'il veut acquérir. Ce choix se concrétise dans la définition de *l'ensemble des objectifs spécifiques d'apprentissage*. Suite aux choix de l'apprenant, le système d'apprentissage permet la création d'un parcours²⁸² d'apprentissage adapté au profil de l'apprenant.

5.2.4. Modélisation de l'objet d'évaluation

Les objectifs d'apprentissage permettent de spécifier des étapes dans le progrès de l'apprentissage. Le niveau d'accomplissement de ces étapes et l'acquisition des objectifs d'apprentissage fournissent une image de l'évolution du processus d'apprentissage. L'évaluation consiste dans la confrontation d'une performance par rapport à un objectif fixé. L'évaluation mesure la qualité du processus d'apprentissage. Les *objets d'évaluation* fournissent les moyens d'évaluation permettant la vérification du niveau d'acquisition des objectifs d'apprentissage.

Les objectifs d'apprentissage fixent les résultats à atteindre, l'évaluation « mesure ces résultats, ou plutôt l'écart entre les objectifs et les résultats » (Bonniol et Vial, 2009, p. 300). C'est une « *opération* qui mesure l'écart entre un objectif prédéterminé que l'on poursuit et le résultat obtenu », donc l'évaluation *mesure l'erreur* (Tyler, 1942; cité par

²⁸² RelFinder est un exemple d'application pour trouver des relations entre les concepts dans le Web sémantique : <http://www.visualdataweb.org/refinder/refinder.php>, <http://www.uni-due.de/~s400268/RelFinder-SAMT09.pdf>

Bonniol et Vial, 2009, p. 301). Évaluer consiste à « recueillir un ensemble d'informations reconnues comme suffisamment pertinentes, valides et fiables et à examiner le degré d'adéquation entre cet ensemble d'informations et un ensemble de critères jugés suffisamment adéquats aux objectifs fixés au départ ou ajustés en cours de route, en vue de fonder une prise de décision », car « toute évaluation autour des apprentissages a pour fonction principale de prendre une décision d'action. » (J.-M. de Ketele, 1993, p. 68). L'évaluation est un jugement de valeur argumenté dans le but de prendre une décision en comparant un référé (le résultat obtenu) à un référent (le résultat visé). (Barbier, 1985)

Dans le cadre de la pédagogie par objectifs, le modèle *tylérienne*²⁸³ suppose huit étapes : « (1) la détermination des objectifs ; (2) le classement des objectifs dans un système de catégorisation (taxonomie) ; (3) la définition des objectifs en termes comportementaux ; (4) l'établissement des situations et des conditions dans lesquelles la maîtrise des objectifs peut être démontrée ; (5) l'explication des buts et des fondements de la stratégie au personnel impliqué dans les situations choisies ; (6) le choix et le développement des techniques de mesure appropriées ; (7) la collecte des données de performances ; (8) la comparaison de celles-ci aux objectifs comportementaux » (J.-M. de Ketele, 1993, p. 63).

Les grands lignes de la pédagogie de la maîtrise proposée par B. Bloom, sont toujours centrées autour de l'objectif d'apprentissage et de l'évaluation : « (1) préciser clairement les résultats attendus à la fin d'un cours ou d'une séquence d'apprentissage ; (2) préparer les apprenants pour qu'ils puissent entrer fructueusement dans la séquence d'apprentissage ; (3) enrichir l'apprentissage de rétroactions fréquentes et de démarches correctives ; (4) ne pas passer à l'apprentissage ultérieur si l'apprentissage actuel n'est pas suffisamment maîtrisé » (ibid., p. 64). Dans le contexte systémique, l'évaluation est un moyen de *rétroaction*, de *régulation du processus d'apprentissage*. Les moyens d'évaluations servent à *enrichir* l'apprentissage et à s'assurer que *l'apprentissage actuel est bien maîtrisé*.

²⁸³ D'après R. Tyler (1902-1924) : « père des objectifs comportementaux, spécialiste d'évaluation » (Raynal et Rieunier, 2010, p. 459)

Selon ces deux approches, le processus d'évaluation implique donc la comparaison du *résultat* avec un *objectif* et le conditionnement de passage au nouvel objectif donné par le résultat de l'évaluation. Les objets d'évaluation permettent d'établir les *résultats*, en tant qu'accomplissement des *objectifs* d'apprentissage. La réalisation satisfaisante d'un objectif d'apprentissage selon son échelle des niveaux entraîne l'acquisition de l'objectif d'apprentissage et l'évolution du profil de l'apprenant. Selon le type de l'apprentissage (e.g. autonome, avec un tuteur), le résultat de la comparaison permet au système d'apprentissage d'intervenir sur le parcours d'apprentissage (e.g. ajouter des objets pédagogiques afin d'assurer l'accomplissement des objectifs) ou d'assister l'enseignant dans ses décisions.

La Figure 5-13 présente les principales classes et propriétés en relation avec l'objet central de cette modélisation : *l'objet d'évaluation*.

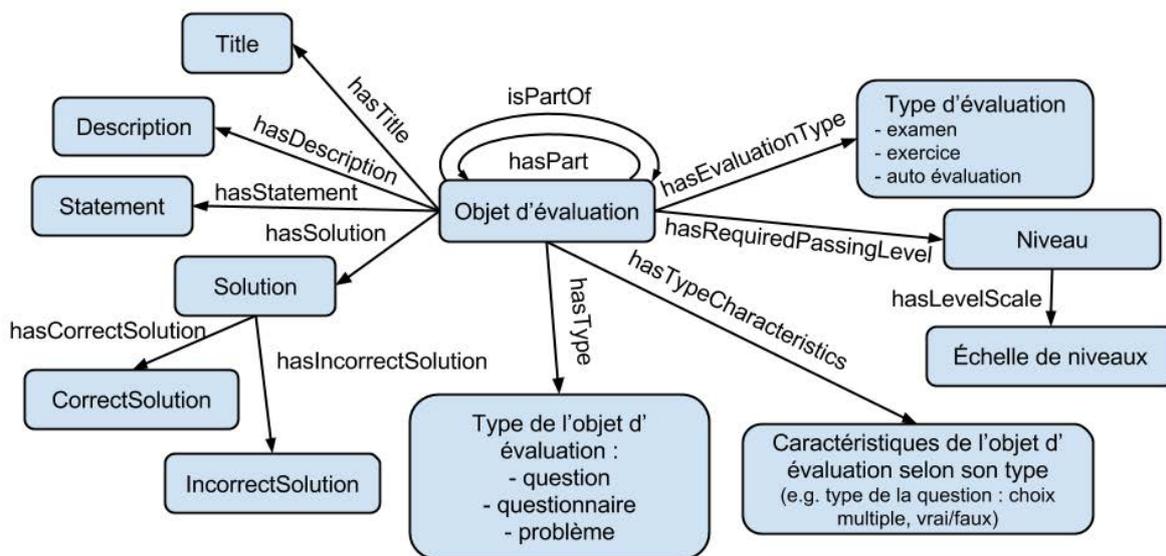


Figure 5-13. La modélisation de l'objet d'évaluation

Parmi les objets d'évaluation les plus répandus nous rappelons *la question*, *le questionnaire* ou *le problème*. Un objet d'évaluation de type *questionnaire* est composé de plusieurs objets de type *question* (e.g. les relations `hasPart` et `isPartOf`). Ces types d'objets reprennent les propriétés de la classe générique *objet d'évaluation* comme le *titre*, la *description* ou le *niveau* nécessaire d'accomplissement. Les objets d'évaluation peuvent être autonomes ou partie intégrante de l'objet pédagogique (e.g. exercices).

5.2.5. Modélisation de l'objet d'annotation

De façon générale, une annotation est une « note critique ou explicative qui accompagne un texte » (Le Robert *et al.*, 2010, p. 100). Les annotations sont des « marques inscrites par un individu sur un document dans un but spécifique » (Mille, 2005, p. 402). En tant qu'objet, l'annotation est un « élément de la mémoire externe de formation des apprenants, qui permet de soulager leur mémoire humaine ». L'annotation en tant qu'activité caractérise la « *lecture active* » (*ibid.*). Les annotations peuvent prendre plusieurs formes telles que symboles, notes, signes, etc.

Dans le processus de régulation du système d'apprentissage, nous considérons l'annotation comme un des éléments clés. Fournissant des retours sur les ressources annotées et représentées dans le cadre du système, l'annotation permet l'évaluation et la classification des ressources selon les usages encouragés dans le Web social.

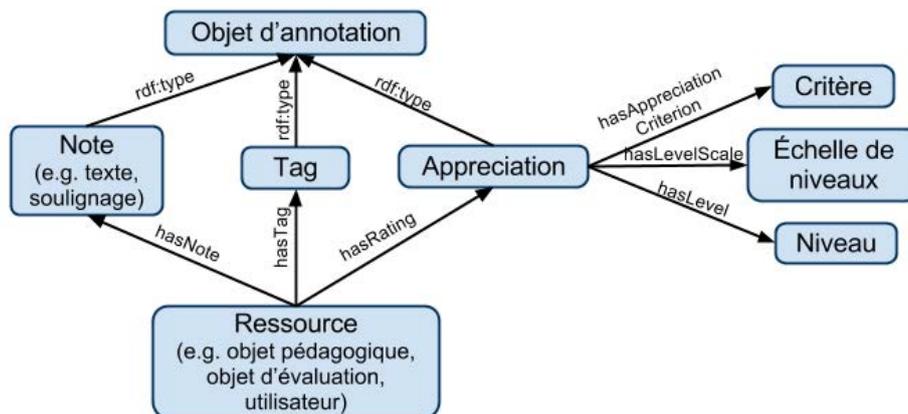


Figure 5-14. Modélisation d'objet d'annotation

Nous considérons trois types d'annotations principales (v. Figure 5-14) :

- La *note* est une marque de l'apprenant qui peut consister en l'ajout de contenu (e.g. texte, dessin) ou la mise en évidence (e.g. soulignage).
- Le *tag* permet d'associer des étiquettes à des ressources, cela permettant la construction des *folksonomies*.
- L'*appréciation* est une marque qui permet à l'apprenant d'évaluer une ressource selon des critères spécifiés et sur une échelle donnée.

Une appréciation par rapport à une ressource d'apprentissage permet de construire la réputation de la ressource en question. Le type d'utilisateur (e.g. apprenant, enseignant), son historique d'annotation et son profil de compétences permettent d'influencer le

pois de l'appréciation de la ressource annotée et ensuite influencer le degré de modification de la réputation²⁸⁴.

La corrélation d'un objet d'annotation avec un profil d'apprenant permet d'identifier des sensibilités de l'apprenant en rapport avec la ressource annotée (e.g. sujet traité, type de ressource). Ces sensibilités se traduisent par une meilleure personnalisation du parcours d'apprentissage.

5.2.6. Modélisation du cadre d'enseignement

Pour faciliter l'intégration du système d'apprentissage avec le modèle d'enseignement réel, nous proposons une modélisation du cadre d'enseignement en concordance avec le système d'éducation traditionnel. Dans le but de faciliter l'emploi ou la mobilité des étudiants, les pays de l'Union Européenne signent la *Déclaration de Bologne* en 19 juin 1999. Parmi les six objectifs²⁸⁵ définis dans la déclaration nous rappelons l'adoption des cycles d'enseignement (e.g. licence, master, doctorat) et l'établissement du système ECTS (European Credit Transfer System).

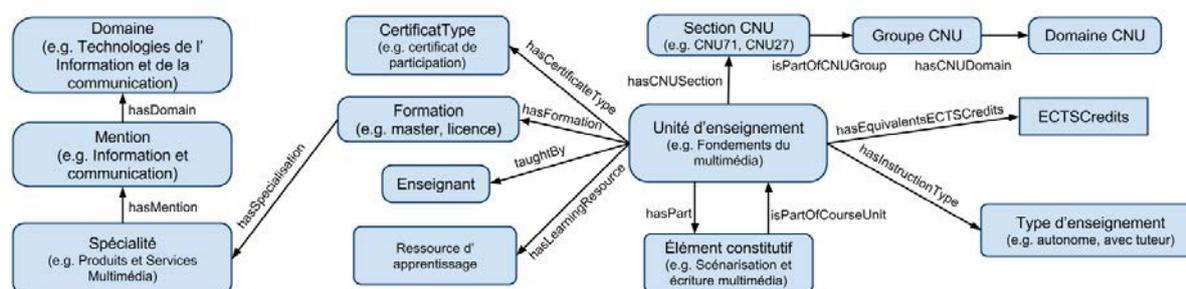


Figure 5-15. Modélisation du cadre d'enseignement

Le modèle du cadre d'enseignement permet de situer l'unité d'enseignement (e.g. cours) dans son contexte d'apprentissage (v. Figure 5-15). Décliné dans des unités d'enseignements et éléments constitutifs de ces unités, le cours est caractérisé par son

²⁸⁴ Appréciations dans les plateformes de type Question-Réponse (e.g. stackoverflow), système de réputation à travers les votes pour les questions.

²⁸⁵ Six objectifs retenus pour construire l'espace européen d'apprentissage : (1) l'adoption d'un système de diplômes comparables ; (2) l'adoption d'un système de formation basé sur deux cursus : études pré-graduées suivies d'études post-graduées ; (3) la formation pré-graduée, d'une durée minimale de 3 ans, conduit à un diplôme permettant l'accès au marché du travail ; (4) la généralisation du système de crédits ECTS ; (5) la promotion de la mobilité pour les étudiants et les enseignants ; (6) la promotion de la collaboration en matière d'évaluation de la qualité ; (7) la promotion de la dimension européenne de l'enseignement supérieur.

encadrement correspondant à la formation qui le propose (e.g. master, licence) et à la classification du *Conseil National des Universités (CNU)*. La spécialité permet de grouper les compétences que le participant est censé acquérir en parcourant les unités d'enseignement associées.

5.3. Ontologies des domaines dans le système d'apprentissage

Pour rendre cette modélisation interprétable par les logiciels ou les agents intelligents (e.g. système informatique d'apprentissage), nous utilisons le langage de construction d'ontologies OWL. Pour l'écriture des ontologies nous abordons une approche modulaire autour des modélisations présentées précédemment. Cela nous permet de séparer les domaines modélisés, d'augmenter la flexibilité dans la déclaration des entités et de faciliter la maintenance via une approche distribuée.

Pour la spécification des termes des ontologies nous utilisons l'espace de nom dérivé de l'adresse IRI de l'équipe OUN : <http://semlearn.pu-pm.univ-fcomte.fr/ontologies/>. C'est sous ce nom de domaine que nous enregistrons chaque ontologie dans un fichier différent. Les préfixes des vocabulaires utilisés dans les ontologies sont présentés dans l'Extrait de code 5-1.

```
@prefix : <http://semlearn.pu-pm.univ-fcomte.fr/ontologies/learner.owl#> .
@prefix : <http://semlearn.pu-pm.univ-fcomte.fr/ontologies/learning-
object.owl#> .
@prefix : <http://semlearn.pu-pm.univ-fcomte.fr/ontologies/learning-
objective.owl#> .
@prefix : <http://semlearn.pu-pm.univ-fcomte.fr/ontologies/evaluation-
object.owl#> .
@prefix : <http://semlearn.pu-pm.univ-fcomte.fr/ontologies/annotation-
object.owl#> .
@prefix : <http://semlearn.pu-pm.univ-fcomte.fr/ontologies/learning-
environnement-ontology.owl#> .

@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix xml: <http://www.w3.org/XML/1998/namespace> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
```

Extrait de code 5-1. Les préfixes des vocabulaires utilisés dans la construction des ontologies

Afin de faciliter la lecture des ontologies, les extraits de code que nous présentons sont écrits en utilisant la syntaxe *Turtle*.

5.3.1. Ontologie de l'apprenant

L'ontologie de l'apprenant permet l'expression du modèle d'apprenant. Dans le cadre du système informatique d'apprentissage, l'utilisateur peut prendre quatre *rôles* différents, chacun avec des droits correspondants. L'*apprenant* est un type *d'utilisateur* dont le rôle est *étudiant*. La classe :Learner est une sous-classe de la classe :User ayant une contrainte sur la propriété :hasRole (v. ANNEXE H).

Les caractéristiques nécessaires pour identifier l'utilisateur dans le système sont liées à la classe User. La classe User est une sous-classe de la classe Person, qui fournit des caractéristiques générales sur la personne en question, comme *le nom, prénom* ou *date de naissance*.

Le profil de connaissances d'un apprenant est représenté en spécifiant les objectifs d'apprentissage ou les spécialisations qu'il a complétés ou acquis. Les principales classes sont Specialisation, GenericLearningObjective, Course et LearningResource. Les différentes propriétés permettent de spécifier si les entités de ces types font partie de l'historique de l'apprenant ou des objectifs de l'apprenant. Les relations qui illustrent le rapport de l'apprenant avec les instances de ces classes sont les suivantes
hasCompletedSpecialisation, hasEnrolledSpecialisation,
hasAcquiredLearningObjective, hasLearningObjective, hasCompletedCourse,
hasEnrolledCourse, hasCoveredLearningResource.

Les classes principales pour représenter le style d'apprentissage sont représentées dans l'Extrait de code 5-2. Un exemple de la modélisation des dimensions et directions de ces styles est présenté dans l'ANNEXE N.

```

:LearningStyle rdf:type owl:Class .
:LearningStyleDimension rdf:type owl:Class .
:LearningStyleDimensionDirection rdf:type owl:Class .
:LearningStyleDimensionDirectionValue rdf:type owl:Class .

:FelderSilvermanLearningStyle rdf:type :LearningStyle ,
                                owl:NamedIndividual .
:HoneyMumfordLearningStyle rdf:type :LearningStyle ,
                                   owl:NamedIndividual .
:MyersBriggsTypeIndicator rdf:type :LearningStyle ,
                                 owl:NamedIndividual .

```

Extrait de code 5-2. Styles d'apprentissage dans le cadre du système

Les caractéristiques liées aux préférences de l'apprenant sont groupées dans la classe *Preference*. La modélisation des préférences permet au système d'enregistrer le type et l'engagement de l'apprenant (e.g. motivation, temps passé).

L'ontologie de l'apprenant est au cœur de la personnalisation dans le SASA. Ayant établi ces caractéristiques de l'apprenant, nous pouvons répondre aux questions sur *les spécialisations* et *les cours complétés* par l'apprenant, les *compétences* et les *objectifs d'apprentissage* acquis, ainsi que sur les *ressources d'apprentissage* parcourues. Ces informations fournissent le profil de connaissances de l'apprenant ce qui permet de proposer les objectifs d'apprentissage en concordance avec son historique d'apprentissage. De plus, les caractéristiques concernant le *style d'apprentissage* ou les *préférences de l'apprenant* permettent la personnalisation du parcours d'apprentissage.

5.3.2. Ontologie de l'objet pédagogique

L'ontologie de l'objet pédagogique concerne le savoir contenu dans l'objet. De ce point de vue, cette ontologie est un méta-savoir. Elle offre au système d'apprentissage les informations nécessaires afin de manipuler le savoir existant dans les objets pédagogiques. Dans ce qui suit, nous présentons des extraits de code de l'ontologie, selon la modélisation proposée, pour (1) l'identification de l'objet pédagogique, (2) la description des caractéristiques pédagogiques de l'objet et (3) l'association de l'objet pédagogique aux différents systèmes de classification, notamment avec le parcours des objectifs d'apprentissage. Le préfixe utilisé pour l'ontologie de l'objet pédagogique est présenté dans l'Extrait de code 5-3.

@prefix : <http://semlearn.pu-pm.univ-fcomte.fr/ontologies/learning_object.owl#> .

Extrait de code 5-3. Préfix de l'ontologie de l'objet pédagogique

L'ontologie de l'objectif d'apprentissage encadre la modélisation des ressources d'apprentissage en général (classe *LearningResource*) ainsi que leurs caractéristiques. L'objet pédagogique est un type spécifique de *ressource d'apprentissage*, principalement caractérisé par l'existence d'un *objectif d'apprentissage* en relation avec cette ressource. La restriction sur la propriété `owl:someValuesFrom` permet d'imposer l'existence d'un objectif d'apprentissage associé à l'objet pédagogique (v. Extrait de code 5-4).

```
:LearningResource rdf:type owl:Class .
:LearningObject rdf:type owl:Class ;
  owl:equivalentClass [ rdf:type owl:Class ;
    owl:intersectionOf ( :LearningResource
      [ rdf:type owl:Restriction ;
        owl:onProperty :hasPurpose ;
        owl:someValuesFrom :EducationalPurpose
      ]
    )
  ] .
```

Extrait de code 5-4. La classe des objets pédagogiques

L'objectif pédagogique est un *critère de classification* de l'objet pédagogique. Ces critères sont spécifiés dans la catégorie 9.1. *Purpose* conforme au standard LOM (v. ANNEXE I). Dans l'ontologie proposée, si la ressource d'apprentissage est classifiée en utilisant les termes `Purpose-educationalObjective` ou `Purpose-competence`, la ressource d'apprentissage est considérée comme un *objet pédagogique*. Étant spécialement conçu pour l'apprentissage, le statut d'objet pédagogique augmente la visibilité d'une ressource d'apprentissage.

Les caractéristiques listées dans la catégorie 5. *Pédagogique* du LOM sont particulièrement intéressantes pour la description de l'objet pédagogique. Elles permettent au système de qualifier la ressource d'apprentissage dans son contexte pédagogique. Ainsi ces propriétés sont des sous-propriétés de la propriété `:hasDescription-Educational` (v. Extrait de code 5-5)

```
:hasDescription-Educational rdf:type owl:ObjectProperty ;
  rdfs:range :EducationalDescription ;
  rdfs:domain :LearningResource .
```

Extrait de code 5-5. La propriété `hasDescription-Educational`

Les propriétés permettant d'exprimer les caractéristiques pédagogiques sont présentées dans l'ANNEXE I.

Nous présentons plus de détails concernant trois propriétés permettant d'évaluer le type et le niveau d'interactivité. La propriété `:hasInteractivityType` spécifie le type d'interactivité et les valeurs possibles sont les suivantes :

- `InteractivityType-active` spécifie que l'objet pédagogique engage l'apprenant dans l'interaction avec le contenu d'apprentissage (e.g. simulations, questionnaires) ;
- `InteractivityType-expositive` spécifie que l'objet pédagogique est destiné à exposer un contenu d'apprentissage ;
- `InteractivityType-mixed` spécifie que l'objet pédagogique intègre les deux types d'interactivité présentés ci-dessus.

La propriété `:hasInteractivityLevel` spécifie le niveau d'interactivité correspondant. Les valeurs possibles de cette propriété sont échelonnées sur cinq niveaux. La propriété `:hasLearningResourceType` permet de spécifier le type de la ressource d'apprentissage. Nous présentons dans ANNEXE N la représentation de quelques possibilités de valeurs pour cette propriété. Les types possibles sont représentés comme *individus nommés* (`owl:NamedIndividual`), sont membres de la classe `:LearningResourceType` et sont associés à une ressource d'apprentissage via la propriété `:hasLearningResourceType`.

5.3.3. Ontologie de l'objectif d'apprentissage

Conformément à la modélisation présentée dans le § 5.2.3, l'ontologie de l'objectif d'apprentissage permet de quantifier le niveau de l'apprentissage dans le cadre du système. Ainsi le portfolio des objectifs d'apprentissage acquis reflète le niveau d'évolution de l'apprenant dans son parcours d'apprentissage. L'acquisition de l'objectif d'apprentissage lui permet d'évaluer son état et d'influencer son futur parcours d'apprentissage.

Trois classes principales se distinguent dans notre modélisation. La classe `:GenericLearningObjective` est utilisée pour la déclaration des propriétés générales

d'un objectif d'apprentissage (:isPrerequisiteFor, :hasPrerequisite, :hasTitle, :hasDescription, :hasLevel). Les classes de l'objectif d'apprentissage (:LearningObjective), de l'objectif spécifique d'apprentissage (:SpecificLearningObjective) et de la compétence (:Competence) sont déclarées comme sous-classes de la classe générale (:GenericLearningObjective) et héritent ses caractéristiques (v. l'ANNEXE J).

La propriété :isPrerequisiteFor et son inverse :hasPrerequisite permettent de spécifier l'ordre d'apparence dans le parcours de l'apprenant des objets pédagogiques associés à ces objectifs. Les propriétés :isPartOf et :hasPart permettent la composition des objectifs d'apprentissage. Par exemple la compétence *La maîtrise de la langue française* est composée de : (1) la capacité à lire et comprendre des textes variés, (2) la qualité de l'expression écrite, (3) la maîtrise de l'expression orale, (4) l'apprentissage de l'orthographe et de la grammaire et (5) l'enrichissement quotidien du vocabulaire. Ces objectifs peuvent, à leur tour, avoir d'autres objectifs plus spécifiques dans leur composition. En conformité avec la modélisation des objectifs d'apprentissage, les différents niveaux associés aux compétences sont représentés en utilisant les propriétés présentées dans l'ANNEXE N. La propriété hasLevel permet de définir le niveau d'un objectif d'apprentissage sur une échelle spécifiée avec la propriété hasLevelScale. Par exemple, le niveau d'une compétence sur l'échelle de la taxonomie de Bloom est spécifié à travers cette propriété en associant l'objectif à un des six niveaux (e.g. se rappeler, comprendre, appliquer, analyser, évaluer, créer). La propriété :hasLearningObjectiveType (v. l'ANNEXE N) permet l'association d'un objectif spécifique à un type ou dimension de l'objectif (e.g. cognitive, psychomoteur, attitude).

5.3.4. Ontologie de l'objet d'évaluation

Ayant pour but la spécification des entités utilisées lors de l'évaluation de l'apprenant, cette ontologie spécifie les principaux objets, classes et caractéristiques intervenant dans la relation avec *l'objet d'évaluation*. Les principales classes sont présentées dans l'ANNEXE K. Les entités de types *objet d'évaluation* sont soumises à la restriction d'avoir déclaré le type de l'évaluation. Ceci est réalisé en utilisant la restriction d'existentialité owl:someValuesFrom sur la propriété :hasType. Par rapport à l'*ontologie du quizz*

présentée dans (R. Balog-Crisan, Roxin, et Szilagyi, 2009), et qui est basée sur deux objets spécifiques principaux (*question* et *quiz*), cette ontologie offre plus de flexibilité permettant l'extensibilité et la représentation de plusieurs types d'entités d'évaluation. Ces types sont spécifiés comme instances de la classe `:EvaluationObjectType`.

5.3.5. Ontologie de l'objet d'annotation

Les trois types d'objets d'annotation considérés sont représentés dans l'ontologie sous forme de trois classes distinctes associées à leur type. La classe `:AnnotationObjectType` englobe les différents types d'annotation (v. ANNEXE L). La classe `:Appreciation` permet une appréciation générale sur la ressource d'apprentissage. La classe `:Tag` permet de regrouper les tags appliqués sur les ressources et fournit des informations sur la perception de l'apprenant sur le sujet de la ressource d'apprentissage. La classe `:Note` permet aux apprenants d'enregistrer leur propres commentaires et notes sur les ressources d'apprentissage. La classe `:AppreciationCriterion` permet la spécification de différents critères d'appréciation. Ces types d'appréciation permettent l'augmentation ou la diminution de la visibilité d'une ressource (e.g. objet pédagogique, ressource d'information, question). Les types d'appréciation considérés sont les suivants : (1) générale ; (2) sur le contenu et (3) sur la forme de présentation. La corrélation entre ces appréciations et les profils des apprenants permet une meilleure adaptation et personnalisation du contenu d'apprentissage.

5.3.6. Ontologie du cadre d'enseignement

L'ontologie du cadre d'enseignement (v. ANNEXE M) permet d'enregistrer les différents contextes de l'apprentissage. Cette ontologie est basée sur la proposition présentée dans (Balog-Crisan, 2011), qui décrit les entités participantes à l'apprentissage selon le modèle réel (e.g. classification CNU) et ajoute des propriétés pour faciliter l'intégration dans le contexte en ligne (e.g. `CercitificationType`, `LearningType`, `LearningPath`). Les propriétés `containLearningResource` et `hasNext` (v. l'ANNEXE N) permettent d'associer et d'ordonner les ressources d'apprentissage dans les parcours d'apprentissage.

Ainsi la requête SPARQL présentée dans l'Extrait de code 5-6 permet d'obtenir la première ressource d'apprentissage d'un parcours d'apprentissage.

```
PREFIX          le:          <http://semlearn.pu-pm.univ-
fcomte.fr/ontologies/learning-environnement-ontology.owl#>

SELECT DISTINCT ?s
  WHERE {
    ?lp a le:LearningPath .
    ?lp le:containLearningResource ?s .
    ?s le:hasNext ?o .
    FILTER NOT EXISTS {?x le:hasNext ?s }
  }
```

Extrait de code 5-6. Requête SPARQL permettant l'identification de la première ressource d'apprentissage existant dans un parcours d'apprentissage

Les ressources d'apprentissage suivantes sont obtenues par le biais de la propriété `hasNext`, déclarée comme *fonctionnelle*, acceptant donc une seule valeur.

5.4. Ontologies au niveau d'application du SASA

Des ontologies spécifiques interviennent au niveau de l'application, notamment l'ontologie de liaisons et l'ontologie d'application. Au niveau de l'application, ces ontologies assurent l'accès du système informatique à l'ensemble des entités déclarées dans le système conceptuel du SASA.

5.4.1. Ontologie de liaison

L'ontologie de liaison permet de relier les concepts appartenant aux différentes ontologies (e.g. ontologies de domaine, ontologie d'application). Les équivalences entre ces concepts sont réalisées à trois niveaux. Au niveau des classes les équivalences sont réalisées en utilisant les propriétés suivantes :

- `rdfs:subClassOf` – contrainte partielle – affirme que toutes les instances d'une classe sont aussi les instances de la classe supérieure ;
- `owl:equivalentClass` – contrainte complète – affirme que deux classes contiennent exactement le même ensemble d'instances.

Au niveau des propriétés, les équivalences sont réalisées en utilisant les propriétés suivantes :

- `rdfs:subPropertyOf` – toutes les ressources reliées par une propriété sont aussi reliées par la propriété supérieure ;
- `owl:equivalentProperty` – les propriétés reliées ont les mêmes extensions (domaine et portée).

Au niveau des instances, les équivalences sont réalisées en utilisant la propriété `owl:sameAs`²⁸⁶. Elle affirme que deux instances sont identiques. L'utilisation de la propriété `owl:sameAs` pour les classes et pour les propriétés n'est pas acceptée en OWL DL. Traitant les classes comme instances cette propriété appliquée aux classes est acceptée seulement en OWL Full.

Les équivalences réalisées dans l'ontologie de liaison permettent au système conceptuel d'assimiler de nouvelles entités et définitions des concepts. L'intégration et l'harmonisation des ontologies dans le SASA sont réalisées en utilisant ces équivalences. Elles permettent *l'alignement* des ontologies et l'exploitation de l'ensemble de connaissances à travers une seule ontologie au niveau de l'application.

5.4.2. Ontologie d'application

L'interaction du système informatique d'apprentissage avec le système conceptuel d'apprentissage est réalisée au niveau d'un noyau sémantique à travers l'ontologie d'application (v. Figure 5-16).

²⁸⁶ Le site <http://sameas.org/> permet de trouver les équivalences entre les instances déclarées dans différentes bases de connaissances. Actuellement il compte 150 406 368 URIs qui représentent 53 054 359 d'entités uniques (« sameAs », s. d.).

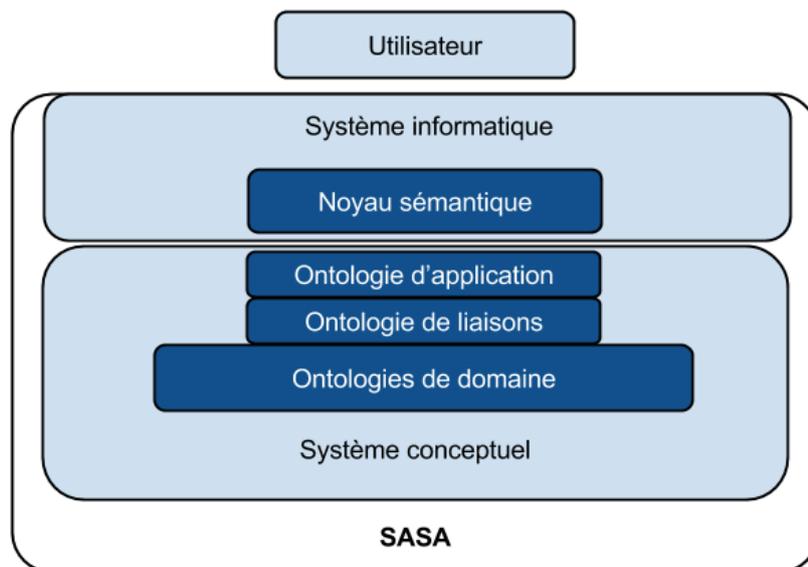


Figure 5-16. Relations entre les systèmes et les ontologies du SASA

Au niveau du noyau sémantique, le moteur d'inférence permet de représenter les relations définies dans les ontologies de domaine en utilisant le vocabulaire de l'ontologie d'application. Ainsi au niveau du système informatique, les requêtes SPARQL sont écrites en utilisant les termes définis dans l'ontologie d'application. Elle fournit des termes spécifiques à la fonctionnalité offerte par le système informatique. Elle est modélisée et peut évoluer en prenant compte des fonctionnalités au niveau du système informatique d'apprentissage.

Chapitre 6. Implémentation du SASA

« Le meilleur moyen de prédire le futur est de
l'inventer »
Alan KAY

La modélisation et les ontologies du système conceptuel d'apprentissage présentées dans le Chapitre 5 sont exploitées par le SASA. Les systèmes informatiques sont particulièrement intéressants pour la résolution des problèmes de recherche d'information ou de tri des données à partir des critères, ou encore dans le cas de systèmes informatiques à base de connaissances, pour rendre explicite la connaissance implicite d'un ensemble des connaissances selon des règles spécifiques.

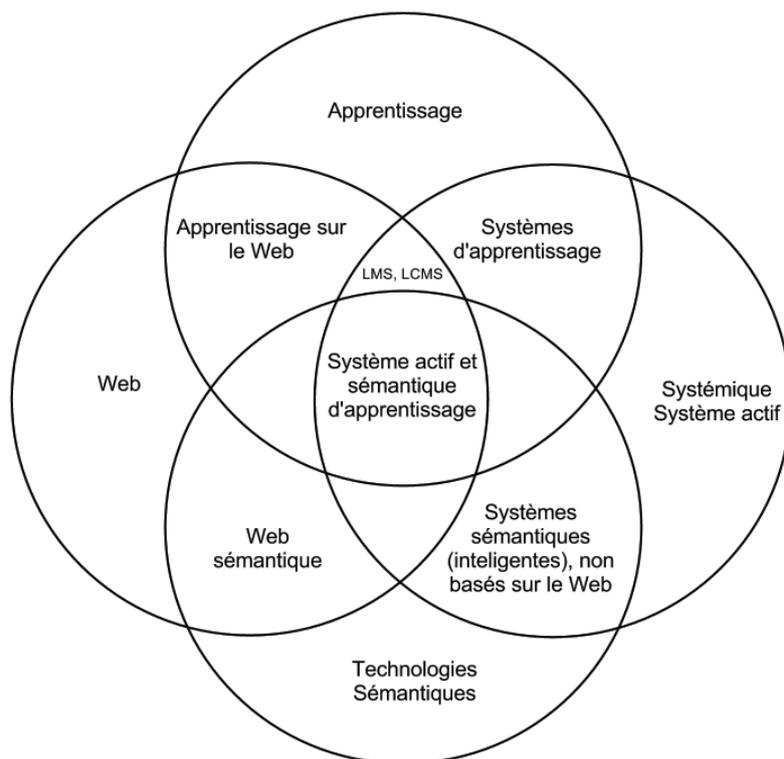


Figure 6-1. Le SASA à l'intersection de l'apprentissage, de la systémique et du Web sémantique

Le SASA se trouve à l'intersection de plusieurs domaines et approches technologiques (v. Figure 6-1). L'objectif du SASA est d'améliorer l'expérience d'apprentissage en fournissant un parcours d'apprentissage persistant et adapté aux besoins de l'apprenant.

SASA accompagne l'apprenant dans son expérience d'apprentissage, tout en déchargeant sur la machine une partie du travail intellectuel humain de l'enseignant ou du tuteur humain. Dans cette perspective, le SASA est un *agent intelligent et personnel d'apprentissage*. Il est intelligent parce qu'il fournit des fonctions d'inférence sur l'ensemble des données enregistrées, et personnel parce qu'il exploite une modélisation de l'apprenant. Notre système est *actif*, parce qu'il modifie et adapte le parcours d'apprentissage selon des caractéristiques de l'apprenant. Son but est de stimuler les activités cognitives de l'apprenant à travers la personnalisation du contenu d'apprentissage. Imaginons que le système identifie une difficulté de l'apprenant par rapport à un objectif d'apprentissage. Dans ce cas, le système génère des parcours d'apprentissage enrichis de sorte que l'apprenant dépasse cette difficulté. Le but du SASA dans ce cas est la réussite de cet objectif d'apprentissage par l'apprenant.

Pour atteindre son but, le système d'apprentissage interprète les actions de l'apprenant selon les modèles spécifiés et, ensuite, il propose une personnalisation du parcours d'apprentissage en utilisant des mécanismes d'adaptation. Le système d'apprentissage trouve sa place dans le rôle de médiateur, d'accompagnateur de l'apprenant dans son expérience d'apprentissage.

Dans le cadre d'un apprentissage formel, le système d'apprentissage est un assistant pour *l'enseignant*. Il assiste l'enseignant dans son rôle de médiateur tout en prenant en charge une partie de ses fonctions. L'enseignant est très présent dans la phase préparatoire du cours (e.g. objets pédagogiques, parcours d'apprentissage), mais après le commencement du cours, il prend un rôle d'observation, il se concentre sur le suivi de l'apprenant.

Dans le cadre de l'apprentissage informel, le système prend en charge toutes les fonctions de l'enseignant. Le système d'apprentissage « organise » les cours, en choisissant les objets pédagogiques et en établissant l'ordre de ces objets dans le parcours d'apprentissage, en concordance avec les objectifs de l'apprenant.

6.1. Organisation du système

Dans cette section nous présentons les principaux composants du SASA ainsi que leur organisation au sein du système. Selon les tâches à effectuer par chaque composant du système, nous structurons le système en plusieurs niveaux. Nous présentons le noyau du système, ainsi que les modules intégrant l'ensemble des connaissances.

6.1.1. L'architecture générale du SASA

Selon la nature et le rôle des différents composants du système d'apprentissage, nous avons structuré l'architecture générale du SASA en trois niveaux distincts (v. Figure 6-2).

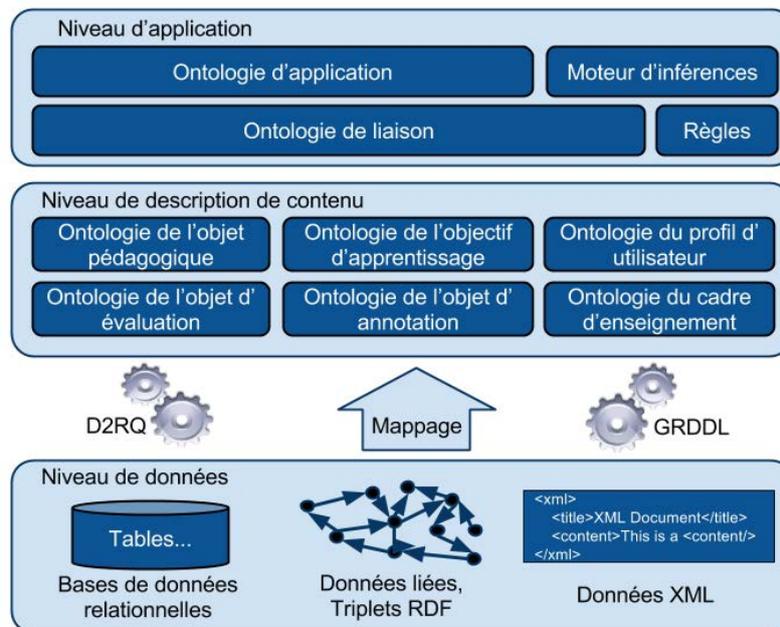


Figure 6-2. L'architecture sur niveaux du SASA

Le premier niveau, le niveau de données, comprend les *données effectives*. À ce niveau les données peuvent être représentées sous forme de triplets RDF, structurées en format XML ou dans des bases de données relationnelles. Afin de rendre ces données exploitables par le SASA, il est nécessaire de les transformer au format RDF. Ainsi le langage déclaratif D2RQ²⁸⁷ permet d'associer le modèle de données d'une base de données relationnelle aux termes définis dans un vocabulaire RDFS ou dans une

²⁸⁷ La plateforme D2RQ est composée de trois parties : (1) le langage de mappage D2RQ, (2) le moteur D2RQ (plug-in pour la librairie Jena) et (3) un serveur D2R (qui permet de requêtes HTTP). <http://d2rq.org/>

ontologie. Suivant cette association et l'exécution de la transformation, les données provenant de la base de données peuvent être représentées en format RDF. De la même manière, le langage GRDDL ²⁸⁸ (*Gleaning Resource Descriptions from Dialects of Languages*) permet d'extraire les données représentées au format XML et de les exprimer en triplets RDF. Les règles de transformation seront explicitées dans un fichier de transformation XSLT (eXtensible Stylesheet Language Transformations). A travers ces deux langages (D2RQ et GRDDL), nous pouvons transformer en triplets RDF, non seulement les données présentes dans le premier niveau du système mais aussi les données fournies par les différents services Web.

Le deuxième niveau de l'architecture du SASA assure la description du contenu. Il englobe les vocabulaires, les schémas et les ontologies utilisés dans le cadre du système. Les données intégrées au premier niveau du SASA sont interprétées à travers les définitions des termes spécifiés dans le deuxième niveau. Ce niveau est principalement constitué de six modules intégrant les modélisations des entités utilisées dans le SASA.

Enfin, le troisième niveau, celui d'application, intègre l'ontologie d'application, l'ontologie de liaison, les règles et le moteur d'inférence. L'ontologie de liaison permet l'agrégation des modèles définis dans les différents modules et les données associées à ces modèles. La spécification des règles à ce niveau permet au moteur d'inférence d'explicitier des connaissances composées à partir de plusieurs modules.

Une architecture d'un système capable d'intégrer des ressources d'apprentissage à partir de plusieurs LOR (Learning Object Repository) est déjà présentée par P. Vidal et J. Broisin. Ils présentent une architecture sur quatre niveaux qui permet de fédérer des ressources d'apprentissage à partir de différents LOR (Learning Object Repository). Les quatre niveaux présentés sont : (1) interface pédagogique, (2) couche de médiation, (3) services de fédération des LORs et (4) stockage des ressources pédagogiques (Vidal et Broisin, 2005). La couche de fédération assure l'agrégation des ressources pédagogiques, alors que la couche de médiation permet aux différentes plateformes d'apprentissage d'exploiter ces ressources à travers des API spécifiques. Ils anticipent l'utilisation des technologies du Web sémantique dans ces systèmes : « Actuellement en

²⁸⁸<http://www.w3.org/TR/grddl/>

plein essor, une interopérabilité entre le Web sémantique et les LOR augmenterait de façon incommensurable le nombre de ressources disponibles » (ibid.).

6.1.2. Les modules du SASA

L'approche modulaire du SASA permet l'administration individuelle de chaque modélisation. Les modules du système, correspondant aux six ontologies déclarées, permettent l'administration de l'ensemble des connaissances associées. Un septième module permet l'intégration de données externes au système. La Figure 6-3 présente les modules du système d'apprentissage. Chaque module peut être vu comme une base de connaissances. Il contient les connaissances terminologiques déclarées dans l'ontologie correspondante, ainsi qu'un ensemble d'instances et de ressources en tant que connaissances assertionnelles. Ainsi, le module des objets pédagogiques permet d'enregistrer les données et les métadonnées des *ressources d'apprentissage* selon les descriptions définies dans l'*ontologie de l'objet pédagogique*.

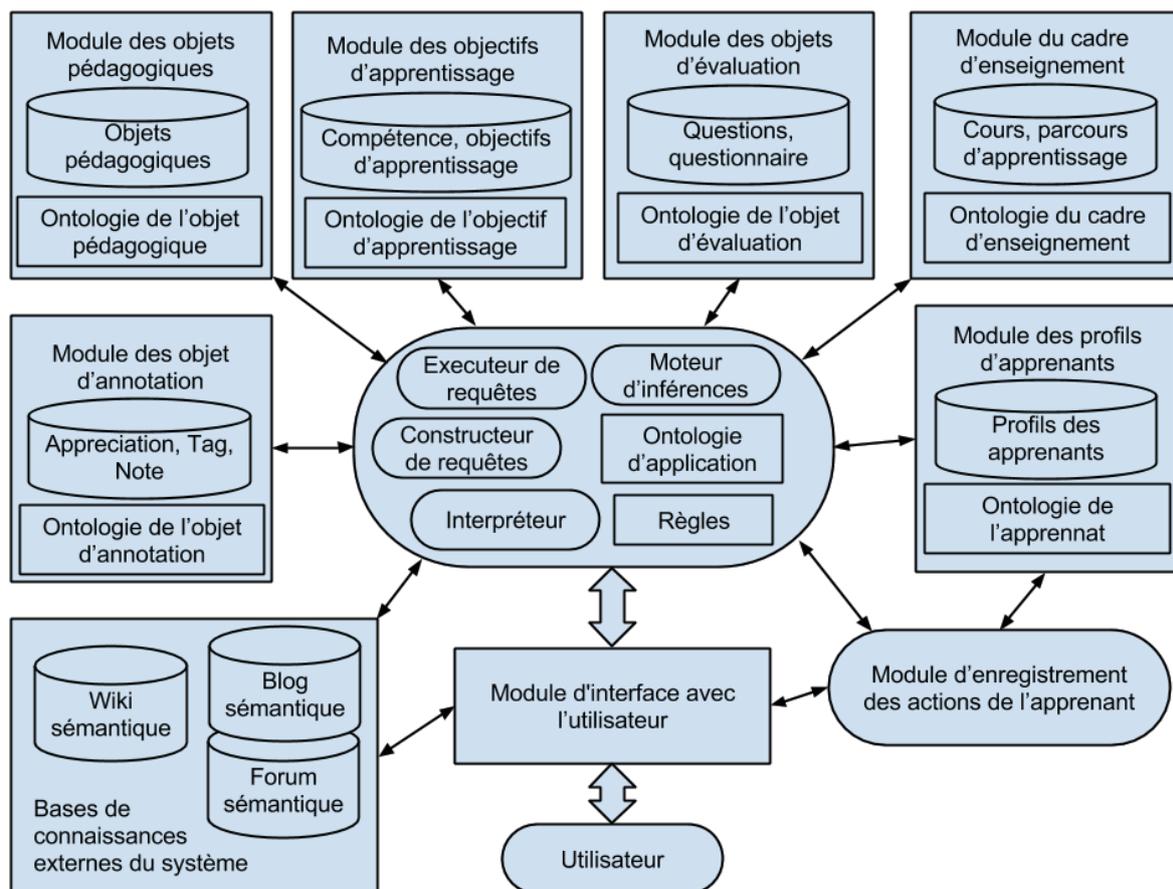


Figure 6-3. Les modules du SASA

Toutes les données enregistrées dans la « boîte d’assertions » correspondant à ce module, devraient être représentées en utilisant le vocabulaire défini dans l’ontologie associée. Chaque module du système peut intégrer un moteur d’inférence propre. Ainsi, les tâches d’inférences sont distribuées au niveau du système. Au centre du système, *le noyau sémantique*, permet de relier les modules du système et d’effectuer des requêtes sur l’ensemble des connaissances.

6.1.3. Le noyau sémantique : un agent intelligent pour l’apprentissage

La partie centrale du SASA est le *noyau sémantique*. Il détient la logique du système et assure la communication entre l’utilisateur et les différents modules. Le module d’interface utilisateur assure l’échange d’information et l’interaction de l’utilisateur avec le noyau sémantique. À travers cette interface, un objet pédagogique interactif peut initier une action afin de lancer une requête dans la base de connaissances.

Au niveau du noyau sémantique, l’interprétation des actions de l’apprenant ou des requêtes lancées par l’objet pédagogique est réalisée par la composante *interprète*. Cette composante du noyau sémantique évalue les actions de l’utilisateur et fournit les détails pour la construction de la requête à exécuter sur la base de connaissances. Il réalise l’association entre les signaux reçus du module d’interface et les termes utilisés pour la construction de la requête SPARQL.

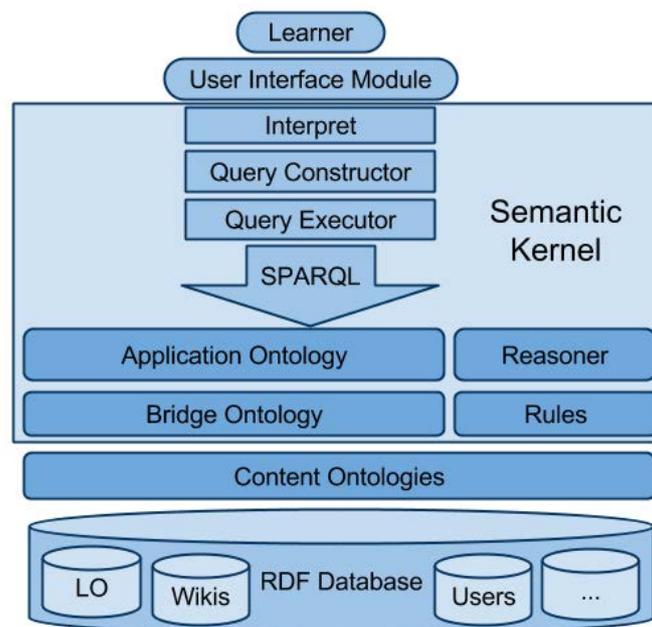


Figure 6-4. Le noyau sémantique

Le constructeur de requêtes construit dynamiquement les requêtes selon l'interprétation des signaux reçus. Ainsi deux catégories de requêtes se distinguent conformément au langage SPARQL : les requêtes qui nécessitent d'exécuter une écriture dans les bases de connaissances (e.g. requêtes de type DELETE ou INSERT) ou les requêtes qui exécutent une lecture des données (e.g. SELECT, ASK). Enfin, l'*exécuteur de requêtes* lance la requête sur la base de connaissances en utilisant les termes définis dans l'ontologie d'application. En implémentant SPARQL, le noyau sémantique présente l'avantage de fournir un accès direct à la base des connaissances pour l'ensemble des agents (e.g. objets pédagogiques) qui constituent l'environnement où se déroule l'apprentissage. Ainsi, l'évolution et les traces de l'apprenant peuvent être enrichies et exploitées à partir de plusieurs agents. Les ontologies d'application et de liaison sont exploitées au niveau du noyau sémantique. L'ontologie de liaison permet de *connecter* et de *mapper* les termes déclarés dans les différentes ontologies à l'ontologie d'application.

Le noyau sémantique englobe la logique du système. L'agrégation de l'ensemble des connaissances, l'exploitation de modélisations des objets participants à l'apprentissage ou la spécification des règles spécifiques pour l'adaptation du parcours d'apprentissage rendent intelligent le système d'apprentissage.

6.2. SASA : personnalisation du parcours d'apprentissage

En ajoutant des règles spécifiques au niveau du noyau sémantique, le moteur d'inférence permet de modifier (e.g. ajouter, supprimer) des relations entre les entités définies dans la base de connaissances. La modification de ces relations modifie *l'état du système conceptuel* et provoque des changements dans l'organisation du système. Les principaux changements de l'organisation du système qui influencent l'apprenant sont sur son *parcours d'apprentissage*. Concrètement, la personnalisation de l'apprentissage consiste en l'adaptation des ressources d'apprentissage composant son parcours. Pour adapter le parcours d'apprentissage, ces ressources d'apprentissage sont choisies en fonction de ses *objectifs d'apprentissage, ses objectifs atteints, son style d'apprentissage et les caractéristiques des ressources d'apprentissage*.

La réorganisation du système conceptuel suite aux excitations extérieures (e.g. changements dans le profil de l'apprenant, l'ajout des ressources d'apprentissage) et conformément aux règles déclarées indique la capacité du système d'agir (système *actif*). En fonction du moment de l'exécution des actions par rapport à l'initiative de l'apprenant, deux facettes du système se distinguent (v. Figure 6-5).

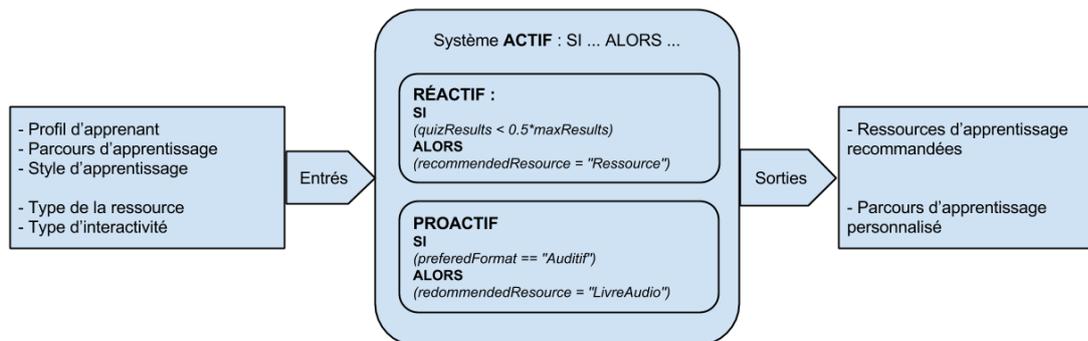


Figure 6-5. Système actif d'apprentissage

Le système est *réactif* lorsqu'il modifie le parcours d'apprentissage *suite* aux actions de l'apprenant. C'est le cas des modifications du profil d'apprenant suite à la réalisation d'objectifs d'apprentissage, à la complétion de questionnaires ou d'examens. Ces modifications du profil d'apprenant permettent d'identifier les forces et les faiblesses de l'apprenant, utilisées dans l'adaptation du parcours d'apprentissage.

Le système est *proactif* puisqu'il modifie le parcours d'apprentissage en se basant sur des règles préalablement définies, sans attendre que l'utilisateur le demande. Le système associe des ressources d'apprentissage aux profils d'apprenants selon les caractéristiques décrites dans leurs modélisations et conformément aux règles définies. C'est le cas de la personnalisation du parcours d'apprentissage conforme au style d'apprentissage de l'apprenant. Le système d'apprentissage *anticipe* et identifie les ressources d'apprentissage adaptées et procède à la modification du parcours.

Du fait de ces types d'actions, le SASA est un système de recommandation basé sur le contenu. La définition et la conception de règles et d'algorithmes de recommandation sortant du cadre de cette thèse, nous présentons ici seulement quelques scénarios de base pour la personnalisation du parcours d'apprentissage.

6.2.1. Scénarios pour la personnalisation du parcours d'apprentissage

Une première approche consiste à identifier le parcours d'apprentissage de l'apprenant, qui peut être relevé à différents niveaux. L'ensemble des objets pédagogiques que l'apprenant a parcouru constitue son parcours d'apprentissage au niveau des ressources d'apprentissage. L'ensemble des objectifs d'apprentissage qu'il a atteint constitue son parcours d'apprentissage au niveau des compétences. Dans la suite nous donnons un exemple en considérant le parcours d'apprentissage en termes d'*objectifs atteints*.

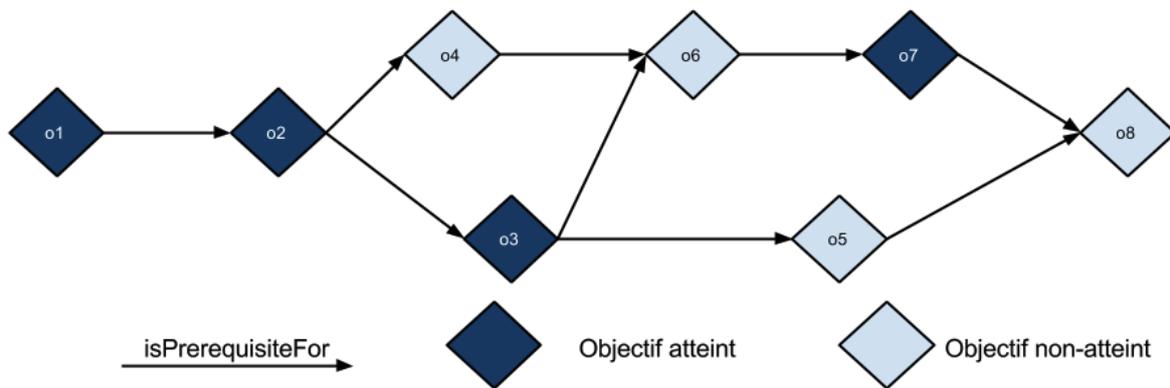


Figure 6-6. Un exemple de graphe RDF avec d'objectifs d'apprentissage

Considérons les objectifs d'apprentissage et leurs relations de *prérequis* présentées sous forme de graphe dans la Figure 6-6. Les objectifs atteints sont représentés avec une couleur foncée et les objectifs non-atteints avec une couleur plus claire. Considérons le terme `:isPrerequisiteFor` pour définir les relations de prérequis entre les objectifs d'apprentissage. Une première requête consiste à trouver tous les objectifs d'apprentissage prérequis pour accéder à l'objectif d'apprentissage désiré. Considérons l'objectif désiré `res:objectif8`, la requête SPARQL qui permet de répondre à la question : « Quelles sont (tous) les prérequis nécessaires pour atteindre l'objectif d'apprentissage `res:objectif8` » est la suivante :

```
SELECT DISTINCT ?objectifs_all
WHERE {
res:objectif8 ^co:isPrerequisiteFor+ ?objectifs_all .
}
```

Extrait de code 6-1. Exemple de requête²⁸⁹ SPARQL - trouver les objectifs d'apprentissage prérequis

Etant donnée la boîte d'assertion proposée, la réponse à cette requête consiste en une liste de tous les objectifs d'apprentissage. La requête suivante permet de retrouver la liste des objectifs atteints par un apprenant²⁹⁰ :

```
SELECT DISTINCT ?objectifs_acquired
WHERE {
res:apprenant co:hasAcquired ?objectifs_acquired .
}
```

Extrait de code 6-2. Exemple de requête SPARQL - trouver les objectifs d'apprentissage atteints

Considérant le graphe dans la Figure 6-6, cette requête va retourner les objectifs 1, 2, 3 et 7. La requête ci-dessous permet d'obtenir la liste des objectifs d'apprentissage qui sont *pré requis* pour l'objectif désiré `res:objectif8`, mais *non-atteints* par l'apprenant (définie comme la différence entre les deux listes d'objectifs précédentes).

```
SELECT DISTINCT ?objectif
WHERE {
res:objectif8 ^co:isPrerequisiteFor+ ?objectif .
FILTER NOT EXISTS { res:apprenant co:hasAcquired ?objectif .
}
}
```

Extrait de code 6-3. Exemple de requête SPARQL – trouver les objectifs d'apprentissage prérequis et non atteints

La liste des objectifs d'apprentissage non-atteints par l'apprenant et auxquels il peut donc accéder est obtenue avec la requête suivante (nous retournons les objectifs non-atteints pour lesquels l'apprenant a validé au moins un des pré-requis) :

²⁸⁹ Le signe + de la propriété `co:isPrerequisiteFor` permet de trouver toutes les entités qui peuvent être liées à partir du `res:objectif8` en utilisant cette propriété, alors que le signe ^ permet *d'inverser le chemin*.

²⁹⁰ Considérons que la propriété `co:hasAcquired` permet d'associer un apprenant aux objectifs déjà atteints.

```

SELECT DISTINCT ?objectif
WHERE {

res:apprenant co:hasAcquired ?pre .
?pre co:isPrerequisiteFor ?objectif .

FILTER NOT EXISTS { res:apprenant co:hasAcquired ?objectif . }
}

```

Extrait de code 6-4. Exemple de requête SPARQL - trouver les objectifs d'apprentissage non-atteints mais accessibles

Le choix des objectifs d'apprentissage permet ensuite de proposer les ressources d'apprentissage correspondantes, et les intégrer dans le parcours d'apprentissage. Nous présentons dans l'ANNEXE O des exemples qui illustre comment est obtenue la liste des objets pédagogiques correspondant aux objectifs d'apprentissage auxquels l'apprenant peut accéder. Par la suite la liste des objets d'évaluation associés à ces objectifs d'apprentissage permet de vérifier l'état d'avancement de l'apprenant. La validation de l'accomplissement d'un objectif d'apprentissage déclenche le changement du profile d'apprenant, donc des nouvelles modifications dans la liste des ressources d'apprentissage à parcourir.

6.2.2. Système réactif : adaptation basée sur les résultats de l'apprenant

Les évaluations permettent de modifier le profil de l'apprenant en y ajoutant les compétences acquises (objectifs d'apprentissage atteints) et les nouveaux objectifs d'apprentissage. Les résultats donnés par l'exécution d'un *objet d'évaluation* permettent de vérifier le niveau d'accomplissement d'un objectif d'apprentissage (e.g. le rapport entre le résultat obtenu et le maximum de points possible d'un objet d'évaluation). Toutefois, l'objet d'évaluation peut spécifier un seuil à partir duquel l'objectif vérifié est considéré comme atteint. Selon la granularité de l'objet d'évaluation et ses relations déclarées avec les objectifs d'apprentissage, l'identification des sujets spécifiques fournira avec pertinence, les points faibles et, de la même manière, les points forts d'un apprenant.

```

res:objectif5 eo:isValidatedByEvaluationObject res:evalObj1 .
res:evalObj1 eo:hasEvaluationPointsMax 20 ;
eo:hasMinValidationPoints 10 .

```

Extrait de code 6-5. La représentation de la relation entre l'objectif d'apprentissage et l'objet d'évaluation

Les relations présentées dans l'Extrait de code 6-5 affirment que l'objectif `res:objectif5` est validé par l'objet d'évaluation `res:evaObj1`, par l'obtention d'un nombre minimum de points déclaré à l'aide de la propriété `eo:hasMinValidationPoints` et d'une manière analogue pour l'objectif d'apprentissage `res:objectif6` et `res:objectif2`.

L'association de l'objectif d'apprentissage avec l'objet pédagogique à travers la relation `eo:isValidatedByEvaluationObject`, et l'association de l'objectif d'apprentissage avec l'objet pédagogique par la relation `lo:hasLearningObjective` permettent d'établir une relation entre l'objet d'évaluation et l'objet pédagogique. En outre, les enregistrements des évaluations d'un apprenant sont réalisées de la manière présentée dans l'Extrait de code 6-6. Cela permet d'enregistrer aussi le moment de l'évaluation (`eo:evaluationDate`) et le résultat obtenu lors de l'évaluation (`eo:obtainedPoints`).

```
res:apprenant eo: hasAttended [
    eo:attendedEvaluationObject res:evalObj5 ;
    eo:startedTime "2013-10-15T10:00:00Z"^^xsd:dateTime ;
    eo:completedTime "2013-10-15T12:00:00Z"^^xsd:dateTime ;
    eo:obtainedPoints 17
] .
```

Extrait de code 6-6. La représentation des relations de l'apprenant à des objets d'évaluation

Les différentes requêtes SPARQL sont construites afin de répondre aux fonctionnalités implémentées au niveau de l'application. La requête présentée dans l'Extrait de code 6-7 permet d'obtenir la liste des objets d'évaluation que l'apprenant a parcourus sans validation.

```
SELECT ?evalObj
WHERE {
res:apprenant eo:hasAttended ?bObj .
?bObj eo:attendedEvaluationObject ?evalObj .
?bObj eo:obtainedPoints ?obtainedPoints .

?evalObj eo:hasMinValidationPoints ?minValPoints .
FILTER (?obtainedPoints < ?minValPoints)
}
```

Extrait de code 6-7. Exemple requête SPARQL - liste des objets d'évaluation que l'apprenant n'a pas validés

Le système d'apprentissage est capable de fournir une liste des ressources d'apprentissage en relation avec l'objectif d'apprentissage que l'apprenant n'a pas validé. L'enrichissement avec des métadonnées (e.g. tags) d'un contenu d'apprentissage (e.g. objet pédagogique, objet d'évaluation) permet d'identifier davantage de ressources associées. D'autres requêtes peuvent être construites pour identifier les points faibles des apprenants, ou encore les objets d'évaluation que plusieurs apprenants n'ont pas validés. Ces indicateurs permettent de déterminer les ressources d'apprentissage efficaces et d'améliorer la sélection des contenus d'apprentissage. Le système d'apprentissage proposé étant dans la phase de prototype expérimental, les données renseignées dans sa base de connaissances sont fictives. Toutefois, les descriptions des ressources d'apprentissage dans le SASA combinées avec la flexibilité du langage SPARQL permettent l'amélioration de l'activité du système en rapport avec l'apprenant.

6.2.3. Système proactif : adaptation basée sur des préférences prédéfinies

Nous présentons dans ce qui suit le fonctionnement *proactif* du SASA, selon les actions de l'apprenant. Le système *anticipe* les possibilités d'amélioration du parcours d'apprentissage et fournit des suggestions adaptées avant que l'apprenant interagisse avec le contenu d'apprentissage. Ses suggestions sont basées sur : a) les informations concernant le style d'apprentissage ou les préférences de l'apprenant (modèle d'apprenant) ; b) les caractéristiques concernant la présentation des ressources d'apprentissage (modèle de la ressource d'apprentissage) ; c) les associations entre les caractéristiques des ressources d'apprentissage et les caractéristiques des apprenants (appréciations données par les apprenants aux ressources d'apprentissage).

Pour donner un exemple, nous choisissons le style d'apprentissage Felder-Silverman, proposé dans la modélisation du profil d'apprenant. En conformité avec ce style d'apprentissage, l'apprenant est caractérisé par quatre dimensions (v. § 1.2.3) : (1) participation de l'apprenant (actif – réflexif), (2) perception (sensoriel – intuitif), (3) format (visuel – verbal) et (4) ordre de présentation (séquentiel – globale) (v. Figure 6-8).

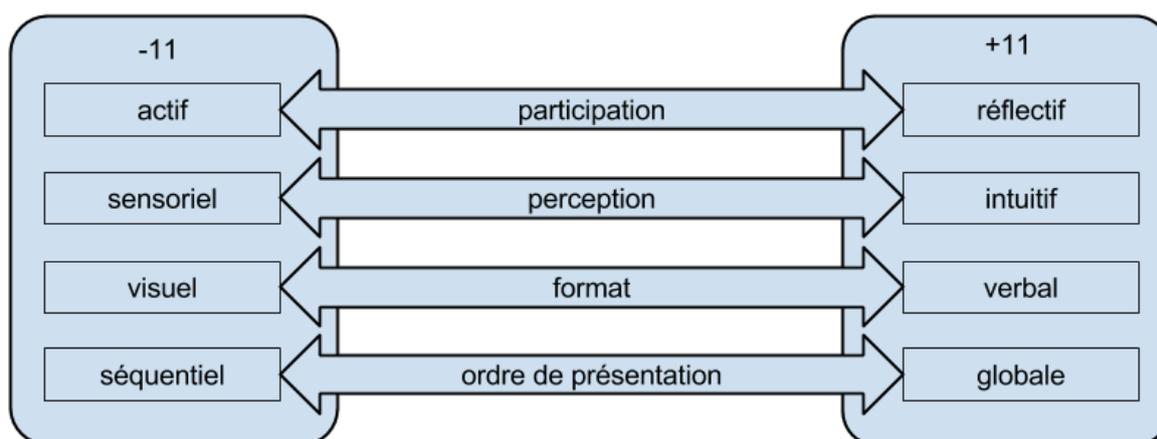


Figure 6-7. Dimensions du style d'apprentissage Felder-Silverman

Nous choisissons d'utiliser le questionnaire²⁹¹ fourni par Barbara A. Soloman et Richard M. Felder pour construire le vecteur du style d'apprentissage de l'apprenant. Ainsi, chaque dimension de ce vecteur peut prendre une valeur entre -11 et +11 (v. Figure 6-7). Le style d'apprentissage d'un apprenant peut être représenté sous la forme suivante :

$$V_a = [d_1 d_2 d_3 d_4], d_i \in [-11, 11], i \in \{1, 2, 3, 4\} \text{ où :}$$

d_1 : participation (actif – réflectif)

d_2 : perception (sensoriel – intuitif)

d_3 : format (visuel – verbal)

d_4 : ordre de présentation (séquentiel – globale)

Afin d'associer les styles d'apprentissage aux ressources d'apprentissage, celles-ci seront décrites en utilisant des caractéristiques pertinentes correspondant aux dimensions du style d'apprentissage utilisé. Ces caractéristiques doivent être *en concordance* avec les dimensions du style d'apprentissage. Pour calculer la compatibilité (*le niveau d'adéquation*) entre une ressource d'apprentissage et un style d'apprentissage, nous considérons quatre dimensions conformément aux quatre dimensions du style d'apprentissage Felder-Silverman. Notons donc le niveau d'adéquation d'une ressource d'apprentissage sous forme de vecteur de la même manière que les dimensions du style d'apprentissage²⁹².

²⁹¹ Le questionnaire se trouve à l'adresse <http://www.engr.ncsu.edu/learningstyles/ilsweb.html>

²⁹² Dans le cas de la dimension de *présentation* -11 représente l'adéquation pour la valeur *séquentiel*, et +11 l'adéquation pour la valeur *globale*)

$$V_r = [a_1 a_2 a_3 a_4], a_i \in [-11, +11], i \in \{1, 2, 3, 4\} \text{ où :}$$

a_1 : adéquation pour la dimension participation (actif – réflexif)

a_2 : adéquation pour la dimension perception (sensoriel – intuitif)

a_3 : adéquation pour la dimension format (visuel – verbal)

a_4 : adéquation pour la dimension ordre de présentation (séquentiel – globale)

Afin de trouver les entités « proches » nous proposons de calculer la *distance euclidienne*. Cette méthode peut être utilisée aussi pour calculer les distances entre les ressources d'apprentissage et les styles d'apprentissage des apprenants (Sangineto *et al.*, 2008; Limongelli *et al.*, 2009). Ainsi la *distance* entre une ressource d'apprentissage, caractérisée par le vecteur V_R , et un style d'apprentissage de l'apprenant, caractérisé par le vecteur V_A , est calculée en utilisant la formule suivante :

$$E_{A-R} = \sqrt{(d_1 - a_1)^2 + (d_2 - a_2)^2 + (d_3 - a_3)^2 + (d_4 - a_4)^2}$$

Selon cette formule, et en conformité avec les valeurs proposées pour les dimensions représentées (entre -11 et 11), la distance euclidienne maximale peut être de 44. Plus la distance est petite, plus les entités s'approchent indiquant une bonne adéquation de la ressource d'apprentissage au style d'apprentissage de l'apprenant.

Tableau 6-1. La « distance » entre les ressources pédagogiques et le style d'apprentissage

				Participation	-11	-10	11							
				Perception	4	-7	7							
				Format	-9	-8	9							
				Présentation	8	-7	8							
Participation	Perception	Format	Présentation	Apprenants	Anne	Bruno	Claude							
				Ressources										
				-11				0	-10	8	LO1	4.12	16.70	29.90
				11				8	0	4	LO2	24.43	29.17	9.90
-2	-1	4	-5	LO3	21.07	15.75	20.66							

Dans le Tableau 6-1 nous présentons quelques exemples de calcul de la distance euclidienne entre trois objets pédagogiques et trois styles d'apprentissage correspondant aux trois apprenants. Les résultats indiquent que l'objet pédagogique

L01 est le plus adapté au style d'apprentissage d'*Anne*, *L02* est adapté pour *Claude* et *L03* correspond au style d'apprentissage de *Bruno*. Ces associations prennent en compte uniquement le style d'apprentissage de l'apprenant et les niveaux d'adéquation des objets pédagogiques aux différentes dimensions du style d'apprentissage.

L'expression peut être modifiée en intégrant un indice (W) afin de pondérer l'importance de chaque dimension dans le résultat final. Considérons le vecteur d'indice :

$$W = [w_1 w_2 w_3 w_4], \quad 0 \leq w_i \leq 1, \quad i \in \{1, 2, 3, 4\} \text{ où :}$$

w_1 : l'importance pour la dimension participation (actif – réflexif)

w_2 : l'importance pour la dimension perception (sensoriel – intuitif)

w_3 : l'importance pour la dimension format (visuel – verbal)

w_4 : l'importance pour la dimension ordre de présentation (séquentiel – globale)

Pour calculer la distance entre le style d'apprentissage d'un apprenant et une ressource d'apprentissage en pondérant les différentes dimensions on utilise la formule suivante :

$$E_{A-R} = \sqrt{(d_1 - a_1)^2 \cdot w_1 + (d_2 - a_2)^2 \cdot w_2 + (d_3 - a_3)^2 \cdot w_3 + (d_4 - a_4)^2 \cdot w_4}$$

Donnant la valeur 1 aux indices d'importance, on retrouve la formule antérieure de la distance euclidienne. Toutefois, assignant à un indice d'importance une valeur inférieure à 1 (ou 0) diminue (ou néglige) la dimension associée à cet indice dans le résultat final.

D'une part ce modèle de personnalisation du parcours d'apprentissage implique l'existence des descriptions du style d'apprentissage de l'apprenant et des valeurs d'adéquation des ressources d'apprentissage correspondant à ce style d'apprentissage. Pour l'identification du style d'apprentissage d'un apprenant il existe deux approches : a) les méthodes explicites (e.g. questionnaire de Solomon et Felder) ; b) les méthodes implicites, basées sur des statistiques et analyses observées sur l'apprenant (Popescu, Badica, et Moraret, 2010, p. 453). Une méthode implicite permettra la mise à jour dynamique du style d'apprentissage, fur et à mesure de l'interaction de l'apprenant dans le cadre du système.

L'enregistrement du style d'apprentissage de l'apprenant dans la base de connaissances est réalisé en utilisant les propriétés `up:hasLearningStyleDimensionValue`, `up:hasLearningStyle`, `up:hasDimension`, `up:hasDimensionValue` (v. ANNEXE N). Ainsi les valeurs et les dimensions du style d'apprentissage peuvent être interrogées avec une requête SPARQL de type :

```
SELECT DISTINCT ?dimensionValue ?dimension ?time
WHERE {
res:apprenant up:hasLearningStyleDimensionValue ?ls .
?ls          up:hasDimension          ?dimension ;
            up:hasDimensionValue     ?dimensionValue ;
            up:atTime                 ?time .
}
```

Extrait de code 6-8. Exemple requête SPARQL - dimensions et valeurs du style d'apprentissage

D'autre part, la modélisation des ressources d'apprentissage doit fournir les caractéristiques équivalentes des dimensions des styles d'apprentissage. La modélisation de l'objet pédagogique dans la forme proposée par IEEE ou LRMI permet l'enregistrement des caractéristiques suivantes : le type et le niveau d'interactivité ; le type de la ressource d'apprentissage ; la structure ; la densité sémantique ; la difficulté ou la tranche typique d'âge. Si la dimension *participation* du style d'apprentissage peut être naturellement associée au *niveau d'interactivité* d'un objet pédagogique, d'autres dimensions, comme *le format*²⁹³, *la perception* ou *l'ordre de présentation* ne trouvent pas d'équivalences explicites dans la modélisation de l'objet pédagogique. Pour éliminer cette insuffisance (Carver *et al.*, 1999) demande aux enseignants de renseigner, sur une échelle de 0 à 100, l'adéquation de chaque catégorie de ressource d'apprentissage (e.g. hypertexte, fichiers audio, fichiers graphiques) aux trois dimensions du style d'apprentissage du Felder-Silverman (perception, format et ordre de présentation). (Sanginetto *et al.*, 2008) et (Limongelli *et al.*, 2009). Il était aussi demandé aux enseignants d'annoter avec un *poids* indiquant le niveau d'adéquation des ressources d'apprentissage pour chaque dimension du style d'apprentissage de Felder-Silverman.

²⁹³ La caractéristique 4.1. *Format* du standard LOM spécifie le format d'un point de vue *technique* (e.g. *format du fichier*), alors que la dimension *format*, utilisée dans le style d'apprentissage, caractérise le format selon son « canal » de transmission (e.g. visuel, verbal). Cependant la version française du LOM, introduit une nouvelle caractéristique appelée *Type documentaire*. Elle est située dans la catégorie 1. Général, avec le chiffre d'indexation 1.9.

L'adéquation d'un objet pédagogique à un style d'apprentissage spécifique est enregistrée en utilisant les propriétés :

```
lo:hasSuitabilityForLearningStyleDimensionValue,  
lo:suitabilityDimension et lo:suitabilityValue.
```

Certaines de ces caractéristiques peuvent être déduites en utilisant des statistiques basées sur l'annotation et le filtrage collaboratif, ainsi que des règles d'inférences. Le style d'apprentissage d'un apprenant, son annotation sur une ressource d'apprentissage, ou encore ses résultats suite à l'accomplissement d'une ressource d'apprentissage, permet de rendre plus visible les niveaux d'adéquation de la ressource en relation avec le style d'apprentissage de l'apprenant. Nous présentons quelques règles simples implémentées au niveau du noyau sémantique. Ces règles permettent d'établir le niveau d'adéquation d'un objet pédagogique à la dimension concernant le *format* préféré déclaré dans le style d'apprentissage Felder-Silverman. Ces valeurs peuvent être utilisées en cas d'absence d'information sur les niveaux d'adéquation. Pour une personnalisation pertinente, chaque niveau d'adéquation doit être renseigné par l'enseignant.

6.2.4. Règles et inférences dans le SASA

La spécification des règles au niveau du noyau sémantique permet d'associer les entités appartenant aux différents modules du système d'apprentissage. Le niveau d'adéquation relatif à la dimension de *participation (actif-réflexif)* du style d'apprentissage peut être déduit, jusqu'à un certain degré, à partir des valeurs modélisées par LOM. Le Tableau 6-2 propose une estimation du niveau d'adéquation selon les valeurs enregistrées dans les propriétés 5.1. *Type d'interactivité*, 5.2. *Type de la ressource d'apprentissage* et 5.3. *Niveau d'interactivité de LOM*. Les valeurs positives approchent la ressource d'apprentissage de l'extrémité *réflexif*, alors que les valeurs négatives approchent la ressource de l'extrémité *actif*.

Tableau 6-2. Caractéristiques des objets pédagogiques associées aux dimensions du style d'apprentissage

Variable dans l'OP	Valeur OP	Valeur du niveau d'adéquation	Dimension dans le style d'apprentissage Felder Silverman
LOM 5.1 Type d'interactivité	passive mixed active	+3 0 -3	Participation-
LOM 5.2 Type de la ressource d'apprentissage	exercice simulation questionnaire diagramme figure graph index présentation tableau texte narratif examen exprimment énoncé de problème auto évaluation lecture	-1 -1 -1 +1 +1 +1 +1 +1 +1 +1 -1 -1 +1 -1 +1	Participation
LOM 5.3 Niveau d'interactivité	très faible, faible, moyen, élevé, très élevé	0 -1 -2 -3 -4	Participation

Considérons un objet pédagogique ayant la description présentée dans Extrait de code 6-9. Au départ, la valeur du niveau d'adéquation (`lo:suitabilityValue`) correspondant aux dimensions du style d'apprentissage est *neutre* (i.e. a la valeur 0).

```
res:learningObject1      lo:hasEducationalDescription [
    lo:hasInteractivityType      lo:active ;
    lo:learningRessourceType      lo:exercice ;
    lo:learningRessourceType      lo:questionnaire ;
    lo:learningRessourceType      lo:exam ;
    lo:hasIteractivityLevel      lo:high ;
] .

res:learningObject1
lo:hasSuitabilityForLearningStyleDimensionValue [
    lo:suitabilityDimension up:ParticipationFSLs ;
    lo:suitabilityValue 0 ;
] .
```

Extrait de code 6-9. Représentation des propriétés pédagogiques d'une ressource d'apprentissage

Nous définissons trois groupes de règles correspondant aux trois caractéristiques du LOM illustrées dans le Tableau 6-2. Dans le premier groupe de règles, nous illustrons la règle qui permet de définir le niveau d'adéquation d'une ressource d'apprentissage selon le type d'interactivité. En ce sens, la règle affirme que *si* la valeur de la propriété `lo:hasInteractivityType` est `lo:active`, *alors* la valeur du niveau d'adéquation de cette ressource est modifiée dans la direction du sens *actif* de la dimension *participation* du style d'apprentissage de Felder-Silverman. Cette règle est présentée dans l'Extrait de code 6-10 et est structurée en trois parties : un nom de règle, le corps de la règle et l'entête de la règle. Le corps et l'entête de la règle sont séparés par le symbole `->`. La règle affirme que *si* toutes les affirmations avant le symbole `->` (les conditions) sont satisfaites, *alors* les conclusions énumérées après le symbole `->` vont avoir lieu. Trois fonctions fournies par le moteur d'inférence sont utilisées dans cet exemple. La fonction `noValue(?s ?p ?o)` est vrai si, dans la base de connaissances, n'existe pas un triplet correspondant aux valeurs `(?s ?p ?o)`. La fonction `difference(?x,?y,?d)` permet d'assigner, à la troisième variable, la différence entre les deux premières : `?d = ?x - ?y`. La fonction `remove(n)` permet d'enlever le *n*^{ième} triplet appartenant au corps de la règle. Enfin, la fonction `hide(p)` permet de déclarer que les triplets utilisant le prédicat `p` sont « cachés ». Cette fonction permet de signaler les prédicats utilisés pour le contrôle de déclenchement des règles. Dans ce cas, le prédicat `app:firedRule` permet de signaler quelle règle était appliquée à quelle ressource.

```
[rule1:
(?s rdf:type lo:LearningObject),
(?s lo:hasEducationalDescription ?ed),
(?ed lo:hasInteractivityType lo:active),
(?s lo:hasSuitabilityForLearningStyleDimensionValue ?t),
(?t lo:suitabilityDimension up:ParticipationFSLs),
(?t lo:suitabilityValue ?v), difference(?v,3,?nv),
noValue(?t app:firedRule res:rule1)
->
remove(5),
(?t lo:suitabilityValue ?nv),
(?t app:firedRule res:rule1),
hide(app:firedRule)
]
```

Extrait de code 6-10. Règle de définition du niveau d'adéquation

Pour compléter le niveau d'adéquation, les autres règles sont implémentées de la même manière. Toutefois ces règles permettent d'approximer le niveau d'adéquation d'une ressource d'apprentissage seulement pour la dimension *participation* du style d'apprentissage de Felder-Silverman. Pour les autres dimensions, l'intervention de l'enseignant est nécessaire.

D'autres indices sur l'adéquation des ressources d'apprentissage sont donnés par le niveau d'appréciation des apprenants. Ainsi une appréciation élevée sur une ressource d'apprentissage donnée par un apprenant ayant un style d'apprentissage accentué sur certaines dimensions permet d'élever l'adéquation de cette ressource pour le style d'apprentissage correspondant. Cela pose les bases pour une recommandation des ressources d'apprentissage basée sur un filtrage collaboratif. Les appréciations des apprenants constituent aussi une dimension qui permet de mesurer l'activité de l'apprenant dans le cadre du système d'apprentissage. Les différentes interactions de l'apprenant sont interprétées par le noyau sémantique et intégrées dans la base de connaissances correspondant au profil d'apprenant.

6.3. Agrégation des connaissances dans le SASA

Selon la nature des données, deux types d'informations sont enregistrées dans le système d'apprentissage. Dans un premier temps, les informations concernant l'apprenant sont intégrées à travers *l'interpréteur* qui est implémenté au niveau du noyau sémantique. Dans un second temps, l'agrégation d'information sur les ressources d'apprentissage est réalisée à travers des requêtes SPARQL ou, dans le cas de services spécifiques, en utilisant des interfaces de programmation de l'application.

6.3.1. Acquisition d'informations sur les apprenants

L'interpréteur du noyau sémantique « traduit » le *comportement* de l'apprenant et met à jour son profil. Les aspects de behaviourisme interviennent à cette étape pour identifier des indices correspondant aux dimensions du style d'apprentissage. Ainsi, pour les quatre préférences du style d'apprentissage Felder-Silverman nous reprenons les indicateurs présentés en (Popescu, 2008) :

Préférences sur la *participation* :

- préférence active : accès à une grande quantité de ressources d'apprentissage de type exercice ou de ressources qui demandent beaucoup d'interaction de la part de l'apprenant ;
- préférence passive : accès au contenu théorique avant d'accéder au contenu pratique.

Préférences sur la *perception du contenu* :

- préférence sensorielle : accès au concret, aux exemples en premier ;
- préférence intuitive : accès au contenu théorique ou abstrait, performance élevée dans des exercices applicatifs de la théorie.

Préférences sur le *format du contenu* :

- préférence visuelle : beaucoup de temps passé avec des ressources ayant un contenu graphique riche (e.g. images, schémas graphiques, vidéos) ;
- préférence auditive : beaucoup de temps passé avec des ressources ayant un contenu textuel ou auditif riche (e.g. ressources textuelles, blogs, forums, livres audio).

Préférences sur *l'ordre de présentation* :

- préférence séquentielle : navigation linéaire, accès rare aux ressources externes ;
- préférence globale : modèle de navigation non-linéaire, accès fréquent aux ressources additionnelles.

Les questionnaires d'évaluation du style d'apprentissage représentent un des moyens principaux pour l'évaluation du profil d'apprenant.

6.3.2. Agrégation des ressources d'apprentissage externes

Selon (DUMEZ, 2010) deux approches différentes existent pour la composition des ontologies et services basés sur le Web sémantique : *l'orchestration* et *la chorégraphie* des ontologies. L'orchestration des ontologies et des services sémantiques résulte en un nouveau service, dit composé, qui peut être défini comme l'agrégation de plusieurs autres services (atomiques ou composés). La chorégraphie des ontologies consiste à concevoir une coordination décentralisée des applications et des ontologies. Le noyau sémantique du SASA réalise l'orchestration des différentes ontologies appartenant aux modules intégrés dans le système.

En dehors des sources internes, nous avons pris en considération trois sources externes afin d'enrichir le contenu d'apprentissage : Freebase²⁹⁴, DBpedia²⁹⁵ et YouTube²⁹⁶. L'intégration des articles de Wikipédia, la plus grande encyclopédie en ligne, est réalisée à travers le DBpedia. Deux modalités existent pour effectuer des requêtes dans la base de données de DBpedia. Une première méthode consiste en l'interrogation du point d'accès SPARQL de DBpedia : <http://dbpedia.org/sparql>. Avec des requêtes en format SPARQL, cette méthode est appliquée si le noyau sémantique possède l'identifiant IRI de la ressource d'intérêt. Sinon, l'apprenant doit faire un choix en interrogeant le service de recherche par mots-clés de DBpedia. Ce service est accessible via une interface de programmation à l'adresse : <http://lookup.dbpedia.org/api/search.asmx/KeywordSearch>. La fonction qui exécute des interrogations sur ce service est illustrée dans ANNEXE P. Le service de DBpedia accepte trois paramètres :

- QueryString : la chaîne de caractères pour laquelle l'IRI DBpedia doit être retrouvé ;
- QueryClass : la classe DBpedia conformément à l'ontologie utilisée en DBpedia²⁹⁷ ;
- MaxHits : le nombre maximal de résultats retournés.

Les résultats de l'interrogation sont retournés sous forme d'une chaîne de caractères en format JSON²⁹⁸ (JavaScript Object Notation).

L'exploitation de la base de connaissances Freebase est réalisée de la même manière en utilisant l'adresse <https://www.googleapis.com/freebase/v1/search>. Parmi les paramètres²⁹⁹ qui peuvent être utilisés pour cette interface nous utilisons le paramètre `query` qui représente le terme à rechercher dans la base de connaissances et le paramètre `filter`, qui permet la spécification des contraintes au niveau de la requête.

²⁹⁴ <http://www.freebase.com/>

²⁹⁵ <http://dbpedia.org/About>

²⁹⁶ <http://www.youtube.com/>

²⁹⁷ <http://wiki.dbpedia.org/Ontology?v=d38>

²⁹⁸ <http://www.json.org/>

²⁹⁹ La liste complète de paramètres se retrouve à l'adresse : <https://developers.google.com/freebase/v1/search>.

Pour accéder aux métadonnées des vidéos de YouTube, Google met à disposition Youtube Data API³⁰⁰ et une librairie Java³⁰¹ pour faciliter l'implémentation de l'interface. Les différentes métadonnées caractérisant une vidéo sur YouTube sont spécifiées sur <https://developers.google.com/youtube/v3/docs/videos>. Nous présentons dans la Figure 6-8 la représentation de ces métadonnées sous forme de graphe RDF.

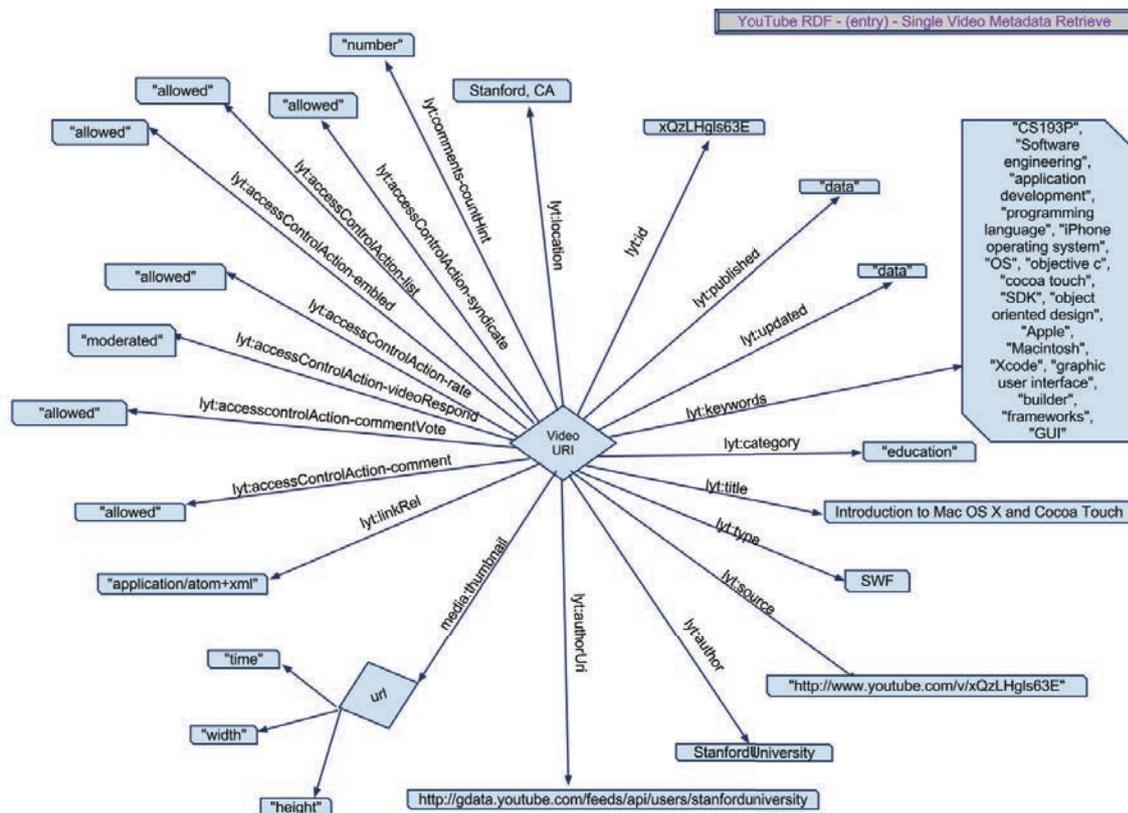


Figure 6-8. Métadonnées d'une ressource de type vidéo sur YouTube

La transformation de la représentation des métadonnées en format RDF est réalisée en utilisant GRDDL (Gleaning Resource Descriptions from Dialects of Languages). Dans l'ANNEXE Q nous présentons quelques séquences du fichier de transformation (XSLT). Pour chaque nœud récupéré, le processeur de transformation applique des modèles de transformation définis dans le document à l'aide des éléments `xsl:template`³⁰². La transformation est appliquée pour chaque instance (`atom:entry`) existant dans les résultats obtenus de l'API de YouTube. Le processeur assigne un IRI local sous la forme http://semlearn.pu-pm.univ-fcomte.fr/resource/youtube_id/id, où

³⁰⁰<https://developers.google.com/youtube/v3/>

³⁰¹<https://developers.google.com/resources/api-libraries/download/youtube/v3/java>

³⁰²Le préfixe `xsl` correspond à <http://www.w3.org/1999/XSL/Transform>.

id est l'identificateur utilisé sur Youtube. La transformation des métadonnées en format RDF permet au noyau sémantique d'exécuter des requêtes SPARQL sur l'ensemble des graphes enregistrés.

6.4. Développement du SASA

Nous présentons, dans ce chapitre, les éléments clés dans le développement du SASA, ainsi que son état de développement actuel. Le développement du système est structuré en trois parties distinctes correspondant au modèle de programmation *Model-View-Controller*. La première partie traite de la réalisation du système conceptuel (*model*), notamment les ontologies du système. Dans la deuxième partie nous présentons l'architecture de classes développées ainsi que les bibliothèques utilisées (*controller*). Dans la troisième partie nous présentons des éléments concernant la construction de l'interface d'utilisateur (*view*).

6.4.1. Développement du modèle et de la base de connaissance de SASA

Plusieurs outils existent pour la création et le développement des ontologies. Dans (Kapoor et Sharma, 2010) est présentée une comparaison entre quatre outils de construction d'ontologies : *Apollo*, *Protégé*, *IsaViz* et *Swoop*. Dans (Alatrish, 2012), la comparaison inclut cinq outils : *Apollo*, *OntoStudio*, *Protégé*, *Swoop* et *TopBraid Composer*. D'autres publications offrent des comparaisons entre plusieurs outils disponibles³⁰³. Alani *et al.* qualifient *Protégé* d'« application qui tue » (en anglais « killer app ») (Alani, O'Hara, et Shadbolt, 2005). Nous avons donc choisi le logiciel *Protégé* pour la construction de modèles ontologiques. *Protégé*³⁰⁴ est développé à l'université de Stanford en collaboration avec l'université de Manchester. Il est librement accessible et ne nécessite pas l'acquisition d'une licence.

Pour simplifier l'écriture des ontologies, *Protégé* utilise la syntaxe Manchester³⁰⁵. Ainsi les opérateurs logiques sont écrits en utilisant les mots clés : `and`, `or` et `not`. Le mot clé

³⁰³ http://www.xml.com/2004/07/14/examples/Ontology_Editor_Survey_2004_Table_-_Michael_Denny.pdf

³⁰⁴ <http://protege.stanford.edu/>

³⁰⁵ <http://www.w3.org/TR/owl2-manchester-syntax/>

pour exprimer la condition d'existentialité (`owl:someValuesFrom`) est `some`, et pour la condition d'universalité (`owl:allValuesFrom`), il s'agit du mot clé `only`.

Avant d'intégrer les ontologies dans le système d'apprentissage, nous avons importé dans le logiciel Protégé l'ensemble des ontologies (v. Figure 6-9). Cela nous permet de lancer le moteur d'inférence sur l'ensemble des ontologies développées.

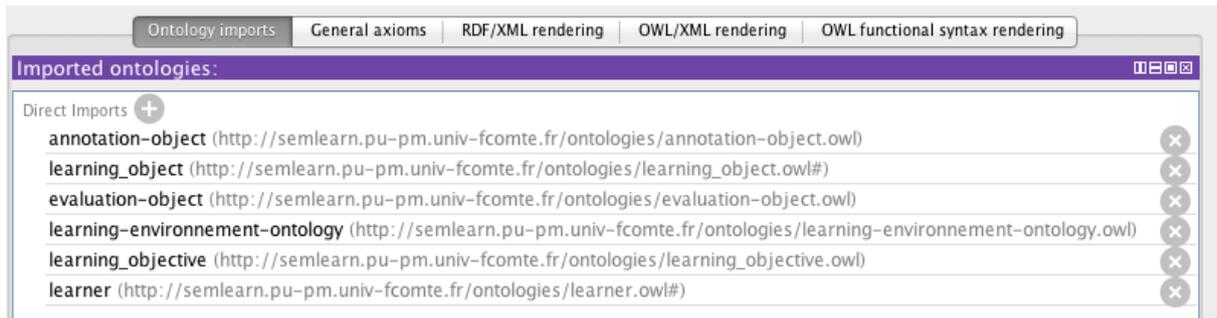


Figure 6-9. L'importation des ontologies en Protégé

Pour faciliter l'administration et l'agrégation des ontologies au niveau de l'ontologie d'application, nous avons déclaré des classes correspondant à chaque ontologie importée (v. Figure 6-10).

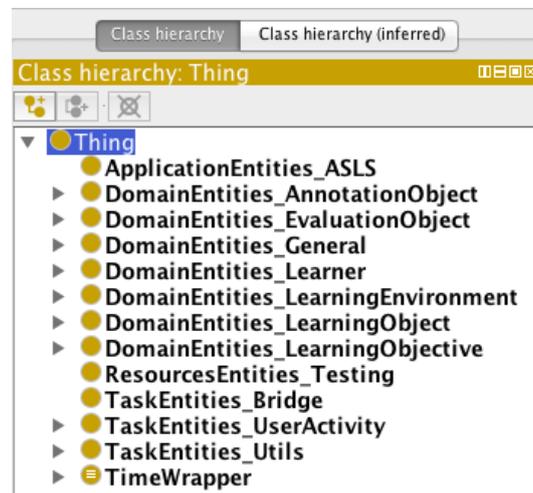


Figure 6-10. L'intégration des ontologies réalisées au niveau de l'ontologie d'application

La manipulation des ontologies dans le SASA est réalisée en utilisant la librairie Jena³⁰⁶. Écrite en langage Java, Jena est une librairie qui permet la construction des applications du Web sémantique.

³⁰⁶ <http://jena.apache.org/>

6.4.2. Développement de la logique du SASA : les contrôleurs et les interfaces

L'implémentation du système d'apprentissage est concrétisée dans une application Web. La librairie Jena permet la manipulation des modélisations et des ontologies déclarées. Pour faciliter le développement de l'interface d'utilisateur du SASA nous avons utilisé la librairie GWT³⁰⁷ (Google Web Toolkit).

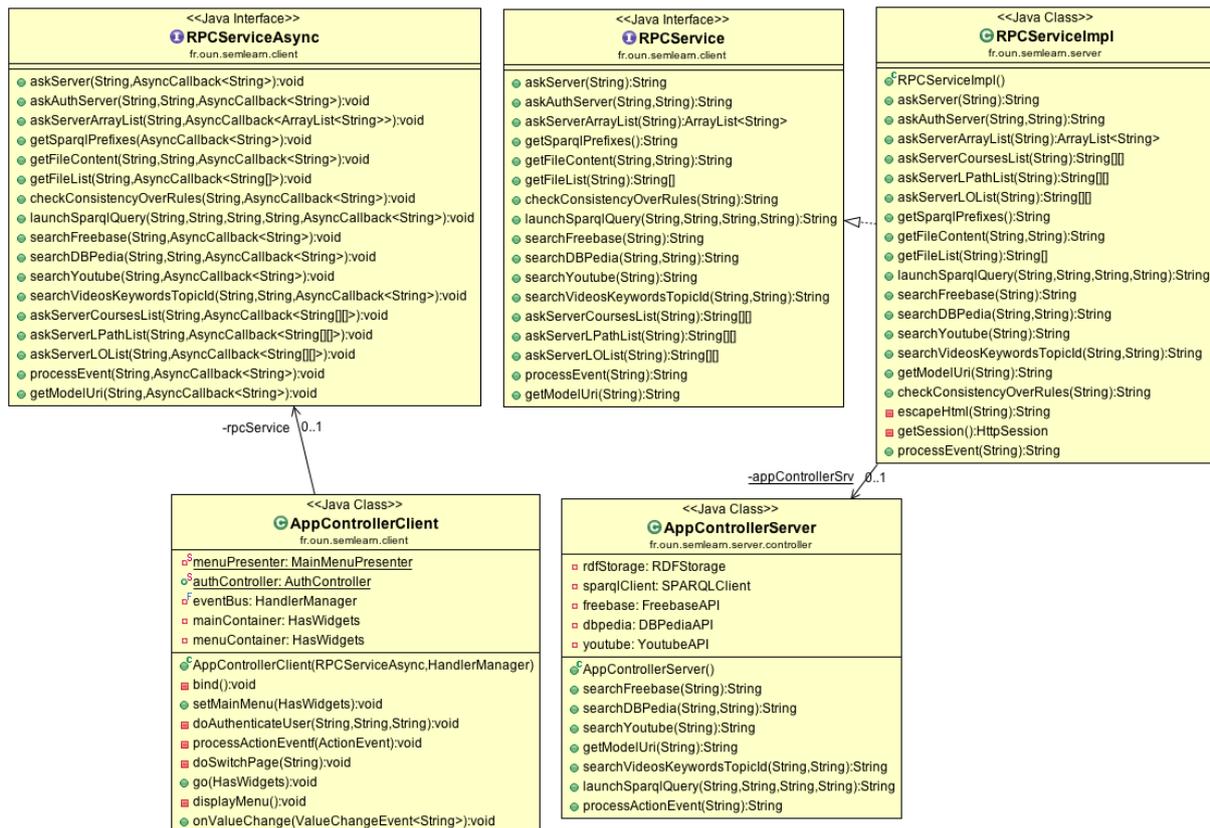


Figure 6-11. Les contrôleurs dans le SASA

Suivant le paradigme de programmation MVC, nous avons structuré le code correspondant en deux parties : la partie serveur et la partie client. L'échange d'informations entre ces deux parties est réalisé en utilisant l'interface `RPCService` (v. Figure 6-11). Chaque partie de l'application dispose d'un contrôleur qui gère la logique de la partie correspondante. La partie du code qui interagit avec les modèles déclarés (les ontologies), ainsi qu'avec les sources de données extérieures se trouve sur le serveur. La Figure 6-12 présente les principales classes déclarées sur la partie serveur.

³⁰⁷ L'utilisation de GWT nous permet d'écrire le code en langage Java. Le compilateur de GWT transforme le code Java en code JavaScript optimisé pour les différents navigateurs Web. <http://www.gwtproject.org/>.

Ainsi les classes `SPARQLClient` et `RDFStorage` permettent au système de manipuler les données en format RDF. Les classes `DBpediaAPI`, `FreebaseAPI` et `YoutubeAPI` permettent au système d'intégrer des ressources extérieures au système d'apprentissage.

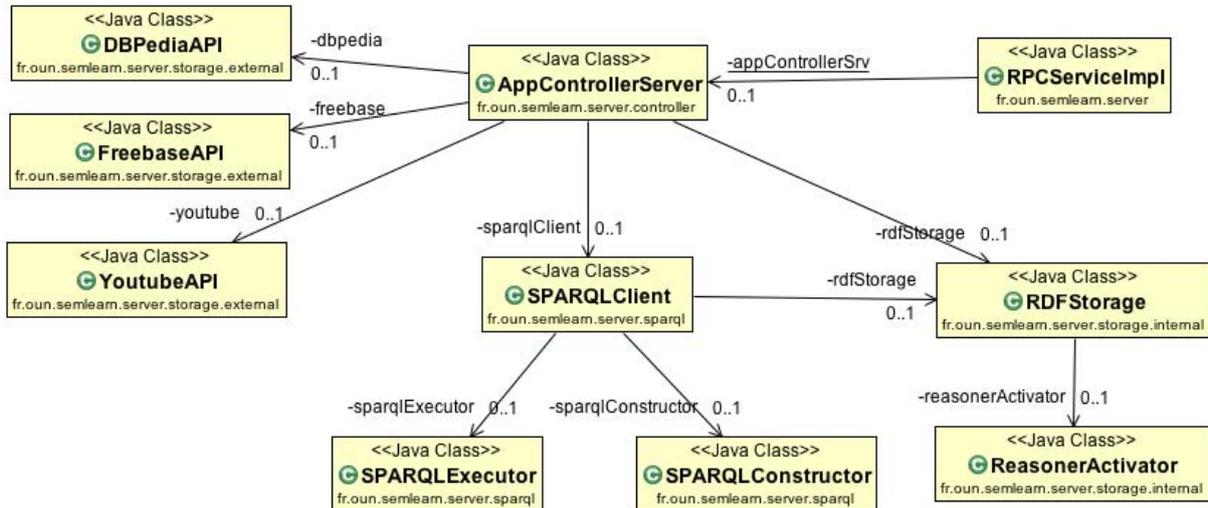


Figure 6-12. L'architecture des classes sur le serveur

L'utilisation de la librairie GWT nous a facilité le processus d'écriture du code pour la partie client. Chaque page affichée dispose de son propre contrôleur (v. Figure 6-13).

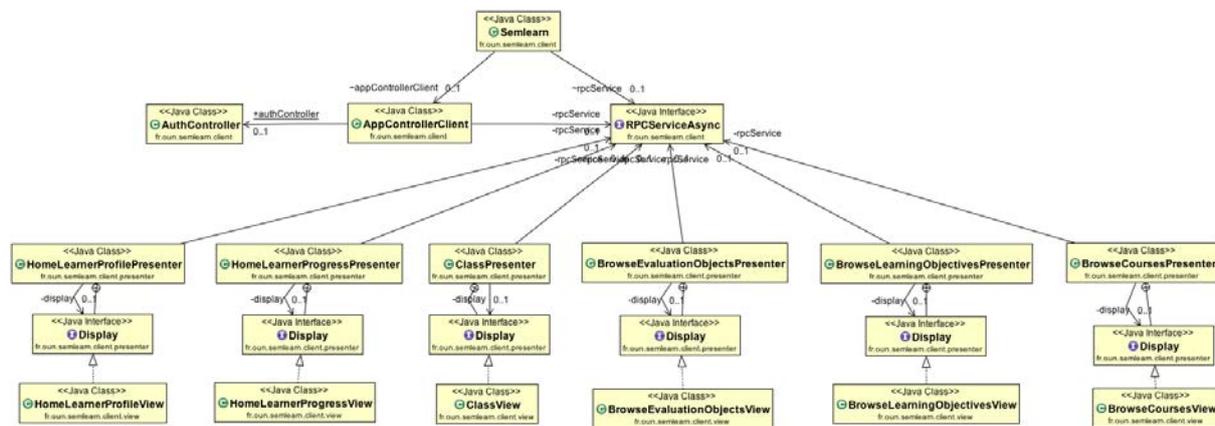


Figure 6-13. L'architecture des classes sur le client

Sur la partie client, la gestion de la logique est structurée sur deux niveaux : le contrôleur de chaque page (e.g. `HomeLearnerProfilePresenter`) et le contrôleur de la partie client (i.e. `AppControllerClient`).

Chapitre 7. Évaluation du SASA

La complexité des situations d'apprentissage et la diversité des facteurs impliqués dans le processus d'apprentissage imposent la spécification de différentes méthodes et approches pour l'évaluation des systèmes d'apprentissage. Dans l'ensemble, l'évaluation d'un système d'apprentissage peut être structurée en plusieurs niveaux (e.g. qualité du contenu d'apprentissage, ergonomie de l'interface du système, modèle de présentation du contenu d'apprentissage). L'objectif de nos travaux étant de vérifier, proposer et spécifier un cadre d'utilisation des technologies sémantiques dans les systèmes d'apprentissage, nous évaluons notre contribution en perspective du SASA et nous apprécions la valeur ajoutée par les technologies du Web dans notre système.

Dans le cadre des systèmes d'apprentissage adaptatifs Brusilovsky *et al.* (Brusilovsky, Karagiannidis, et Sampson, 2004) structurent l'évaluation de ces systèmes sur deux niveaux : (1) l'évaluation de la modélisation de l'utilisateur et (2) l'évaluation de la prise de décision en vue d'adaptation. Ainsi nous réalisons une évaluation du SASA selon ce modèle en niveaux. Dans le premier niveau d'évaluation nous incluons tous les modèles déclarés dans le système d'apprentissage (e.g. apprenant, objet pédagogique, objectif d'apprentissage). Le deuxième niveau d'évaluation concerne le noyau sémantique, notamment l'utilisation des technologies sémantiques à ce niveau et la capacité d'influer le parcours d'apprentissage selon les modèles et les règles définies.

7.1. Évaluation du modèle du SASA

L'évaluation du modèle utilisé en SASA est une évaluation de la validité des informations enregistrées dans le modèle. L'objectif est de vérifier la conformité du modèle du monde à modéliser avec le monde modélisé formellement. Pour le SASA, l'évaluation du modèle est réalisée en conformité avec le métamodèle correspondant. Utilisant les ontologies pour la construction des modèles en SASA, l'évaluation revient à une analyse et validation des ontologies. La question la plus simple donne aussi le critère de base de cette évaluation : est-ce que l'ontologie remplit son but ? Trois groupes de critères sont définis dans (Hitzler *et al.*, 2009, p. 317) : (1) critères sur la logique ; (2) critères sur la

structure et la forme ; et (3) critères sur la fidélité (l'exactitude avec laquelle l'ontologie capte le monde à modéliser). D'après (Gómez-Pérez, 1999) l'évaluation des ontologies est concernée par la construction correcte du contenu de l'ontologie, au fait que les définitions (écrites en langage naturel ou formel) implémentent correctement les exigences de l'ontologie et que l'ontologie s'applique correctement au modèle du monde modélisé. Selon le but et la complexité de l'ontologie, son évaluation peut être catégorisée de la manière suivante (Brank, Grobelnik, et Mladenić, 2005) :

- comparaison de l'ontologie avec un standard de référence, qui peut être aussi une ontologie ;
- basé sur l'utilisation de l'ontologie dans une application et l'évaluation des résultats ;
- comparaison avec un ensemble des données (e.g. collection de documents) dont l'ontologie est spécifiée ;
- l'évaluation est réalisée par des humains, qui essayent de déterminer dans quelle mesure l'ontologie répond à un ensemble prédéfini de critères.

Nous retenons l'idée que l'évaluation ultime devrait être réalisée par des humains tout en utilisant des méthodes et des outils adaptés aux différents critères et tests.

Rappelons que l'objectif de la construction du modèle dans le cadre du SASA est de fournir une représentation qui permet au noyau sémantique d'intervenir sur le parcours d'apprentissage d'un apprenant selon le modèle d'adaptation spécifié dans le noyau sémantique. Le modèle que nous proposons contient d'un part la modélisation de l'apprenant et d'autre part la modélisation des entités participants à l'apprentissage. La corrélation entre les caractéristiques présentes dans ces modèles permet l'adaptation du parcours d'apprentissage. Ainsi la dimension de l'adéquation des ressources d'apprentissage aux dimensions spécifiques du style d'apprentissage de l'apprenant répond à cet objectif.

Le pire scénario pour le moteur d'inférence du système d'apprentissage est d'avoir un modèle dont la classe de complexité correspond à OWL 2 DL. À ce niveau, l'expressivité du modèle est maximale : à savoir *SROIQ(D)*. Comme présenté dans le § 4.1.3 la classe de complexité correspondante à ce niveau d'expressivité est *2NEXPTIME-Complete*. Après

avoir inséré l'ensemble des ontologies en Protégé, les métriques nous indiquent l'expressivité SROIQ(D) (v. Figure 7-1).

Metrics	
Axiom	753
Logical axiom count	442
Class count	106
Object property count	42
Data property count	16
Individual count	111
DL expressivity	SROIQ(D)

Figure 7-1. Mesures sur les ontologies du SASA

Nous énumérons les principales circonstances modélisées qui conduisent à cette classe d'expressivité :

- $\mathcal{S} = \mathcal{ALLC}$ + transitivité des rôles : l'enchaînement des ressources d'apprentissage ou des objectifs d'apprentissage (e.g. `:hasNext`, `:isPrerequisiteFor`) ; cela nous permet de retrouver tous les objets pédagogiques suivants ou toutes les compétences prérequisées à partir d'une instance donnée ;
- \mathcal{R} – inclusion des rôles
 - o $hasValidated \circ hasLearningObjective \sqsubseteq hasAcquired$;
- \mathcal{H} – hiérarchies des rôles (inclut par \mathcal{R})
 - o $hasLearningStyle \sqsubseteq learnigStyleObjectProperties$;
- \mathcal{O} – nominaux, énumération d'individus ;
- \mathcal{I} – rôles inverses (e.g. $hasLSDimension \equiv isLSDimensionFor^-$) ;
- \mathcal{N} – restrictions de cardinalité (e.g. $\geq 2hasCreated$) ;
- \mathcal{Q} – restrictions qualifiées de cardinalité (e.g. $\geq 2hasValidated.Competence$) ;
- \mathcal{F} – rôles fonctionnels (e.g. `:isAuthorOf`) ;
- \mathcal{D} – types de données (e.g. `dateTime`, `Literal`).

La caractéristique *fonctionnelle* d'une propriété est une caractéristique intrinsèque, alors que la restriction de cardinalité sur la propriété permet plus de flexibilité. La décision d'utiliser la restriction qualifiée de cardinalité est nécessaire dans les situations quand nous voulons spécifier le type d'entité sur lequel la cardinalité s'applique. L'utilisation des types de données est nécessaire pour l'interprétation des types de données (e.g. temps, littéraux) par le système. Au niveau de base, la fidélité ou la

pertinence du modèle du système SASA revient au compromis entre une complexité acceptable et un temps de calcul raisonnable.

Dans la phase de conception, pour la validation de la cohérence de l'ontologie nous avons utilisé trois moteurs d'inférences accessibles en Protégé : FAcT++, HermiT et Pellet. Pendant la construction des ontologies les moteurs d'inférences cités n'ont pas signalé de problèmes d'incohérences au niveau des ontologies. La librairie Jena fournit son propre moteur d'inférence tout en permettant l'intégration des moteurs externes comme Pellet. Dans nos simulations, sur une base de connaissances contenant moins de 10000 triplets, l'exécution du moteur d'inférences n'impose pas des temps d'attente importants. Dans une perspective de test et de simulation du modèle du SASA nous n'avons pas examiné les possibilités d'optimisations au niveau technique. Dans une approche de production, le temps d'exécution acceptable peut influencer le niveau d'expressivité du modèle.

L'évaluation est une étape nécessaire dans le processus répétitif d'affinage d'une ontologie. Cette étape est indispensable dans le processus de développement de l'ontologie, mais intervient aussi dans le cycle de vie de l'application. Toutefois, la validation finale est donnée par l'adoption et l'utilisation de l'ontologie par la communauté.

7.2. Tests et simulations du noyau sémantique

Pour vérifier la fonctionnalité du SASA nous avons effectué deux tests. Le premier test concerne la capacité du système à modifier le parcours de l'apprenant conformément à son profil. Le deuxième test concerne la fonction de recherche de données dans les bases de connaissances extérieures au SASA.

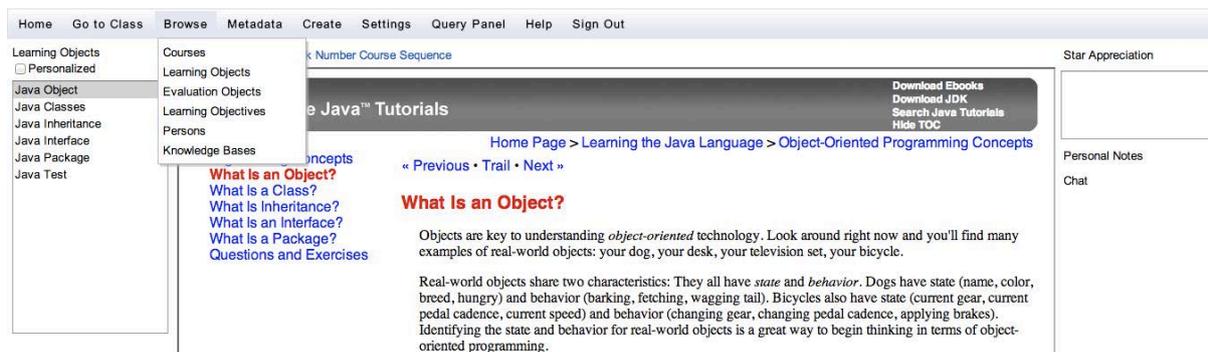


Figure 7-2. Parcours d'apprentissage sans adaptation

La Figure 7-2 illustre une capture d'écran d'un cours en Java, tel que modélisé et implémenté dans notre système. Le parcours d'apprentissage est représenté sous forme d'une liste d'objets pédagogiques. Sans prendre en compte les caractéristiques de l'apprenant, ce parcours d'apprentissage est le même pour chaque apprenant choisissant ce cours. Le menu *Browse* (*parcourir en anglais*) permet à l'apprenant de feuilleter les différentes entités représentées dans le SASA (e.g. cours, objets pédagogiques, objets d'évaluation). L'accès au menu *Go to Class*, permet à l'apprenant de se rendre dans la salle virtuelle de classe.

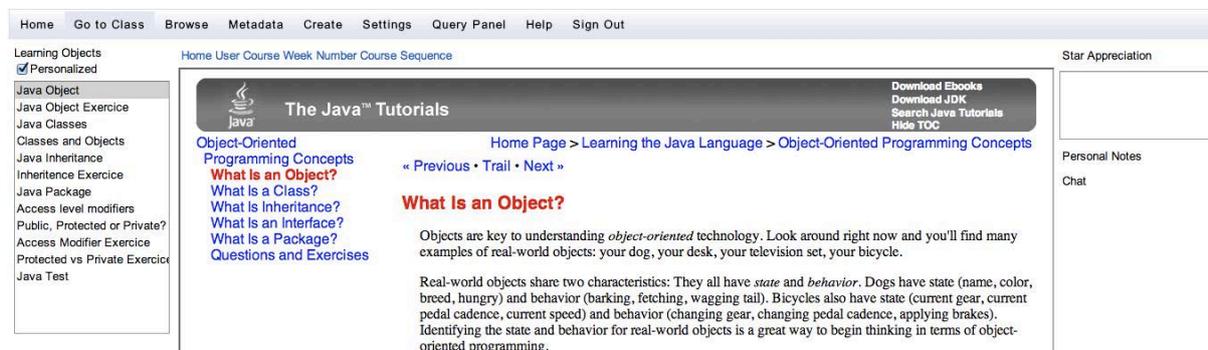


Figure 7-3. Parcours d'apprentissage adapté

La case à cocher *Personalized* (v. Figure 7-3) permet d'insérer les propositions du SASA dans le parcours d'apprentissage. Nous avons choisi d'intégrer un objet d'évaluation (e.g. exercice) en relation avec l'objectif d'apprentissage de chaque objet pédagogique. Cela rend l'apprenant actif à chaque étape du parcours d'apprentissage et fournit des indicateurs sur son avancement dans le processus d'apprentissage. Ainsi, une fois la case *Personalized* sélectionnée, le parcours d'apprentissage est enrichi avec des objets pédagogiques conformes aux règles définies au niveau du noyau sémantique. Un changement dans les dimensions du style d'apprentissage de l'apprenant produit une

modification du parcours d'apprentissage proposé. Étant dans la phase de conception les fonctionnalités au niveau de l'interface sont limitées. En perspective, l'idée est de permettre à l'apprenant de paramétrer les informations et le niveau d'adaptation offert par la plateforme.

Dans le contexte des données liées, DBPedia et Freebase sont deux projets similaires qui extraient des données structurées de Wikipedia afin de les rendre disponibles en format RDF. Ainsi l'exploitation de l'encyclopédie Wikipedia peut être réalisée à partir de ces deux ensembles des données. Nous avons intégré dans les fonctionnalités du SASA des possibilités de recherche des entités dans ces ensembles de données, comme l'illustre la Figure 7-4. Ainsi cette fonctionnalité permet une recherche simultanée sur les trois ensembles des données.

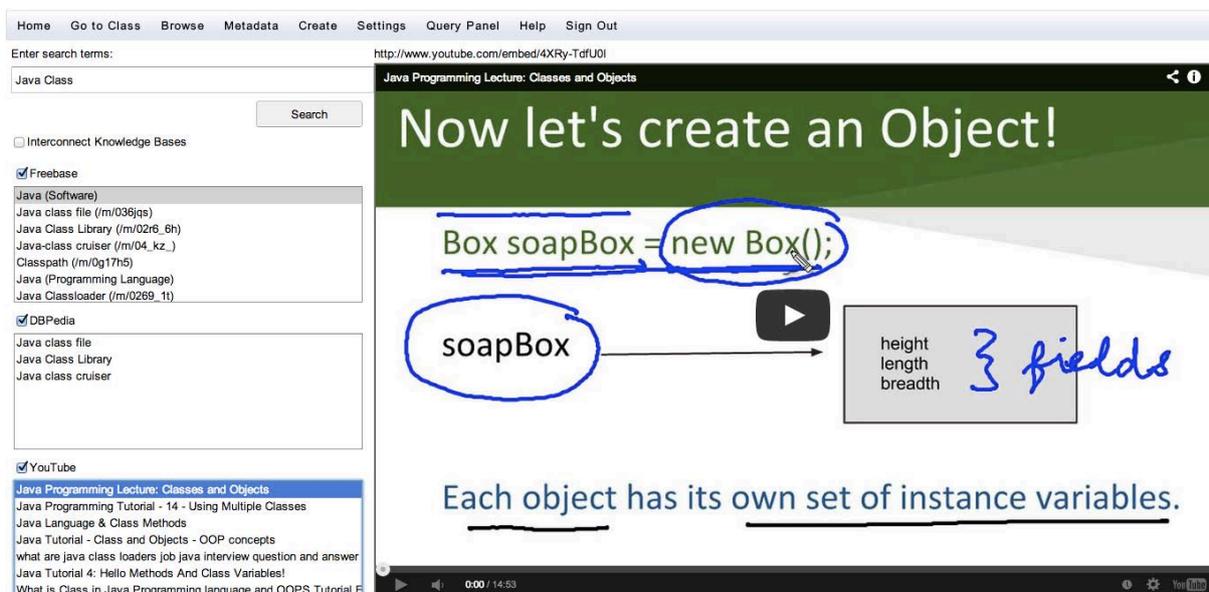


Figure 7-4. Identification des ressources externes au SASA

La librairie utilisée pour la recherche des vidéos sur Youtube peut être configurée pour accepter comme variable d'entrée un *sujet* (*topic* en anglais) fourni par Freebase. La case à cocher *Interconnect Knowledge Bases*, permet une recherche dans l'ensemble de données en utilisant des sujets ou des identifiants IRI. Ces liaisons au niveau d'identifiants permettent d'éviter l'erreur qui peut être introduite par les algorithmes de traitement du langage naturel. Ainsi nous pouvons identifier les ressources correspondantes appartenant à des ensembles de données différents (e.g. Wikipédia, Youtube). Les simulations effectuées ont donné les réponses attendues, dans le sens que

les ressources proposées par le système suivant une recherche croisée sur ces trois ensembles de données ont permis d'enrichir et de suggérer des résultats pertinents. Nous considérons l'expansion du système vers d'autres ensembles des connaissances. L'intégration des termes spécifiques du vocabulaire Schema.org définis pour la description des contenus pédagogique va permettre d'élargir la quantité des ressources d'apprentissage identifiables avec les technologies sémantiques.

Conclusions et perspectives

*« Toute connaissance acquise sur la connaissance devient un moyen de connaissance éclairant la connaissance qui a permis de l'acquérir »
(Morin, 1992, p. 232).*

Le travail de recherche présenté dans ce mémoire concerne les modalités d'application et d'intégration des technologies du Web sémantique aux systèmes d'apprentissage. Notre hypothèse de recherche est que, dans le contexte actuel, fortement influencé par les systèmes d'information et de communication, et particulièrement par le Web, les systèmes d'apprentissage peuvent bénéficier des technologies sémantiques. Nous présentons le SASA comme un prototype qui s'appuie sur des technologies sémantiques pour permettre à la machine d'améliorer son apport dans le processus d'apprentissage.

Après avoir présenté les principales théories et courants d'apprentissage, nous avons identifié que les changements actuels, la dynamique de la société et la vaste quantité d'informations disponible à l'heure actuelle, ont fait du connectivisme un des principaux concepts d'apprentissage. Le savoir-où devient tout aussi important que le savoir-quoi ou le savoir-faire.

Pour la construction de notre prototype de système d'apprentissage, nous avons fait appel à l'approche systémique. Nous avons aussi identifié les technologies du Web sémantique comme pouvant répondre aux défis spécifiés pour ce système, à savoir la recommandation active et proactive de contenus d'apprentissage selon le profil et les actions entreprises par l'apprenant sur la plateforme.

Reposant sur la structure du Web et utilisant le formalisme utilisé dans les langages de logique descriptive (DL), nous avons choisi le langage OWL pour la modélisation des entités composant notre système. Ainsi nous proposons des modélisations pour l'apprenant, l'objet pédagogique, l'objectif d'apprentissage, l'objet d'évaluation, l'objet d'annotation et le cadre d'enseignement. Nous définissons aussi des ontologies

spécifiquement construites pour ces modélisations. Ces ontologies sont vitales pour que le noyau sémantique de notre système puisse répondre aux objectifs que nous nous sommes donnés.

La modification automatique du parcours d'apprentissage conforme au profil ou aux actions de l'apprenant fait de notre prototype SASA un système *actif*. Ainsi, les particularités de l'apprenant, l'identification de points faibles dans son parcours ou ses préférences d'apprentissage permettent au noyau sémantique d'adapter le parcours d'apprentissage en concordance.

Nous considérons que les technologies et les outils du Web sémantique peuvent améliorer les systèmes d'apprentissage sur deux aspects principaux. Dans un premier temps les technologies sémantiques permettent au SASA d'améliorer l'identification et l'organisation des ressources d'apprentissage. L'explicitation des descripteurs sémantiques et des métadonnées dans un langage traitable par la machine rendent plus pertinentes les capacités d'indexation et de recherche du contenu d'apprentissage. L'accord sémantique sur les descripteurs communs et partagés garanti la cohérence et la pertinence des recherches effectuées par le SASA. L'utilisation de plusieurs caractéristiques et descripteurs sémantiques dans la recherche d'une ressource d'apprentissage augmente la précision avec laquelle le système peut identifier cette ressource.

Dans un second temps l'utilisation des langages formels et des logiques de description permet de rendre plus intelligents les systèmes d'apprentissage. Ainsi les inférences automatiques ou les règles appliquées sur les descripteurs des entités permettent une organisation dynamique du système conceptuel. Rendre plus intelligents les systèmes d'apprentissage permettra aux enseignants de décharger les tâches administratives sur la machine et de se concentrer sur les aspects pédagogiques de l'apprentissage. Le but d'un tel système est à la fois de soutenir l'apprenant dans son processus d'apprentissage et d'assister l'enseignant dans son travail d'accompagnement.

La partie la plus laborieuse dans l'élaboration du SASA est la construction des ontologies. Le nombre des caractéristiques, les contraintes logiques peuvent vite

compliquer l'ontologie et rendre difficile la construction d'un modèle pertinent et représentatif pour le monde à modéliser. En perspective, nous considérons la construction automatique des ontologies et la validation de celles-ci par les spécialistes du domaine. Ainsi l'apprentissage de l'ontologie (*ontology learning*), un sous-domaine de l'apprentissage-machine (*machine learning*), facilitera la construction des ontologies à partir des ensembles de données. Une autre perspective d'amélioration du système d'apprentissage consiste en la construction et l'intégration au niveau du noyau sémantiques des ontologies spécifiques à la discipline enseignée. Ainsi l'ontologie des objectifs d'apprentissage pourra être développée et affinée en relation avec l'ontologie de la discipline enseignée.

Principalement, la qualité du parcours d'apprentissage proposée par le SASA est influencée par deux facteurs : le premier est donné par la justesse et le nombre des descripteurs sémantiques et le deuxième par la pertinence de l'algorithme d'adaptation. Sans insister sur l'efficacité de l'approche présentée, nous considérons que les algorithmes d'adaptation constituent un sujet de recherche en lui même. Le nombre de caractéristiques du contenu prises en compte, le poids de chaque caractéristique, les caractéristiques de l'apprenant, l'évaluation de l'adaptation sur d'autres apprenants, l'indications des enseignants constituent des facteurs qui peuvent vite compliquer les algorithmes de personnalisation de l'apprentissage.

Nous retenons un des buts principaux du Web sémantique : améliorer la coopération entre les humaines et les machines. D'autant plus dans le domaine d'apprentissage, les systèmes devraient être un outil encore plus efficace et intelligent. Nous considérons que le Web sémantique peut améliorer ces systèmes d'apprentissage.

Toujours dans l'axe de recherche *plateformes sémantiques d'apprentissage* de l'équipe OUN, nous travaillons sur la construction d'un module d'évaluation dans le cadre du projet SCOLA³⁰⁸ (*Système de Communication Ouvert et Ludique pour les Apprentissages*).

³⁰⁸ Lauréat de l'appel à projet n°2 de la direction générale de l'enseignement scolaire du Ministère de l'Éducation Nationale. Il s'agit d'un projet ancré dans les Technologies de l'Information et de la Communication pour l'Enseignement (TICE), qui a démarré le 01.03.2013 et s'achèvera dans 3 ans. Le financement pour l'équipe OUN est de 107k€. Le consortium du projet SCOLA se compose de partenaires a) universitaires (Laboratoire ADEF, Université de

SCOLA est une plateforme de e-learning innovante fondée sur l'utilisation des *serious games* facilitant l'acquisition et le repérage de compétences. L'équipe OUN participe au projet SCOLA pour réaliser la modélisation sémantique des concepts du module d'évaluation des connaissances et proposer des solutions de normalisation de la plateforme qui intègre les serious games. Pour la construction des modélisations utilisées en SCOLA nous considérons l'ontologie des objectifs d'apprentissage présentée dans ce travail de recherche.

Provence ; équipe OUN-Laboratoire ELLIADD, Université de Franche-Comté ; équipe VORTEX-Laboratoire IRIT, Université Paul Sabatier ; b) industriels (ARIES ; AGIIR Network ; Idées-3Com) ; c) institutionnels (Académie d'Aix-Marseille ; Conseil Régional PACA ; 4 conseils généraux (départements de l'Académie d'Aix-Marseille) ; Mairie de Marseille).

Liste des publications

1. Szilagyi, I., Roxin, A., Roxin, I., « Active Learning System in the Context of the Social Semantic Web » Proceedings of the Sixth International Conference on Developments in E-Systems Engineering ; Abu Dhabi, EAU; 16-18 December 2013
2. Szilagyi, I., Roxin, I., « Learner Ontology for the Active Semantic Learning System » Proceedings of the 12th IEEE International Conference on Advanced Learning Technologies; Rome, ITALY, 2012, pp.393-394.
3. Szilagyi, I., Roxin, I., « Learner and Knowledge Modeling for an Active Semantic Learning System » Proceedings of the 11th International Conference on Informatics in Economy; Bucharest, ROMANIA, 2012, pp.366-372.
4. Szilagyi, I., Balog-Crisan, R., Roxin, A., Roxin, I., « Ontologies and Knowledge Aggregation in the Active Semantic Learning System » Proceedings of the 11th IEEE International Conference on Advanced Learning Technologies, Athens, ATLANTA, USA, 2011, pp.388-392.
5. Greffier, F., Szilagyi, I., Domenget, J.C., « Pour des parcours personnalisés via un système actif et sémantique d'apprentissage » Echanger Pour Apprendre en Ligne, EPAL 2011, Grenoble, France, 2011
http://w3.u-grenoble3.fr/epal/dossier/06_act/pdf/epal2011-greffier-et-al.pdf
6. Szilagyi, I., Greffier, F., Domenget, J.C., « Apprentissage personnalisé via le web sémantique », Atelier EIAH – 24 mai 2011 – Mons, Belgique
http://liris.cnrs.fr/~mlefevre/ActesAtelierEIAH2011/EIAH2011_AtelierPersonnalisationApprentissage_Approche_WebSemantique.pdf
7. Domenget, J.C., Szilagyi, I., Greffier, F., « Web sémantique et système d'apprentissage : apports d'un wiki sémantique », Journée d'étude du groupe TICIS-SFSIC : « Le web a-t-il un sens ? » 14-15 décembre 2010 l'Université Paris 13, Paris, France
8. Szilagyi, I., Roxin, I., « Model for an Active Semantic Learning System » Proceedings of the IADIS International Conference, eLearning 2010, Freiburg, Germany, 2010, pp.247-250. <http://www.iadisportal.org/digital-library/model-for-an-active-semantic-learning-system>
9. Szilagyi, I., Balog-Crisan, R., Roxin, I. « Kernel for a Semantic Learning Platform with Adapted Suggestions » Proceedings of the 10th IEEE International Conference on Advanced Learning Technologies, Sousse, Tunisie, 2010, pp.400-402.
10. Balog-Crisan, R., Roxin, I., Szilagyi, I. « Ontologies for a Semantic Quiz Architecture » Proceedings of the Ninth IEEE International Conference on Advanced Learning Technologies, Riga, 2009, pp. 492-494.
11. Roxin, I., Szilagyi, I., Balog-Crisan, R. « Kernel Design for Semantic Learning Platform » In Proceedings of the World Conference on Educational Multimedia, Hypermedia and Telecommunications, Honolulu, 2009, pp.1190-1194.

Bibliographie

- 10 Jobs That Didn't Exist 10 Years Ago - Forbes. (2012, mai 11). Retrouvé 3 mai 2013, de <http://www.forbes.com/sites/meghancasserly/2012/05/11/10-jobs-that-didnt-exist-10-years-ago/>
- 5 Pédagogique - Sup-LOMFR. (s. d.). Retrouvé 14 août 2013, de http://www.sup.lomfr.fr/index.php?title=5_P%C3%A9dagogique
- 76 Définitions du signe relevées dans les écrits de C.S. Peirce. (s. d.). Retrouvé 11 juillet 2013, de <http://perso.numericable.fr/robert.marty/semiotique/76-fr.htm>
- A bright future for distance learning: One Touch/Hughes alliance promotes interactive « e-learning » service. (1997). Retrouvé 26 mars 2013, de http://connectedplanetonline.com/mag/telecom_bright_future_distance/
- A Standards-based, Open and Privacy-aware Social Web. (s. d.). Retrouvé 5 juillet 2013, de <http://www.w3.org/2005/Incubator/socialweb/XGR-socialweb-20101206/#Overview>
- About — Connectivism. (2005). Retrouvé 12 mars 2013, de <http://www.connectivism.ca/about.html>
- About this project | Restoring the first website. (s. d.). Retrouvé 2 juillet 2013, de <http://first-website.web.cern.ch/>
- About Wolfram|Alpha: Making the World's Knowledge Computable. (s. d.). Retrouvé 15 septembre 2013, de <http://www.wolframalpha.com/about.html>
- Ackoff, R. L. (Aut). (1989). From Data to Wisdom. *Journal of Applied Systems Analysis* 16., 3–9.
- ADL, P. (Aut). (2001). The SCORM Content Aggregation Model. 140.131.114.243. Retrouvé de <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:The+SCORM+Content+Aggregation+Model#0>
- AFNOR (Aut). (2006). *Norme NF Z76-040 Technologies de l'information pour l'éducation, la formation et l'apprentissage : Profil français d'application du LOM (LOMFR)*.
- AFSCET, G. (Aut). (2003, septembre). L'Approche systémique : de quoi s'agit-il ? Retrouvé 30 avril 2013, de <http://www.afscet.asso.fr/SystemicApproach.pdf>
- Aghaei, S. (Aut). (2012). Evolution of the World Wide Web : From Web 1.0 to Web 4.0. *International journal of Web & Semantic Technology*, 3(1), 1-10. doi:10.5121/ijwest.2012.3101
- Ahmad, K., Corbett, G., Rogers, M., et Sussex, R. (Aut). (1985). *Computers, language learning and language teaching*. Cambridge University Press Cambridge. Retrouvé de <http://www.getcited.org/pub/102472421>
- Alani, H., O'Hara, K., et Shadbolt, N. (Aut). (2005). Common Features of Killer Apps: A Comparison with Protégé. Retrouvé de <http://eprints.ecs.soton.ac.uk/10989>
- Alatrish, E. S. (Aut). (2012). Comparison of Ontology Editors. *eRAF Journal on Computing*, 4. Retrouvé de http://joc.raf.edu.rs/4/_k_0020.pdf
- Alberganti, M. (Aut). (2000). À l'école des robots. *L'informatique, l'école et vos enfants*.
- Alle Kennis van de Wereld (Biography of Paul Otlet)*. (1998). Noorderlicht. Retrouvé de <http://archive.org/details/paulotlet>
- Altet, M. (Aut). (1997). *Les pédagogies de l'apprentissage*. Presses Universitaires de France - PUF.

- Álvarez, M. Á., G-Bustelo, B. C. P., Sanjuán-Martínez, O., et Lovelle, J. M. C. (Aut). (2010). Bridging together Semantic Web and Model-Driven Engineering. Dans A. P. de L. F. de Carvalho, S. Rodríguez-González, J. F. D. P. Santana, et J. M. C. Rodríguez (Éd), *Distributed Computing and Artificial Intelligence* (p. 601-604). Springer Berlin Heidelberg. Retrouvé de http://link.springer.com/chapter/10.1007/978-3-642-14883-5_76
- Anderson, J. B., et Johnnesson, R. (Aut). (2006). *Understanding information transmission* (Vol 18). Wiley-IEEE Press. Retrouvé de <http://books.google.fr/books?hl=en&lr=&id=GD5GY4XyPXIC&oi=fnd&pg=PR5&dq=understanding+information+transmission&ots=8mcjCe7-bK&sig=C2nZG8jcuNCXkp6N8HRMfl9Kgiw>
- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., et Raths, J. (Aut). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives* (Complete ed.). New York: Longman.
- Aroyo, L., et Dicheva, D. (Aut). (2004). The New Challenges for E-learning: The Educational Semantic Web. *Educational Technology & Society*, 7(4), 59-69. Retrouvé de <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.107.7038&rep=rep1&type=pdf#page=64>
- Atkinson, C. (Aut). (2006). Models versus Ontologies - What's the Difference and where does it Matter ? Dans *October* (p. 1-16). Présenté au VORTE2006, Hong Kong.
- Baader, F. (Aut). (2003). *The Description Logic Handbook: Theory, Implementation, and Applications*. Cambridge, UK ; New York: Cambridge University Press. Retrouvé de http://books.google.com/books?hl=en&lr=&id=riSeOKw5I6sC&oi=fnd&pg=PR14&dq=Description+Logic+Programs:+Combining+Logic+Programs+with+Description+Logic&ots=T7Yw_-Swef&sig=J0CXZOUf2r324HrtMZgofT4zfhs
- Baader, F., Brandt, S., et Lutz, C. (Aut). (2008). Pushing the EL envelope further. Retrouvé de <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.158.9530>
- Baader, F., et Sattler, U. (Aut). (2001). An overview of tableau algorithms for description logics. *Studia Logica*, 69(1), 5-40. Retrouvé de <http://link.springer.com/article/10.1023/A:1013882326814>
- Bachimont, B. (Aut). (2000). Engagement sémantique et engagement ontologique: conception et réalisation d'ontologies en ingénierie des connaissances. *Ingénierie des connaissances: évolutions récentes et nouveaux défis*, 305-323. Retrouvé de http://www.cndwebzine.hcp.ma/img/pdf/engagement_semantique_et_engagement_ontologique_conception_et_realisation_d_ontologies_en_ingenierie_des_connaissances.pdf
- Bachimont, B. (Aut). (2004a). *Arts et sciences du numérique : Ingénierie des connaissances et critique de la raison computationnelle* (Habilitation à diriger des recherches). Université de Technologie de Compiègne.
- Bachimont, B. (Aut). (2004b). Pourquoi n'y a-t-il pas d'expérience en ingénierie des connaissances? Dans *Ingénierie des Connaissances* (p. 53-64). Retrouvé de <http://hal.archives-ouvertes.fr/hal-00348389/>
- Bacon, F. (Aut). (1627). *New Atlantis*. Clarendon Press.
- Balatsoukas, P., Morris, A., et O'Brien, A. (Aut). (2008). Learning objects update: Review and critical approach to content aggregation. *Educational Technology & Society*,

- 11(2), 119–130. Retrouvé de http://www.ifets.info/others/download_pdf.php?j_id=39&a_id=848
- Balog-Crişan, R. (Aut). (2011). *La modélisation d'objets pédagogiques pour une plateforme sémantique d'apprentissage*.
- Balog-Crisan, R., et Roxin, I. (Aut). (2007). RDF editor for LOM. Dans M. B. Nunes et M. McPherson (Éd), (p. 333-337). Présenté au IADIS International Conference e-Learning, Lisbon, Portugal.
- Balog-Crisan, R., Roxin, I., et Szilagyi, I. (Aut). (2009). Ontologies for a semantic quiz architecture. Dans *Advanced Learning Technologies, 2009. ICALT 2009. Ninth IEEE International Conference on* (p. 492–494).
- Barbier, J.-M. (Aut). (1985). *L'évaluation en formation*. Paris: PUF.
- Baroni, P., Lamperti, G., Pogliano, P., et Zanella, M. (Aut). (1999). Diagnosis of large active systems. *Artificial Intelligence*, 110(1), 135-183. doi:10.1016/S0004-3702(99)00019-3
- Baxter, S., et Vogt, L. C. (Aut). (2002, mars). Content management system.
- Beck, R. J. (Aut). (2008). What are Learning Objects? Retrouvé 18 juin 2013, de http://www4.uwm.edu/cie/learning_objects.cfm?gid=56
- Bellinger, G., Castro, D., et Mills, A. (Aut). (2004). *Data, information, knowledge, and wisdom*. Retrouvé de <http://courseweb.lis.illinois.edu/~katewill/spring2011-502/502%20and%20other%20readings/bellinger%20on%20ackoff%20data%20info%20know%20wisdom.pdf>
- Bénel, A., Calabretto, S., Iacovella, A., et Pinon, J.-M. (Aut). (2002). Porphyry 2001: Semantics for Scholarly Publications Retrieval. Dans M.-S. Hacid, Z. W. Raś, D. A. Zighed, et Y. Kodratoff (Éd), *Foundations of Intelligent Systems* (p. 351-361). Springer Berlin Heidelberg. Retrouvé de http://link.springer.com/chapter/10.1007/3-540-48050-1_39
- Bergson, H. (Aut). (1900). *Le rire: essai sur la signification du comique*. F. Alcan. Retrouvé de <http://nouvellelecturefrancaise.blogs-de-voyage.fr/wp-content/blogs.dir/12554/files/2007/12/Bergson.2.pdf>
- Berners-Lee, T. (Aut). (1989, mars). Information Management: A Proposal. Retrouvé 17 août 2013, de <http://www.w3.org/History/1989/proposal.html>
- Berners-Lee, T. (Aut). (1990, novembre 12). WorldWideWeb: Proposal for a HyperText Project. Retrouvé de <http://www.w3.org/Proposal.html>
- Berners-Lee, T. (Aut). (1991). Naming -- /DesignIssues. Retrouvé 17 septembre 2013, de <http://www.w3.org/DesignIssues/Naming.html>
- Berners-Lee, T. (Aut). (1997a, janvier 6). Web architecture: Metadata. Retrouvé 27 juillet 2013, de <http://www.w3.org/DesignIssues/Metadata.html>
- Berners-Lee, T. (Aut). (1997b, décembre 3). Realising the Full Potential of the Web. Retrouvé 16 septembre 2013, de <http://www.w3.org/1998/02/Potential.html>
- Berners-Lee, T. (Aut). (1998, septembre). Semantic Web Roadmap. Retrouvé 19 août 2013, de <http://www.w3.org/DesignIssues/Semantic.html>
- Berners-Lee, T. (Aut). (2000, décembre 6). *Semantic Web - XML2000*. Présenté au XML2000, Washington, DC. Retrouvé de <http://www.w3.org/2000/Talks/1206-xml2k-tbl/slide10-0.html>
- Berners-Lee, T. (Aut). (2006, juillet 27). Linked Data - Design Issues. Retrouvé 9 juillet 2013, de <http://www.w3.org/DesignIssues/LinkedData.html>
- Berners-Lee, T. (Aut). (2009, février 4). *Linked Data*. Présenté au TED 2009 Conference, Long Beach, CA, USA. Retrouvé de [http://www.w3.org/2009/Talks/0204-ted-tbl/#\(1\)](http://www.w3.org/2009/Talks/0204-ted-tbl/#(1))

- Berners-Lee, T., Fielding, R., et Masinter, L. (Aut). (2005, janvier). Uniform Resource Identifier (URI): Generic Syntax. Retrouvé 26 juillet 2013, de <http://www.ietf.org/rfc/rfc3986.txt>
- Berners-Lee, T., Hendler, J., et Lassila, O. (Aut). (2001). The Semantic Web. *Scientific American*, 284(5), 34–43. doi:10.1038/scientificamerican0501-34
- Bertalanffy, L. V. (Aut). (1968). *General System Theory: Foundations, Development, Applications* (Revised.). New York: George Braziller Inc. Retrouvé de http://www.scribd.com/document_downloads/direct/93285373?extension=pdf&ft=1379037074<=1379040684&user_id=15156098&uahk=9CN8bwqGgkwbS1bwW5l22CpmamI
- Bertrand, A., et Garnier, P.-H. (Aut). (2005). *Psychologie cognitive*. Levallois-Perret: Studyrama.
- Bézivin, J., et Gerbé, O. (Aut). (2001). Towards a precise definition of the OMG/MDA framework. Dans *Automated Software Engineering, 2001.(ASE 2001). Proceedings. 16th Annual International Conference on* (p. 273–280). Retrouvé de http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=989813
- Bizer, C., Heath, T., et Berners-Lee, T. (Aut). (2009). Linked data-the story so far. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 5(3), 1–22. Retrouvé de <http://www.igi-global.com/article/linked-data-story-far/37496>
- Bjørnåvold, J., et European Centre for the Development of Vocational Training (Aut). (2000). *Making learning visible: identification, assessment and recognition of non-formal learning in Europe*. Thessaloniki : Luxembourg: Cedefop--European Centre for the Development of Vocational Training ; Office for Official Publications of the European Communities [distributeur].
- Blom, P. (Aut). (2005). *Enlightening the world: Encyclopédie, the book that changed the course of history* (1st Palgrave Macmillan ed.). New York: Palgrave Macmillan.
- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., et Krathwohl, D. R. (Aut). (1956). Taxonomy of educational objectives handbook I: cognitive domain. *New York: David McKay*, 19, 56. Retrouvé de <http://www.professorevans.net/uploads/EvansAppliedReadingGuide.pdf>
- Bonniol, J.-J., et Vial, M. (Aut). (2009). *Les modèles de l'évaluation: textes fondateurs avec commentaires*. De Boeck Supérieur. Retrouvé de http://books.google.fr/books?hl=en&lr=&id=Js8jZ0ytyacC&oi=fnd&pg=PA15&dq=%E2%80%9DL%E2%80%99%C3%A9valuation+en+formation%E2%80%9D&ots=4dr9P2_vHC&sig=m6PTByi5MghwvrisDVMSKIWpfig
- Bonwell, C. C. (Aut). (1991). *Active learning: creating excitement in the classroom*. Washington, DC: School of Education and Human Development, George Washington University.
- Borst, W. N. (Aut). (1997, septembre 5). *Construction of Engineering Ontologies for Knowledge Sharing and Reuse* (info:eu-repo/semantics/doctoralThesis). Universiteit Twente, Enschede. Retrouvé de <http://doc.utwente.nl/17864/>
- Bos, B. (Aut). (2005, août 17). The XML data model. Retrouvé 18 septembre 2013, de <http://www.w3.org/XML/Datamodel.html>
- Boulding, K. (Aut). (1956). General Systems Theory - the skeleton of science. Retrouvé 20 juin 2013, de <http://www.panarchy.org/boulding/systems.1956.html>
- Boussagol, H. (Aut). (1996). Des systèmes informatiques aux systèmes d'informations. *SOSI CNRS*.

- Brachman, R. J., et Levesque, H. J. (Aut). (2004). *Knowledge Representation and Reasoning*. Morgan Kaufmann. Retrouvé de http://books.google.com/books?hl=fr&lr=&id=OuPtLaA5QjoC&oi=fnd&pg=PP2&dq=knowledge+representation+and+reasoning&ots=LFMalMUWpr&sig=JuX-U_RnwO_-H_nr-kak8fEpGw
- Brank, J., Grobelnik, M., et Mladenić, D. (Aut). (2005). A survey of ontology evaluation techniques. Retrouvé de <http://eprints.pascal-network.org/archive/00001198/>
- Bratt, S. (Aut). (2007). *Semantic Web, and Other Technologies to Watch*. Retrouvé de [http://www.w3.org/2007/Talks/0130-sb-W3CTechSemWeb/#\(1\)](http://www.w3.org/2007/Talks/0130-sb-W3CTechSemWeb/#(1))
- Bray, T., Paoli, J., Sperberg-McQueen, C. M., Maler, E., et Yergeau, F. (Aut). (2008, novembre 26). Extensible Markup Language (XML) 1.0 (Fifth Edition). Retrouvé 18 septembre 2013, de <http://www.w3.org/TR/xml/>
- Bréal, M. (Aut). (1883). Les lois intellectuelles du langage. Fragment de sémantique. *Annuaire de l'Association pour l'encouragement des études grecques en France*, 17, 132–42. Retrouvé de http://projects.chass.utoronto.ca/langueXIX/breal/breal_lois.htm
- Bréal, M. (Aut). (1897). *Essai de Sémantique (science des significations)*. Paris : Hachette. Retrouvé de <http://archive.org/details/essaidesmantiq00bruoft>
- Bréal, M. (Aut). (1991). *The beginnings of semantics: essays, lectures, and reviews*. Stanford, Calif: Stanford University Press.
- Breslin, J. G., Passant, A., et Decker, S. (Aut). (2009). *The Social Semantic Web*. Springer. Retrouvé de http://books.google.fr/books?hl=en&lr=&id=gxUuB_Tq5p8C&oi=fnd&pg=PA1&dq=the+social+semantic+web&ots=cOwrID-noS&sig=05SXRSCC6ZL6mOfQZuLP48vKxHc
- Brickley, D., et Guha, R. V. (Aut). (2004, février 10). RDF Vocabulary Description Language 1.0: RDF Schema. Retrouvé 19 septembre 2013, de <http://www.w3.org/TR/rdf-schema/>
- Brossard, A., et Perret-Clermont, A.-N. (Aut). (1985). L'intrication des processus cognitifs et des relations interpersonnelles dans les interactions sociales : premiers résultats empiriques à partir de l'étude du regard. Présenté au Dossiers de Psychologie. Retrouvé de <http://doc.rero.ch/record/9527?ln=fr>
- Bruner, J. S. (Aut). (1966). *Toward a Theory of Instruction*. Harvard University Press.
- Bruner, J. S. (Aut). (1986). *Actual minds, possible worlds*. Cambridge, Mass.: Harvard University Press.
- Brusilovsky, P. (Aut). (2000). Adaptive hypermedia: From intelligent tutoring systems to Web-based education. Dans *Intelligent Tutoring Systems* (p. 1–7). Retrouvé de http://link.springer.com/chapter/10.1007/3-540-45108-0_1
- Brusilovsky, P. (Aut). (2001). Adaptive hypermedia. *User modeling and user-adapted interaction*, 11(1-2), 87–110. Retrouvé de <http://link.springer.com/article/10.1023/A%3A1011143116306>
- Brusilovsky, P., Karagiannidis, C., et Sampson, D. (Aut). (2004). Layered evaluation of adaptive learning systems. *International Journal of Continuing Engineering Education and Life Long Learning*, 14(4), 402–421. Retrouvé de <http://inderscience.metapress.com/content/QNDCH7P1LMJ4MN6G>
- Bull, S. (Aut). (1997). See yourself write: a simple student model to make students think. Dans *Sixth International Conference Springer* (p. 315–326). Wien New York: A.

- Jameson, C. Paris & C. Tasso (eds). Retrouvé de <http://www.eee.bham.ac.uk/bull/papers-pdf/UM97-SYW.pdf>
- Bull, S. (Aut). (2004). Supporting Learning with Open Learner Models. *Planning*, 29(14), 1. Retrouvé de <http://ile2005-t3.tripod.com/text/support-OLM.pdf>
- Bull, S., et Kay, J. (Aut). (2010). Open learner models. Dans *Advances in Intelligent Tutoring Systems* (p. 301–322). Springer. Retrouvé de http://link.springer.com/chapter/10.1007/978-3-642-14363-2_15
- Bull, S., et Pain, H. (Aut). (1995). « Did I Say what I Think I Said, and Do You Agree with Me? »: Inspecting and Questioning the Student Model. *DAI RESEARCH PAPER*. Retrouvé de <http://www.dai.ed.ac.uk/papers/documents/rp760.html>
- Burton-Jones, A. (Aut). (1999). *Knowledge capitalism: business, work, and learning in the new economy*. Oxford [England]; New York: Oxford University Press.
- Bush, V. (Aut). (1945). As we may think. *The Atlantic Monthly*, 176(1), 112–124.
- Business and MOOCs*. (2013). Retrouvé de http://www.youtube.com/watch?v=DGaUfWkJdi4&feature=youtube_gdata_player
- Butoianu, V., Catteau, O., Vidal, P., et Broisin, J. (Aut). (2011). Un Système à Base de Traces pour la Recherche Personnalisée d'Objets Pédagogiques: le cas d'Ariadne Finder. *Atelier« Personnalisation de l'apprentissage: quelles approches pour quels besoins? »*, EIAH 2011. Retrouvé de http://liris.cnrs.fr/~mlefevre/ActesAtelierEIAH2011/EIAH2011_AtelierPersonnalisationApprentissage_Approche_SystemeBaseTraces.pdf
- Butoianu, V., Vidal, P., Verbert, K., Duval, E., et Broisin, J. (Aut). (2010). User Context and Personalized Learning : a Federation of Contextualized Attention Metadata. *Computer*, 16(16), 2252–2271.
- Byte-sized graphic guide to data storage. (2013, juin 21). *BBC Future*. Retrouvé 26 juin 2013, de <http://www.bbc.com/future/story/20130621-byte-sized-guide-to-data-storage>
- Calabretto, S. (Aut). (2011). Ontology Alignment in the Urban Domain. Dans *Ontologies in Urban Development Projects* (p. 55–68). Springer. Retrouvé de http://link.springer.com/chapter/10.1007/978-0-85729-724-2_4
- Carver, C. A., Howard, R. A., et Lane, W. D. (Aut). (1999). Enhancing student learning through hypermedia courseware and incorporation of student learning styles. *Education, IEEE Transactions on*, 42(1), 33–38. Retrouvé de http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=746332
- Cascading Style Sheets. (s. d.). Retrouvé 23 septembre 2013, de <http://www.w3.org/Style/CSS/Overview.en.html>
- Cases on Web 2.0 in developing countries: studies on implementation, application, and use*. (2012).
- Cawley, C. (Aut). (2012, avril 25). 10 Most Popular Content Management Systems Online. *MakeUseOf*. Retrouvé 26 juillet 2013, de <http://www.makeuseof.com/tag/10-popular-content-management-systems-online/>
- Cerf, V., Dalal, Y., et Sunshine, C. (Aut). (1974, décembre). Specification of Internet Transmission Control Program. Retrouvé 2 juillet 2013, de <http://tools.ietf.org/html/rfc675>
- CERTU (Aut). (2007, février). Une introduction à l'approche systémique : appréhender la complexité. Retrouvé 29 avril 2013, de http://www.certu-catalogue.fr/catalog/product/view/id/827/?__SID=U&link=406&link=406

- Chekuri, C., Goldwasser, M. H., Raghavan, P., et Upfal, E. (Aut). (1997). Web search using automatic classification. Dans *Proceedings of the Sixth International Conference on the World Wide Web*. Retrouvé de <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.125.7608&rep=rep1&type=pdf>
- Claparède, E. (Aut). (1931). L'éducation fonctionnelle. Retrouvé de <http://psycnet.apa.org/psycinfo/1931-04613-000>
- Clark, C. (Aut). (2008). *Classroom skills for nurse educators*. Jones & Bartlett Learning. Retrouvé de http://books.google.com/books?hl=en&lr=&id=voBvGmXfJdcC&oi=fnd&pg=PR1&dq=%22two+effective+learning+systems%22&ots=t_M0amJtDB&sig=vr1UfHnXVLamMyslRaDZJbWiutg
- Clark, D. (Aut). (2004). Understanding & Performance. Retrouvé 23 juillet 2013, de <http://www.nwlink.com/~donclark/performance/understanding.html>
- Clément, J. (Aut). (1995). Du texte à l'hypertexte: vers une épistémologie de la discursivité hypertextuelle. *Balpe, J.-P.; A. Lelu; I. Saleh (coords.). Hypertextes et hypermédias: Réalisations, poutils, methods. Paris: Hermès* <<http://hypermedia.univ-paris8.fr/jean/articles/discursivite.htm>>. Retrouvé de <http://hypermedia.univ-paris8.fr/jean/articles/discursivite.htm>
- CNRTL, (Centre National de Ressources Textuelles et Lexicales) (Aut). (s. d.). SYSTÉMIQUE : Définition de SYSTÉMIQUE. Retrouvé 19 juin 2013, de <http://www.cnrtl.fr/definition/syst%C3%A9mique>
- Cobb, J. (Aut). (2009). Definition of Learning | Define Lifelong Learning. *A Definition of Learning*. Retrouvé 27 février 2013, de <http://www.missiontolearn.com/2009/05/definition-of-learning/>
- Commission des Communautés Européennes (Aut). (2001a, mars 28). Communication de la Commission au Conseil et au Parlement Européen: Plan d'action eLearning; Penser l'éducation de demain. Retrouvé 8 avril 2013, de <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2001:0172:FIN:FR:PDF>
- Commission des Communautés Européennes (Aut). (2001b, novembre 21). Réaliser un espace européen de l'éducation et de formation tout au long de la vie. Retrouvé 2 mai 2013, de <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2001:0678:FIN:FR:PDF>
- Commission Européenne (Aut). (2002, juin). Rapport Européen sur les indicateurs de qualité de l'éducation et la formation tout au long de la vie. Retrouvé 2 mai 2013, de http://ec.europa.eu/education/lifelong-learning-policy/doc/policy/qualityreport_fr.pdf
- Commission générale de terminologie et de néologie (Aut). (1999, avril 8). Vocabulaire de l'informatique et de l'internet. *Bulletin Officiel de l'Education Nationale*. Retrouvé 15 août 2013, de <http://www.education.gouv.fr/botexte/bo990408/CTNX9903444K.htm>
- Communiquer a distance : le telegraphe Chappe. (s. d.). Retrouvé 26 juin 2013, de <http://telegraphe-chappe.com/chappe/communiquer.html>
- Concepts définissant un système | L'approche systémique. (s. d.). Retrouvé 19 juin 2013, de <http://www.approche-systemique.com/definition-systeme/concepts-principes/>
- Conklin, J. (Aut). (1987). Hypertext: An Introduction and Survey, 17-41. Retrouvé de <http://www.ics.uci.edu/~andre/informatics223s2007/conklin.pdf>

- Corno, L., et Snow, R. E. (Aut). (1986). Adapting teaching to individual differences among learners. *Handbook of research on teaching*, 3, 605–629.
- Cortada, J. W., et Woods, J. A. (Aut). (2000). *The Knowledge Management Yearbook 2000-2001*. Taylor & Francis.
- Course Management Systems and the Reinvention of Instruction. (2004, avril 1). *THE Journal - Transforming Education Through Technology*. Retrouvé 1 août 2013, de <http://thejournal.com/articles/2004/10/01/course-management-systems-and-the-reinvention-of-instruction.aspx>
- Cristea, A. I. (Aut). (2004). What can the Semantic Web do for Adaptive Educational Hypermedia? *Educational Technology & Society*, 7, 40–58.
- Cristea, A. I., et de Mooij, A. (Aut). (2003). LAOS: Layered WWW AHS authoring model and their corresponding algebraic operators. Dans *WWW03 (The Twelfth International World Wide Web Conference), Alternate Track on Education, Budapest, Hungary*. Retrouvé de <http://www.shibbo.ethz.ch/CDstore/www2003/papers/alternate/P301/p301-cristea.pdf>
- Cybernétique, la science des systèmes. (s. d.). Retrouvé 9 mai 2013, de <http://www.syti.net/Cybernetics.html>
- Cyganiak, R., et Wood, D. (Aut). (2013, juillet 23). RDF 1.1 Concepts and Abstract Syntax. Retrouvé 18 septembre 2013, de <http://www.w3.org/TR/rdf11-concepts/>
- Daubigney, L., Geist, M., et Pietquin, O. (Aut). (2011). Apprentissage par renforcement pour la personnalisation d'un logiciel d'enseignement des langues. Dans *Actes de la Conférence sur les Environnements Informatiques pour l'Apprentissage Humain (EIAH 2011)* (p. 1–5). Retrouvé de <http://hal.archives-ouvertes.fr/hal-00652516/>
- Davenport, T. H., et Prusak, L. (Aut). (1998). *Working knowledge: How organizations manage what they know*. Harvard Business Press. Retrouvé de <http://books.google.com/books?hl=fr&lr=&id=QIyIWWhdYoYC&oi=fnd&pg=PR5&dq=working+knowledge&ots=Z22ChjNWWX&sig=e764Ub496i6lmYxiexVLuNUc-ek>
- Davis, R., Shrobe, H., et Szolovits, P. (Aut). (1993). What is a knowledge representation? *AI magazine*, 14(1), 17. Retrouvé de <http://www.aaai.org/ojs/index.php/aimagazine/article/viewArticle/1029>
- DBpedia.org, D. S. (Aut). (s. d.). wiki.dbpedia.org : Datasets / Dataset Statistics. Retrouvé 22 août 2013, de <http://wiki.dbpedia.org/Datasets/DatasetStatistics>
- DCMI Metadata Basics. (s. d.). Retrouvé 26 juillet 2013, de <http://dublincore.org/metadata-basics/>
- De Bra, P., Houben, G.-J., et Wu, H. (Aut). (1999). AHAM: a Dexter-based reference model for adaptive hypermedia. Dans *Proceedings of the tenth ACM Conference on Hypertext and hypermedia: returning to our diverse roots: returning to our diverse roots* (p. 147–156). Retrouvé de <http://dl.acm.org/citation.cfm?id=294508>
- De Landsheere, V. (Aut). (1992). *L'éducation et la formation*. Paris: Presses universitaires de France.
- De Rosnay, J. (Aut). (1974). Le microscope. *Paris, Seuil*, 23, 137. Retrouvé de http://cognisciences.fr/numeros/CSS_n5.pdf
- Dehors, S. (Aut). (2007, février 2). *Exploiting Semantic Web and Knowledge Management Technologies for E-learning*. Retrouvé de <http://tel.archives-ouvertes.fr/docs/00/13/41/14/PDF/These.pdf>

- Denis, B. (Aut). (1994). *Information, communication: théories et pratiques : fiches de synthèse*. Paris: Ed. d'organisation.
- Derycke, A. (Aut). (2009, décembre 28). *Quelques éléments pour la prospective : Mobilités, Ubimédia et Contextes*. Retrouvé de <http://www.slideshare.net/davidcordina/alain-derycke-lille1-lifl-mobilite-et-ubimedia>
- Develay, M. (Aut). (1992). *De l'apprentissage à l'enseignement: pour une épistémologie scolaire*. Paris: ESF.
- Devèze, J. (Aut). (2004). Abraham Moles, un exceptionnel passeur transdisciplinaire. *Hermès: Cognition-comunication-politique*, (39), 189–200. Retrouvé de <http://documents.irevues.inist.fr/handle/2042/9482>
- Dewey, E. (Aut). (1922). *The Dalton Laboratory Plan*. E. P. Dutton & Company Inc. Retrouvé de <http://ia600308.us.archive.org/7/items/daltonlaboratory00deweiala/daltonlaboratory00deweiala.pdf>
- Dibiaggio, L., et Meschi, P. X. (Aut). (2010). *Le management dans l'économie de la connaissance*. Pearson Education France.
- DicoLatin. (s. d.). *DicoLatin*. Retrouvé 6 mai 2013, de <http://www.dicolatin.com>
- Diderot, D., et d' Alembert, J. L. R. (Aut). (1751). *Encyclopédie, ou dictionnaire raisonné des sciences, des arts et des métiers, par une société de gens de lettres*. (Vol Tome premier). Retrouvé de http://cerebro.xu.edu/math/Sources/Dalembert/probability_diderot.pdf
- Dieuzeide, H. (Aut). (1994). *Les Nouvelles Technologies Outils d'enseignement* (Paris: Nathan.). UNESCO.
- Doise, W., et Mugny, G. (Aut). (1979). Individual and collective conflicts of centrations in cognitive development. *European Journal of Social Psychology*, 9(1), 105–108. doi:10.1002/ejsp.2420090110
- Donnadieu, G., et Karsky, M. (Aut). (2002). *La systémique, penser et agir dans la complexité*. Rueil-Malmaison, France: Éditions Liaisons.
- Dublincore.org (Aut). (2012a, juin 14). Dublin Core Metadata Element Set, Version 1.1. Retrouvé 28 juillet 2013, de <http://dublincore.org/documents/dces/>
- Dublincore.org (Aut). (2012b, juin 14). DCMI Metadata Terms. Retrouvé 28 juillet 2013, de <http://dublincore.org/documents/dcmi-terms/>
- Duerst, M., et Suignard, M. (Aut). (2005, janvier). Internationalized Resource Identifiers (IRIs). Retrouvé 16 septembre 2013, de <http://www.faqs.org/rfcs/rfc3987.html>
- DUMEZ, C. (Aut). (2010). *Approche dirigée par les modeles pour la spécification, la vérification formelle et la mise en oeuvre de services Web composés*. Université de Technologie de Belfort-Montbeliard. Retrouvé de <http://hal.archives-ouvertes.fr/tel-00515130/>
- Dunn, R. S., et Dunn, K. J. (Aut). (1978). *Teaching students through their individual learning styles: A practical approach*. Allyn and Bacon Boston. Retrouvé de <http://www.getcited.org/pub/101860721>
- Dunn, R. S., et Dunn, K. J. (Aut). (1993). *Teaching secondary students through their individual learning styles: Practical approaches for grades 7-12*. Allyn & Bacon.
- E-Learning Glossary. (s. d.-a). Retrouvé 27 mars 2013, de <http://www3.imperial.ac.uk/ict/services/e-learning/aboutlearning/elearningglossary>

- E-Learning Glossary. (s. d.-b). Retrouvé 1 août 2013, de <http://www3.imperial.ac.uk/ict/services/e-learning/old%20site/aboutelearning/elearningglossary#l>
- Education nouvelle et pédagogie active. (s. d.). Retrouvé 19 juillet 2013, de <http://ddata.over-blog.com/xxxyyy/3/12/82/15/ORAL-PROFESSIONNEL/pedagogies-nouvelles.pdf>
- Ehrlich, S. (Aut). (1985). Les représentations. *Psychologie française*, 30(3-4), 226–296. Retrouvé de <http://cat.inist.fr/?aModele=afficheN&cpsidt=8636014>
- Escarpit, R. (Aut). (1991). *L'information et la communication: théorie générale*. Librairie Hachette.
- Europa.eu (Aut). (2007). Compétences clés pour l'éducation et la formation tout au long de la vie. Retrouvé 30 août 2013, de http://ec.europa.eu/dgs/education_culture/publ/pdf/ll-learning/keycomp_fr.pdf
- Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies. (2010). Retrouvé 25 mars 2013, de <http://www2.ed.gov/rschstat/eval/tech/evidence-based-practices/finalreport.pdf>
- Feigenbaum, E. A., et McCorduck, P. (Aut). (1983). *The fifth generation*. Addison-Wesley Reading. Retrouvé de <https://saltworks.stanford.edu/assets/vx422bg1298.pdf>
- Felder, R. M., et Silverman, L. K. (Aut). (1988). Learning and teaching styles in engineering education. *Engineering education*, 78(7), 674–681. Retrouvé de <http://www4.ncsu.edu/unity/lockers/users/f/felder/public/Papers/LS-1988.pdf>
- Fensel, D. (Aut). (2001). Ontologies. Dans *Ontologies* (p. 11-18). Springer Berlin Heidelberg. Retrouvé de http://link.springer.com/chapter/10.1007/978-3-662-04396-7_2
- Fernández-López, M., Gómez-Pérez, A., et Juristo, N. (Aut). (1997). Methontology: from ontological art towards ontological engineering. Retrouvé de <http://oa.upm.es/5484/>
- Ferrière, A. (Aut). (1922). L'école active.
- Finin, T. (Aut). (2001, décembre 27). Re: NAME: SWOL versus WOL from tim finin on 2001-12-27 (www-webont-wg@w3.org from December 2001). Retrouvé de <http://lists.w3.org/Archives/Public/www-webont-wg/2001Dec/0169.html>
- Fink, L. D. (Aut). (2003). *Creating significant learning experiences : an integrated approach to designing college courses*. San Francisco, Calif.: Jossey-Bass.
- Floridi, L. (Aut). (2009). Web 2.0 contre Web sémantique : un point de vue philosophique | Knowtex. *Knowtex*. Retrouvé 4 juillet 2013, de http://www.knowtex.com/nav/web-2-0-contre-web-semantique-un-point-de-vue-philosophique_2405
- France Université Numérique - FUN - Découvrir, apprendre et réussir. (s. d.). Retrouvé 5 février 2014, de <http://www.france-universite-numerique.fr/>
- Frankel, D. (Aut). (2003). *Model driven architecture: applying MDA to enterprise computing*. New York: Wiley.
- Frankel, D., Hayes, P., Kendall, E., et McGuinness, D. (Aut). (2004). The model driven semantic web. Dans *1st International Workshop on the Model-Driven Semantic Web (MDSW2004), Monterey, California, USA*. Retrouvé de <ftp://www.ufv.br/dpi/mestrado/Gerais/modeldivensemanticweb.pdf>

- Frayssinhes, J. (Aut). (2012). *L'apprenant adulte à l'ère du numérique: Ou l'art de développer de nouveaux savoir-faire en situation d'apprentissage contextualisé*. Editions L'Harmattan.
- Frege, G. (Aut). (1879). Begriffsschrift, a formula language, modeled upon that of arithmetic, for pure thought. Retrouvé 30 septembre 2013, de <http://dec59.ruk.cuni.cz/~kolmanv/Begriffsschrift.pdf>
- Freinet, C. (Aut). (1954). *L'école moderne française*.
- Friesen, N. (Aut). (2004). The International Learning Object Metadata Survey. *The International Review of Research in Open and Distance Learning*, 5(3). Retrouvé de <http://www.irrodl.org/index.php/irrodl/article/view/195>
- Frigg, R., et Hartmann, S. (Aut). (2012). Models in Science. Dans E. N. Zalta (Éd), *The Stanford Encyclopedia of Philosophy* (Fall 2012.). Retrouvé de <http://plato.stanford.edu/archives/fall2012/entries/models-science/>
- Full Hierarchy - schema.org. (s. d.). Retrouvé 28 juillet 2013, de <http://schema.org/docs/full.html>
- Gandon, F., Faron-Zucker, C., et Corby, O. (Aut). (2012). *Le Web sémantique: comment lier les données et les schémas sur le Web ?* Paris: Dunod.
- Gardner, H. (Aut). (1985). *The mind's new science: a history of the cognitive revolution*. New York: Basic Books.
- Gayeski, D. M. (Aut). (1995). *Designing Communication and Learning Environments*. Educational Technology.
- George, S., et Derycke, A. (Aut). (2005). Conceptions et usages des plates-formes de formation. *Revue STICEF*, 12. Retrouvé de http://edutice.archives-ouvertes.fr/docs/00/19/58/53/PDF/S01_STICEF_EDITO_GEORGE_versions_auteurs.pdf
- Giordan, A. (Aut). (1998). *Apprendre!* Paris: Belin.
- Gómez-Pérez, A. (Aut). (1999). Evaluation of taxonomic knowledge in ontologies and knowledge bases. Retrouvé de http://oa.upm.es/6456/1/Evaluation_of_Taxonomic_K.pdf
- Gomez-Pérez, A., Fernandez-Lopez, M., et Corcho, O. (Aut). (2003). *Ontological engineering: with examples from the areas of knowledge management, e-commerce and the semantic Web* (Vol 1-1). London, Royaume-Uni.
- Gonzalez, C. (Aut). (2004). The Role of Blended Learning in the World of Technology. Retrouvé de <http://www.unt.edu/benchmarks/archives/2004/september04/eis.htm>
- Grandbastien, M., Huynh-Kim-Bang, B., et Monceaux, A. (Aut). (2009). Knowledge framework supporting semantic search of learning resources. Dans *Metadata and Semantics* (p. 259–268). Springer. Retrouvé de http://link.springer.com/chapter/10.1007/978-0-387-77745-0_25
- Grant, M. (Aut). (2012, février 23). *CMS, LMS & LCMS*. Technology. Retrouvé de <http://www.slideshare.net/msquareg/cms-lms-lcms>
- Great New Jobs That Didn't Exist 10 Years Ago. (2011, février 1). *AOL Jobs*. Retrouvé 3 mai 2013, de <http://jobs.aol.com/articles/2011/02/01/great-new-jobs-that-didnt-exist-10-years-ago/>
- Grishman, R., et Sundheim, B. (Aut). (1996). Message Understanding Conference-6: A Brief History. Dans *COLING* (Vol 96, p. 466–471). Retrouvé de <http://acl.ldc.upenn.edu/C/C96/C96-1079.pdf>
- Gruber, T. R. (Aut). (1993). A translation approach to portable ontology specifications. *Knowledge acquisition*, 5(2), 199–220. Retrouvé de

- <http://secs.ceas.uc.edu/~mazlack/ECE.716.Sp2011/Semantic.Web.Ontology.Papers/Gruber.93a.pdf>
- Grunert, J. (Aut). (1997). *The course syllabus: a learning-centered approach*. Bolton, MA: Anker Pub. Co.
- Guarino, N. (Aut). (1997). Semantic matching: Formal ontological distinctions for information organization, extraction, and integration. Dans *Information Extraction A Multidisciplinary Approach to an Emerging Information Technology* (p. 139–170). Springer. Retrouvé de http://link.springer.com/chapter/10.1007/3-540-63438-X_8
- Guarino, N., et Welty, C. (Aut). (2000). Towards a methodology for ontology-based model engineering. Dans *Proceedings of the ECOOP-2000 Workshop on Model Engineering*. Retrouvé de <http://cuiwww.unige.ch/isi/cours/aftsi/articles/01b-guarino00towards.pdf>
- Guéraud, V., Adam, J.-M., Pernin, J.-P., Calvary, G., et David, J.-P. (Aut). (2004). L'exploitation d'Objets Pédagogiques Interactifs à distance: le projet FORMID. *Revue des Sciences et Technologies de l'Information et de la Communication pour l'Education et la Formation (STICEF)*, 11. Retrouvé de <http://halshs.archives-ouvertes.fr/hal-00696400/>
- Guyot, B. (Aut). (2011). Introduction à l'ingénierie documentaire et aux sciences de l'information. Retrouvé de http://archivesic.ccsd.cnrs.fr/sic_00665267
- Halasz, F., Schwartz, M., Grønbaek, K., et Trigg, R. H. (Aut). (1994). The Dexter hypertext reference model. *Communications of the ACM*, 37(2), 30–39. Retrouvé de <http://dl.acm.org/citation.cfm?id=175237>
- Halpin, H. (Aut). (2006). Web Science and the European Web 2.0: Centralization of Decentralization? Retrouvé 5 juillet 2013, de <http://www.ibiblio.org/hhalpin/homepage/presentations/de2007/>
- Hannafin, M., Land, S., et Oliver, K. (Aut). (1999). Open learning environments: Foundations, methods, and models. *Instructional-design theories and models: A new paradigm of instructional theory*, 2, 115–140. Retrouvé de http://books.google.com/books?hl=en&lr=&id=AbJc4Kg6XQoC&oi=fnd&pg=PA115&dq=Hannafin,+Land,+%26+Oliver,+1999+Open+Learning+Environments:+Foundations,+methods,+and+models&ots=WAntm9Scqy&sig=UMQB-BKVM_oYai1GHltOUQB8nuw
- Hansen, C., et McCalla, G. (Aut). (2003). Active open learner modelling. Dans *Proceedings of Learner Modelling for Reflection Workshop International Conference on Artificial Intelligence in Education*. Retrouvé de <http://www.eee.bham.ac.uk/bull/papers-pdf/AIED03-WS-proceedings.pdf#page=56>
- Harald Sack (Aut). (2013, février 22). *OpenHPI 4.4 - Ontology Types*. Retrouvé de <http://www.slideshare.net/lysander07/open-hpi-semweb04part4>
- Hayes, P., et McBride, B. (Aut). (2004, février 10). RDF Semantics. Retrouvé 19 septembre 2013, de <http://www.w3.org/TR/rdf-mt/>
- Heaton, J. (Aut). (2011, décembre 3). The Free and Open Software Behind IBM's Jeopardy Champion Watson | Heaton Research. Retrouvé 20 septembre 2013, de <http://www.heatonresearch.com/content/free-and-open-software-behind-ibm%E2%80%99s-jeopardy-champion-watson>
- Heinderyckx, F. (Aut). (2002). *Une introduction aux fondements théoriques de l'étude des médias* (2e édition., Vol 1). Liège: Editions du Céfal. Retrouvé de <http://books.google.com/books?hl=fr&lr=&id=6ZYDC2LhuaQC&oi=fnd&pg=P>

A5&dq=Une+introduction+aux+fondements+th%C3%A9oriques+de+l%27%C3%A9tude+des+m%C3%A9dias&ots=wp2ApSyQEQ&sig=eMFvf2Fum-HZV1U_LTnbTXW801A

- Helen Parkhurst (Aut). (1922). *Education On The Dalton Plan*. E. P. Dutton & Company Inc. Retrouvé de <http://archive.org/details/educationontheda028244mbp>
- Henri La Fontaine - Biographical. (s. d.). Retrouvé 26 juin 2013, de http://www.nobelprize.org/nobel_prizes/peace/laureates/1913/fontaine-bio.html
- Henry George Liddell, Robert Scott, A Greek-English Lexicon, σύστημα-α. (s. d.). Retrouvé 29 avril 2013, de <http://www.perseus.tufts.edu/hopper/text?doc=Perseus%3Atext%3A1999.04.0057%3Aentry%3Dsu%2Fsthma>
- Henze, N., Dolog, P., et Nejd, W. (Aut). (2004). Reasoning and Ontologies for Personalized E-Learning in the Semantic Web. *Educational Technology & Society*, 7, 82-97.
- Herman, I. (Aut). (2008). *Qu'est-ce le Web sémantique ?* Retrouvé de <http://www.w3.org/People/Ivan/CorePresentations/HighLevelIntro/Slides-fr.pdf>
- Herman, R. L. (Aut). (2012). The MOOCs are Coming. *The Journal of Effective Teaching*, 12(2), 1-3. Retrouvé de http://www.uncw.edu/cte/ET/articles/Vol12_2/Volume1202.pdf
- Hilgard, E. R. (Aut). (1980). The trilogy of mind: Cognition, affection, and conation. *Journal of the History of the Behavioral Sciences*, 16(2), 107-117. doi:10.1002/1520-6696(198004)16:2<107::AID-JHBS2300160202>3.0.CO;2-Y
- Histoire de la Systémique. (s. d.). Retrouvé 29 avril 2013, de <http://web.archive.org/liveweb/http://www.systemique.com/la-systemique/decouvrir/histoire-de-la-systemique.html>
- History of e-Learning. (s. d.). Retrouvé 26 mars 2013, de http://www.leerbeleving.nl/wbts/1/history_of_elearning.html
- History of Ontology from Suárez to Kant (1597-1781). (s. d.). Retrouvé 24 septembre 2013, de <http://www.ontology.co/history.htm>
- Hitzler, P., Krötzsch, M., et Rudolph, S. (Aut). (2009). *Foundations of Semantic Web technologies*. Boca Raton: CRC Press.
- Hodges, W. (Aut). (2013). Model Theory. Dans E. N. Zalta (Éd), *The Stanford Encyclopedia of Philosophy* (Fall 2013.). Retrouvé de <http://plato.stanford.edu/archives/fall2013/entries/model-theory/>
- Hoel, T., et Mason, J. (Aut). (2011). Expanding the Scope of Metadata and the Issue of Quality. Dans *T Hirashima et al (Eds)(2011) Proceedings of the 19th International Conference on Computers in Education Chiang Mai, Thailand: Asia-Pacific Society for Computers in Education*. Retrouvé de http://www.hoel.nu/publications/ICCE_workshop_paper_Hoel_Mason2011-final.pdf
- Holt, P., Dubs, S., Jones, M., et Greer, J. (Aut). (1994). The State of Student Modelling. Dans J. E. Greer et G. I. McCalla (Éd), *Student Modelling: The Key to Individualized Knowledge-Based Instruction* (p. 3-35). Springer Berlin Heidelberg. Retrouvé de http://link.springer.com/chapter/10.1007/978-3-662-03037-0_1
- Home - schema.org. (s. d.). Retrouvé 28 juillet 2013, de <http://schema.org/>

- Honey, P., et Mumford, A. (Aut). (1992). The manual of learning styles. *Maidenhead: Peter Honey Publications Ltd.*
- Horn clause logic. (s. d.). Retrouvé 23 octobre 2013, de <http://cs.nyu.edu/courses/spring03/G22.2560-001/horn.html>
- Houdé, O., Kayser, D., Koenig, O., Proust, J., et Rastier, F. (Aut). (1998). *Vocabulaire de sciences cognitives*. Presses Univ. de France. Retrouvé de <http://www.lavoisier.fr/livre/notice.asp?id=3R2WOKA6K000WX>
- Houssaye, J. (Aut). (2000). *Le triangle pédagogique* (3ème édit.). Bern: Peter Lang.
- How Napoleon's semaphore telegraph changed the world. (2013, juin 16). *BBC*. Retrouvé de <http://www.bbc.co.uk/news/magazine-22909590>
- Huhns, M. N., et Singh, M. P. (Aut). (1997). Ontologies for agents. *Internet Computing, IEEE*, 1(6), 81-83. Retrouvé de http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=643942
- IBM What is big data? - Bringing big data to the enterprise. (2013, juin 12). CT000. Retrouvé 7 juillet 2013, de <http://www-01.ibm.com/software/data/bigdata/>
- IBM100 - Deep Blue. (2012, mars 7). CTB14. Retrouvé 26 août 2013, de <http://www-03.ibm.com/ibm/history/ibm100/us/en/icons/deepblue/>
- IEEE (Aut). (2002). Draft Standard for Learning Object Metadata, ((LOM) - Document IEEE P1484.12.1-2002), 44.
- IEEE (Aut). (2006). IEEE 1484.20.1/Draft 4. Draft Standard for Learning Technology Standard for Reusable Competency Definitions. Retrouvé de http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1378611
- In the 21st Century, Data Is King*. (2013). Retrouvé de http://www.youtube.com/watch?v=71Ej41sw6fw&feature=youtube_gdata_player
- Inria - Michel Serres - 2007 - Intro*. (2010). Retrouvé de <http://www.youtube.com/watch?v=kRFXFDmqCqY>
- Inria - Michel Serres - 2007 - Les Hommes : le cognitif (1)*. (2010). Retrouvé de http://www.youtube.com/watch?v=1wsYbQxppRc&feature=youtube_gdata_player
- Inria - Michel Serres - 2007 - Les Hommes : le cognitif (3)*. (2010). Retrouvé de http://www.youtube.com/watch?v=w5OqlbrXiOE&feature=youtube_gdata_player
- Jackson, P. (Aut). (1999). *Introduction to expert systems* (3rd ed.). Harlow, England ; Reading, Mass: Addison-Wesley.
- Jacobs, I., et Walsh, N. (Aut). (2004, décembre 15). Architecture of the World Wide Web, Volume One. Retrouvé 28 juin 2013, de <http://www.w3.org/TR/2004/REC-webarch-20041215/#def-information-resource>
- Jeopardy - Watson vs. The Humans - Day 1*. (2011). Retrouvé de http://www.youtube.com/watch?v=yJptrlCVDHI&feature=youtube_gdata_player
- Jonassen, D. H., et Grabowski, B. L. (Aut). (1993). *Handbook of individual differences: Learning & instruction*. Hillsdale, NJ: Lawrence Earlbaum Associates. ISBN: 0-8058-1412-4. 0-8058-1413-2.
- Kapoor, B., et Sharma, S. (Aut). (2010). A comparative study ontology building tools for semantic web applications. *International journal of Web & Semantic Technology (IJWest)*, 1(3). Retrouvé de <http://www.airccse.org/journal/ijwest/papers/0710ijwest01.pdf>

- Kaushik, D. (Aut). (s. d.). IPv6.com - IPv6 History and Timeline. Retrouvé 26 juin 2013, de <http://www.ipv6.com/articles/general/timeline-of-ipv6.htm>
- Kay, J., Halin, Z., Ottomann, T., et Razak, Z. (Aut). (1997). Learner know thyself: student models to give learner control and responsibility. Dans *Proceedings of International Conference on Computers in Education* (p. 17–24). Retrouvé de <http://www.it.usyd.edu.au/~judy/Homec/Pubs/icce.pdf>
- Ketele, J.-M. de (Aut). (1993). L'évaluation conjugquée en paradigmes. *Revue française de pédagogie*, 103(1), 59-80. doi:10.3406/rfp.1993.1298
- Ketele, J.-M. D., Chastrette, M., et Cros, D. (Aut). (1989). *Guide du fomateur*. De Boeck & Larcier s.a.
- Kifer, M., et Boley, H. (Aut). (2013, février 5). RIF Overview (Second Edition). Retrouvé 22 octobre 2013, de <http://www.w3.org/TR/2013/NOTE-rif-overview-20130205/>
- Kobielus, J. (Aut). (2013, mai 9). Measuring the Business Value of Big Data. Retrouvé 1 août 2013, de <http://www.ibmbigdatahub.com/blog/measuring-business-value-big-data>
- Kolb, D. A. (Aut). (1973). On management and the learning process. Retrouvé de <http://dspace.mit.edu/handle/1721.1/48673>
- Kolb, D. A. (Aut). (1984). *Experiential learning: experience as the source of learning and development*. Englewood Cliffs, New Jersey: Prentice Hall. Retrouvé de <http://www.learningfromexperience.com/images/uploads/process-of-experiential-learning.pdf> ! (date of download: 31.05.2006)
- Koohang, A., et Harman, K. (Aut). (2007). *Learning Objects and Instructional Design*. Informing Science.
- Korzybski, A. (Aut). (1958). *Science and sanity: An introduction to non-Aristotelian systems and general semantics*. Institute of GS. Retrouvé de <http://books.google.com/books?hl=fr&lr=&id=KN5gvaDwrGcC&oi=fnd&pg=P R13&dq=an+introduction+to+non-aristotelian+systems+and+general+semantics&ots=r021Bues87&sig=w9qSYS ruxBwGqcPWPNoaXyYYrqw>
- Krathwohl, D. R. (Aut). (2002). A revision of Bloom's taxonomy: An overview. *Theory into practice*, 41(4), 212–218. Retrouvé de http://www.tandfonline.com/doi/pdf/10.1207/s15430421tip4104_2
- Krathwohl, D. R., Bloom, B. S., et Masia, B. B. (Aut). (1964). *Taxonomy of educational objectives handbook II: affective domain*. New York: David McKay. Retrouvé de <http://library.wur.nl/WebQuery/clc/982120>
- Kurzweil, R. (Aut). (2005). *The singularity is near: when humans transcend biology*. New York: Penguin.
- L'activité d'apprentissage. (s. d.). Retrouvé 25 mars 2013, de <http://csrdsn.qc.ca/discas/Apprentissage/activiteApprentissage.html>
- La Borderie, R., Paty, J., et Sembel, N. (Aut). (2000). *Les sciences cognitives en éducation*. Paris: Nathan.
- Lassila, O., et McGuinness, D. (Aut). (2001). The role of frame-based representation on the semantic web. *Linköping Electronic Articles in Computer and Information Science*, 6(5), 2001. Retrouvé de ftp://ftp.ksl.stanford.edu/pub/KSL_Reports/KSL-01-02.html
- Laurillard, D. (Aut). (1995). Multimedia and the changing experience of the learner. *British Journal of Educational Technology*, 26(3), 179–189. doi:10.1111/j.1467-8535.1995.tb00340.x

- Lautrey, J. (Aut). (2006). *Psychologie du développement et psychologie différentielle*. Presses Universitaires de France - PUF.
- Laverde, A. C., Cifuentes, Y. S., et Rodríguez, H. Y. R. (Aut). (2007). Toward an instructional design model based on learning objects. *Educational Technology Research and Development*, 55(6), 671–681. Retrouvé de <http://link.springer.com/article/10.1007/s11423-007-9059-0>
- Le Deist, F. D., et Winterton, J. (Aut). (2005). What Is Competence? *Human Resource Development International*, 8(1), 27–46. doi:10.1080/1367886042000338227
- Le Moigne, J.-L. (Aut). (1995). *Les épistémologies constructivistes* (Vol Collection "Que sais-je?"). Paris: Presses universitaires de France.
- Le Ny, J. F. (Aut). (1989). *Science cognitive et compréhension du langage*. Presses universitaires de France.
- Le Robert, Rey-Debove, J., et Rey, A. (Aut). (2010). *Le petit Robert dictionnaire alphabétique et analogique de la langue française*. Paris: le Robert.
- Learning environment - EduTech Wiki. (s. d.). Retrouvé 27 mars 2013, de http://edutechwiki.unige.ch/en/Learning_environment
- Learning Resource Metadata Initiative :: About the LRMI. (s. d.). Retrouvé 14 août 2013, de <http://www.lrmi.net/about>
- Learning Resource Metadata Initiative :: FAQ. (s. d.). Retrouvé 14 août 2013, de <http://www.lrmi.net/about/faq>
- Learning Resource Metadata Initiative :: The Specification. (s. d.). Retrouvé 14 août 2013, de <http://www.lrmi.net/the-specification>
- Lecomte, J. (Aut). (1993). *Les mécanismes de l'apprentissage* (Vol Sciences Humaines n°32).
- Lee, Y. T. (Aut). (1999). Information modeling: From design to implementation. Dans *Proceedings of the second world manufacturing congress* (p. 315–321). Retrouvé de <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.120.5829&rep=rep1&type=pdf>
- Leeuwen, J. (Éd). (1990a). *Handbook of Theoretical Computer Science: Algorithms and complexity*. Elsevier.
- Leeuwen, J. (Éd). (1990b). *Handbook of Theoretical Computer Science: Formal models and semantics*. Elsevier.
- Lefevre, M., Broisin, J., Butoianu, V., Daubias, P., Daubigny, L., Greffier, F., ... Terrat, H. (Aut). (2012). Personnalisation de l'apprentissage: comparaison des besoins et approches à travers l'étude de quelques dispositifs. *Revue STICEF*, 19. Retrouvé de http://sticef.univ-lemans.fr/num/vol2012/06-lefevre-individualisation/sticef_2012_NS_lefevre_06p.html
- Lefevre, M., Jean-Daubias, S., et Guin, N. (Aut). (2011). Adapte, un logiciel pour aider l'enseignant à proposer des activités personnalisées à chacun de ses apprenants. *Atelier Personnalisation de l'apprentissage, EIAH*. Retrouvé de http://liris.cnrs.fr/~mlefevre/ActesAtelierEIAH2011/EIAH2011_AtelierPersonnalisationApprentissage_Approche_Adapte.pdf
- Legendre, R. (Aut). (1993). *Dictionnaire actuel de l'éducation*. Montréal: Guérin.
- Leitzelman, M., et Dou, H. (Aut). (1998). Typology of Information Systems Essai de typologie des Systèmes d'Informations. *International Journal of Information Sciences for Decision Making*, 2. Retrouvé de <http://isdmln.fr/PDF/isdmln2/isdmln2.pdf#page=57>

- Les quatre web de Joel de Rosnay, du 1.0 au 4.0. (s. d.). *Dailymotion*. Retrouvé 28 juin 2013, de http://www.dailymotion.com/video/x47z47_les-quatre-web-de-joel-de-rosnay-du_tech?start=84
- Lesourne, J. (Aut). (1976). *Les systèmes du destin*. Retrouvé de <http://documents.irevues.inist.fr/handle/2042/32241>
- Levie, F. (Aut). (2006). *L'homme qui voulait classer le monde: Paul Otlet et le Mundaneum*. Bruxelles: Impressions nouvelles.
- Levy, M. (Aut). (1997). *Computer-assisted language learning: Context and conceptualization*. Oxford University Press. Retrouvé de http://books.google.com/books?hl=en&lr=&id=RRGgrjteVjUC&oi=fnd&pg=PR8&dq=Computer-Assisted+Language+Learning,+Context+and+Conceptualization&ots=8p7hG5jAoa&sig=_RaxyAWsMzHg3bvqsGrTkLSGoll
- Lévy, P. (Aut). (1990). *Les Technologies de l'intelligence : l'avenir de la pensée à l'ère informatique*. Paris: La Découverte.
- Lewis, R. (Aut). (2007, mai 31). Dereferencing HTTP URIs. Retrouvé 17 septembre 2013, de <http://www.w3.org/2001/tag/doc/httpRange-14/2007-05-31/HttpRange-14>
- Licklider, J. C., et Taylor, R. W. (Aut). (1968). The computer as a communication device. *Science and technology*, 76(2), 1–3. Retrouvé de <http://www.cc.utexas.edu/ogs/alumni/events/taylor/licklider-taylor.pdf>
- Ligue Internationale pour l'Éducation Nouvelle. (1922). *Pour l'ère nouvelle*. *Revue internationale d'éducation nouvelle*. Retrouvé de <http://lelien.org/malonne/Textes/LIEN1.pdf>
- Limongelli, C., Sciarrone, F., Temperini, M., et Vaste, G. (Aut). (2009). Adaptive learning with the LS-plan system: a field evaluation. *Learning Technologies, IEEE Transactions on*, 2(3), 203–215. Retrouvé de http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4967566
- Linkeddata.org (Aut). (s. d.). Linked Data | Linked Data - Connect Distributed Data across the Web. Retrouvé 20 août 2013, de <http://linkeddata.org/>
- LMS and LCMS Demystified. (2006, mai 9). Retrouvé 1 août 2010, de www.brandon-hall.com/free_resources/lms_and_lcms.shtm
- Locke, J. (Aut). (1689). John Locke: An Essay Concerning Human Understanding. Retrouvé 23 juillet 2013, de <http://enlightenment.supersaturated.com/johnlocke/BOOKIVChapterI.html>
- Malcuit, G., Maurice, P., et Pomerleau, A. (Aut). (1995). *Psychologie de l'apprentissage: termes et concepts*. [Saint-Hyacinthe, Québec]: Edisem.
- Manson, P. (Aut). (2007). Technology-enhanced learning: supporting learning in the 21st century. *Technology-Enhanced Learning*, 71(3). Retrouvé de <http://www.cepis.org/upgrade/files/issue%20III-2008-manson.pdf>
- Manzano, M. (Aut). (1999). *Model Theory*. Oxford University Press.
- Markus, B. (Aut). (2005). BUILDING SPATIAL KNOWLEDGE INFRASTRUCTURE. Retrouvé de <http://www.isprs.org/proceedings/XXXVI/4-W6/papers/65-70Belamarkus-A103.pdf>
- Marton, F., et Booth, S. (Aut). (1998). The learner's experience of learning. *The handbook of education and human development : new models of learning, teaching and schooling*, 534–564.
- Masters, K. (Aut). (2011). A Brief Guide To Understanding MOOCs. *The Internet Journal of Medical Education*, 1(2). doi:10.5580/1f21

- Mattelart, A. (Aut). (2009). *Histoire de la société de l'information*. Paris: La Découverte. Retrouvé de <http://ezproxy.usherbrooke.ca/login?url=http://www.cairn.info/histoire-de-la-societe-de-l-information--9782707157980.htm>
- Mattelart, A., et Mattelart, M. (Aut). (1995). *Histoire des théories de la communication*. La Découverte. Retrouvé de <http://www.lavoisier.fr/livre/notice.asp?ouvrage=2069816>
- May, M., George, S., et Prévôt, P. (Aut). (2011). TrAVis to enhance online tutoring and learning activities: Real-time visualization of students tracking data. *Interactive Technology and Smart Education*, 8(1), 52–69. Retrouvé de <http://www.emeraldinsight.com/journals.htm?articleid=1921982&show=abstract>
- McGreal, R. (Aut). (2004). Learning objects: A practical definition. *International Journal of Instructional Technology and Distance Learning (IJITDL)*, 9(1). Retrouvé de <http://auspace.athabasca.ca/handle/2149/227>
- Meichel, F. (Aut). (2009, novembre 18). Florence Meichel BLOG: Distinguer Savoir et Connaissance. Retrouvé 4 mai 2013, de <http://florencemeichel.blogspot.fr/2009/11/distinguer-savoir-et-connaissance.html>
- Meirieu, P., Kambouchner, D., et Stiegler, B. (Aut). (2012). *L'école, le numérique et la société qui vient*. Fayard/Mille et une nuits.
- Mendelsohn, P. (Aut). (1994). Le transfert des connaissances: la pierre philosophale de l'enseignant. *Colloque International sur les transferts de connaissances en formation initiale et continue Lyon - 29 septembre 1994*. Retrouvé 27 février 2013, de http://tecfa.unige.ch/tecfa/publicat/mendel-papers/lyon_94-1.html
- Metros, S. E. (Aut). (2005). Learning objects: A rose by any other name. *Educause Review*, 40(4), 12–13.
- Mille, D. (Aut). (2005). Formalisation d'annotations produites par des apprenants. Etape préalable à la réalisation d'un annotateur (p. 401–406). Présenté au Environnements Informatique pour l'apprentissage humain, Montpellier.
- Minier, P. (Aut). (2003). Ancrage historique et développement des courants de pensée de l'apprentissage. Retrouvé 23 juillet 2013, de <http://www.wens.uqac.ca/~pminier/act1/graph1.htm>
- Mitchell, G. (Aut). (s. d.). Bertalanffy's General Systems Theory: The Topology of Mind Development. Retrouvé 23 juin 2013, de <http://www.trans4mind.com/mind-development/systems.html>
- Monteil, J.-M., et Fayol, M. (Aut). (1989). *La psychologie scientifique et ses applications*. Presses universitaires de Grenoble.
- Moore, M. G. (Aut). (1989). Editorial: Three types of interaction. Retrouvé de <http://www.tandfonline.com/doi/abs/10.1080/08923648909526659>
- Morin, E. (Aut). (1977). *La méthode, tome 1 : La Nature de la nature* (SEUIL.). Seuil.
- Morin, E. (Aut). (1992). *La méthode. Tome III. La connaissance de la connaissance. Anthropologie de la connaissance*. Éditions du Seuil.
- Morin, E. (Aut). (1999). *Les sept savoirs nécessaires à l'éducation du futur*. Paris: Seuil. Retrouvé de <http://unesdoc.unesco.org/images/0011/001177/117740fo.pdf>
- Morin, E. (Aut). (2011). *Mes philosophes*. Meaux: Germina.
- Morris, C. W. (Aut). (1938). *Foundations of the Theory of Signs* (Vol 1). University of Chicago Press. Retrouvé de <http://110.81.153.212/9/ishare.down.sina.com.cn/23904113.pdf?ssig=6ckVA>

- jWHEF&Expires=1373644800&KID=sina,ishare&ip=1373532670,82.247.123.&fn=Foundations+of+the+Theory+of+Signs%281938%29+-+Charles+W.Morris.pdf
- Mortimer, L. (Aut). (2002). Learning objects of desire: Promise and practicality. *línea*, Disponible en: <http://www.learningcircuits.org/>. [Consultado: 22-02-2005].
- Motik, B., Patel-Schneider, P. F., et Grau, B. C. (Aut). (2012, décembre 11). OWL 2 Web Ontology Language Direct Semantics (Second Edition). Retrouvé 24 octobre 2013, de <http://www.w3.org/TR/2012/REC-owl2-direct-semantics-20121211/>
- Mulligan, K. (Aut). (2000). Métaphysique et Ontologie. *Précis de Philosophie analytique*, 5-33. Retrouvé de <http://www.unige.ch/lettres/philo/enseignants/km/doc/PrecisMetaph1.pdf>
- Murugesan, S. (Aut). (2007). Understanding Web 2.0. *IT professional*, 9(4), 34–41. Retrouvé de http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4287373
- Nadeau, D., et Sekine, S. (Aut). (2007). A survey of named entity recognition and classification. *Linguisticae Investigationes*, 30(1), 3–26. Retrouvé de <http://www.ingentaconnect.com/content/jbp/li/2007/00000030/00000001/art00002>
- Nardi, D., et Brachman, R. J. (Aut). (2003). An Introduction to Description Logics. Dans *Description Logic Handbook* (p. 1–40). Retrouvé de <http://www.ce.sharif.ac.ir/courses/85-86/2/ce694/resources/root/books/dlhb-01.pdf>
- National Information Standards Organization (U.S.) (Aut). (2004). *Understanding metadata*. Bethesda, MD: NISO Press.
- Nicolle, C., et Cruz, C. (Aut). (2011). Semantic Building Information Model and Multimedia for Facility Management. Dans *Web Information Systems and Technologies* (p. 14–29). Springer. Retrouvé de http://link.springer.com/chapter/10.1007/978-3-642-22810-0_2
- Nilsson, M. (Aut). (2010). *From Interoperability to Harmonization in Metadata Standardization*. Retrouvé de <http://kth.diva-portal.org/smash/get/diva2:369527/FULLTEXT02>
- Nonaka, I., et Takeuchi, H. (Aut). (1995). *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. Oxford University Press, USA. Retrouvé de <http://books.google.com/books?hl=fr&lr=&id=B-qxrPaU1-MC&oi=fnd&pg=PA3&dq=the+knowledge+creating+company&ots=XgXJpwrGGY&sig=pglkX5BRGtXQi3fl-vJCQBj-DRY>
- Noy, N. F., et McGuinness, D. L. (Aut). (2001). *Ontology development 101: A guide to creating your first ontology*. Stanford knowledge systems laboratory technical report KSL-01-05 and Stanford medical informatics technical report SMI-2001-0880. Retrouvé de http://liris.cnrs.fr/alain.mille/enseignements/Ecole_Centrale/What%20is%20an%20ontology%20and%20why%20we%20need%20it.htm
- Nwana, H. S. (Aut). (1990). Intelligent tutoring systems: an overview. *Artificial Intelligence Review*, 4(4), 251–277. Retrouvé de <http://link.springer.com/article/10.1007/BF00168958>
- O'Reilly, T. (Aut). (2005a, septembre 30). What Is Web 2.0 - O'Reilly Media. Retrouvé 4 juillet 2013, de <http://oreilly.com/web2/archive/what-is-web-20.html>

- O'Reilly, T. (Aut). (2005b, octobre 1). Web 2.0: Compact Definition? - O'Reilly Radar. Retrouvé 5 juillet 2013, de <http://radar.oreilly.com/2005/10/web-20-compact-definition.html>
- Ogden, C. K. (Aut). (1930). *Basic English: A general introduction with rules and grammar* (Vol 29). Londres: Paul Treber & Co.
- Ogden, C. K., et Richards, I. A. (Aut). (1923). *The meaning of meaning: A Study of the Influence of Language Upon Thought and of the Science of Symbolism*. Retrouvé de <http://psycnet.apa.org/psycinfo/1923-15001-000>
- Omg.org (Aut). (2009, mai). ODM 1.0. Retrouvé 26 août 2013, de <http://www.omg.org/spec/ODM/1.0/>
- ontologie - Wiktionnaire. (s. d.). Retrouvé 24 septembre 2013, de <http://fr.wiktionary.org/wiki/ontologie>
- Otlet et la bibliologie: XXIe colloque international de l'Association internationale de bibliologie*. (2011). Editions L'Harmattan.
- Otlet, P. (Aut). (1934). *Traité de documentation: le livre sur le livre, théorie et pratique*. Editiones Mundaneum.
- Otlet, P. (Aut). (1935). *Monde: essai d'universalisme. Brussels: Mundaneum*. Retrouvé de http://www.laetusinpraesens.org/uia/docs/otlet_contents.php
- OWL 2 Web Ontology Language Document Overview (Second Edition). (s. d.). Retrouvé 2 octobre 2013, de <http://www.w3.org/TR/2012/REC-owl2-overview-20121211/>
- OWL 2 Web Ontology Language Structural Specification and Functional-Style Syntax (Second Edition). (2012, décembre 11). Retrouvé 2 octobre 2013, de <http://www.w3.org/TR/owl2-syntax/>
- Paquette, G. (Aut). (2002a). *L'ingénierie pédagogique: pour construire l'apprentissage en réseau*. Sainte-Foy: Presses de l'Université du Québec.
- Paquette, G. (Aut). (2002b). *Modélisation des connaissances et des compétences: un langage graphique pour concevoir et apprendre*. Sainte-Foy: Presses de l'Université du Québec.
- Peirce, C. C. S. (Aut). (1978). *Écrits sur le signe*. Paris: Éditions du Seuil.
- Pergnier, M. (Aut). (1997). *Du sémantique au poétique: avec Baudelaire, Cocteau, Magritte*. Editions L'Harmattan.
- Peterson, D., Gao, S., Malhotra, A., Sperberg-McQueen, C. M., et Thompson, H. S. (Aut). (2012, avril 5). W3C XML Schema Definition Language (XSD) 1.1 Part 2: Datatypes. Retrouvé 19 septembre 2013, de <http://www.w3.org/TR/xmlschema11-2/>
- Piaget, J. (Aut). (1969). *Psychologie et pédagogie*. Paris : Denoël.
- Piaget, J. (Aut). (1974). *Réussir et comprendre*. Presses universitaires de France.
- Pire, B. (Aut). (s. d.). LES LOIS DE LA PENSÉE (G. Boole). *Encyclopædia Universalis*. Retrouvé 30 septembre 2013, de <http://www.universalis.fr/encyclopedie/les-lois-de-la-pensee/>
- Poibeau, T. (Aut). (2008). *Des mots aux textes. Analyse sémantique pour l'accès à l'information*. Université Paris-Nord-Paris XIII. Retrouvé de <http://tel.archives-ouvertes.fr/tel-00436064/>
- Poletti, F., Wheeler, N. V., Petrovich, M. N., Baddela, N., Fokoua, E. N., Hayes, J. R., ... Richardson, D. J. (Aut). (2013). Towards high-capacity fibre-optic communications at the speed of light in vacuum. *Nature Photonics*, 7(4), 279-284. doi:10.1038/nphoton.2013.45
- Polyani, M. (Aut). (1966). *The tacit dimension*. Doubleday New York.

- Popescu, E. (Aut). (2008). *Dynamic adaptive hypermedia systems for e-learning*. Université de Technologie Compiègne et University of Craiova. Retrouvé de http://software.ucv.ro/~epopescu/welsa/abstract_en.pdf
- Popescu, E., Badica, C., et Moraret, L. (Aut). (2010). Accommodating Learning Styles in an Adaptive Educational System. *Informatica: An International Journal of Computing and Informatics*, 34(4), 451–462. Retrouvé de http://www.informatica.si/PDF/34-4/06_Popescu-Accommodating%20Learning%20Styles%20in%20an%20Adaptive%20Educational%20System.pdf
- Popescu, E., Trigano, P., et Badica, C. (Aut). (2007). Towards a unified learning style model in adaptive educational systems. Dans *Advanced Learning Technologies, 2007. ICALT 2007. Seventh IEEE International Conference on* (p. 804–808). Retrouvé de http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4281163
- Powel, W., et Gill, C. (Aut). (2003). Web content management systems in higher education. *Educause Quarterly*, 26(2), 43–50. Retrouvé de <http://net.educause.edu/ir/library/pdf/eqm0325.pdf>
- Prax, J.-Y. (Aut). (2000). *Le Guide du knowledge management: concepts et pratiques du management de la connaissance*. Paris: Dunod.
- PRELIMINARY NOTES - ISBD(G):General International Standard Bibliographic Description - Annotated Text - Revised Edition. (1992). Retrouvé 27 juillet 2013, de <http://archive.ifla.org/VII/s13/pubs/isbdg0.htm>
- Project Xanadu®. (s. d.). Retrouvé 16 août 2013, de <http://www.xanadu.com/>
- Psotka, J., Massey, L. D., et Mutter, S. A. (Aut). (1988). *Intelligent tutoring systems: Lessons learned*. Routledge. Retrouvé de <http://books.google.fr/books?hl=en&lr=&id=8Q8z5nV54yEC&oi=fnd&pg=PR13&dq=intelligent+tutoring+systems:+lesson+learned&ots=NrT4qaRazg&sig=mGKkwqgL9RuWbrSNMAQYv4qiv5g>
- Rabardel, P. (Aut). (1995). *Les hommes et les technologies: approche cognitive des instruments contemporains*. Paris: Armand Colin.
- Raynal, F., et Rieunier, A. (Aut). (2010). *Pédagogie, dictionnaire des concepts clés : Apprentissage, formation, psychologie cognitive*. (E. Editeur, Éd) (8e édition).
- Rebetez, C. (Aut). (2003, juillet 17). Définition des concepts de base. Retrouvé 24 mars 2013, de <http://tecfa.unige.ch/staf/staf-i/rebetez/staf11/periode1/Glossaire.pdf>
- Recherche d'information & utopie : de Bacon à Otlet, d'Otlet à Google. (2012, mai 12). *Acteurs, Réseaux & Systèmes*. Retrouvé 3 juillet 2013, de <http://guillaumesire.wordpress.com/2012/05/16/recherche-dinformation-utopie-de-bacon-a-otlet-dolet-a-google/>
- Reuchlin, M. (Aut). (1990). *Les Différences individuelles dans le développement cognitif de l'enfant*. Presses Universitaires de France - PUF.
- Rézeau, J. (Aut). (2001). *Médiatisation et médiation pédagogique dans un environnement multimédia*. Université Victor Segalen Bordeaux 2.
- Richard, J.-F. (Aut). (1990). *Les activités mentales*. Paris. Presses Universitaires de France.
- Richard, J.-F., Bonnet, C., et Ghiglione, R. (Aut). (1990). *Traité de psychologie cognitive 2: le traitement de l'information symbolique*. Dunod.
- Richey, R. C., Fields, D. C., et Foxon, M. (Aut). (2001). *Instructional design competencies: The standards*. ERIC. Retrouvé de <http://www.eric.ed.gov/ERICWebPortal/recordDetail?accno=ED453803>

- Rosnay, J. de (Aut). (1975). *Le Macroscopie: vers une vision globale*. Éditions du Seuil.
- Rowley, J. E. (Aut). (2008). *Organizing knowledge: an introduction to managing access to information* (Fourth Edition.). Ashgate Publishing, Ltd.
- Roxin, I. (Aut). (2003). *Multimédia et Web sémantique au service de l'apprentissage* (Habilitation à diriger des recherches). Université de Franche-Comté.
- Roxin, I., et Mercier, D. (Aut). (2004). *Multimédia: les fondamentaux Introduction à la représentation numérique*. Paris: Vuibert.
- Rueda, U., Larrañaga, M., Arruarte, A., et Elorriaga, J. A. (Aut). (2003). Dynamic Visualization of Student Models Using Concept Maps. Dans *Proceedings of the 11th International Conference on Artificial Intelligence in Education* (p. 89–96). Retrouvé de <http://books.google.com/books?hl=en&lr=&id=bTpLV0KLNTwC&oi=fnd&pg=PA89&dq=dynamic+visualisation+of+student+model+using+concept+maps&ots=-gFpkx3IKd&sig=lbvOj1u51BSTuwa6RELu2gi4qQk>
- Russell, S. J., et Norvig, P. (Aut). (2003). *Artificial intelligence: a modern approach*. Upper Saddle River, N.J.: Prentice Hall/Pearson Education.
- sameAs. (s. d.). Retrouvé 21 novembre 2013, de <http://sameas.org/>
- Sampson, D., et Fytros, D. (Aut). (2008). Competence models in technology-enhanced competence-based learning. Dans *Handbook on Information Technologies for Education and Training* (p. 155–177). Springer. Retrouvé de http://link.springer.com/chapter/10.1007/978-3-540-74155-8_9
- Sampson, D. G., Lytras, M. D., et Diaz, P. (Aut). (2004). Ontologies and the Semantic Web for E-learning. *Educational Technology & Society*, 7, 26–28.
- Sánchez-Ruiz, L. M., Edwards, M., et Ballester-Sarrias, E. (Aut). (2006). Competence learning challenges in Engineering Education in Spain: from theory to practice. Dans *9th International Conference on Engineering Education (ICEE)*. San Juan, Puerto Rico.
- Sandberg, J. (Aut). (1993). Educational paradigms: issues and trends. Dans *Proceedings of the IFIP TC3/WG3. 3 Working Conference on Lessons from Learning* (p. 219–230). Retrouvé de <http://dl.acm.org/citation.cfm?id=716593>
- Sanginetto, E., Capuano, N., Gaeta, M., et Micarelli, A. (Aut). (2008). Adaptive course generation through learning styles representation. *Universal Access in the Information Society*, 7(1-2), 1–23. Retrouvé de <http://link.springer.com/article/10.1007/s10209-007-0101-0>
- Sauermann, L., et Cyganiak, R. (Aut). (2008, décembre 3). Cool URIs for the Semantic Web. Retrouvé 26 juin 2013, de <http://www.w3.org/TR/cooluris/>
- Saur, K. G. (Aut). (1997a). ISBD(ER): International Standard Bibliographic Description for Electronic Resources. Retrouvé 27 juillet 2013, de <http://archive.ifla.org/VII/s13/pubs/isbd.htm>
- Saur, K. G. (Aut). (1997b). ISBD(ER): International Standard Bibliographic Description for Electronic Resources. Retrouvé 27 juillet 2013, de <http://archive.ifla.org/VII/s13/pubs/isbd2.htm>
- Schneider, D. (Aut). (1997). Advanced Learning Environments. Dans *Workshop at Unicamp. TECFA-FPSE. Université de Genève*. http://tecfa.unige.ch/tecfa/research/CMC/brazil97/psfiles/internet-advanced_fm.pdf. Retrouvé de http://129.194.9.47/tecfa/research/CMC/brazil97/psfiles/internet-advanced_fm.ps
- Seidewitz, E. (Aut). (2003). What models mean. *Software, IEEE*, 20(5), 26–32. Retrouvé de http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1231147

- Serres, M. (Aut). (2007, décembre 11). Les nouvelles technologies : révolution culturelle et cognitive - Conférence prospective. Retrouvé 23 avril 2013, de http://www.ac-grenoble.fr/ien.bourgoinashnord/IMG/pdf_Texte_de_la_conference.pdf
- Shannon, C. E. (Aut). (2001). A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5(1), 3–55. Retrouvé de <http://dl.acm.org/citation.cfm?id=584093>
- Siemens, G. (Aut). (2005). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology and Distance Learning*, 2(1), 3–10. Retrouvé de http://www.ingedewaard.net/papers/connectivism/2005_siemens_ALearningTheoryForTheDigitalAge.pdf
- Siemens, G. (Aut). (2006). *Knowing knowledge*. S.l.: www.knowingknowledge.com. Retrouvé de <http://books.google.com/books?hl=en&lr=&id=Pj41TomgKXYC&oi=fnd&pg=PR5&dq=siemens+knowing+knowledge&ots=WsqCGpzYrG&sig=fQG0ti5H7060lC1nr6X3SH0x8x0>
- Simard, C. (Aut). (2002). Normalisation de la formation en ligne: Enjeux, tendances et perspectives. *prepared for the 'Agence Universitaire de la Francophonie' (AUF)*.
- Simon, H. A. (Aut). (1974). *La science des systèmes, science de l'artificiel*. (J.-L. Le Moigne, Trad). Paris: EPI Editeurs. Retrouvé de <http://documents.irevues.inist.fr/handle/2042/29529>
- Sinclair, J., et Cardew-Hall, M. (Aut). (2008). The folksonomy tag cloud: when is it useful? *Journal of Information Science*, 34(1), 15–29. doi:10.1177/0165551506078083
- Skinner, B. F. (Aut). (1968). *The technology of teaching* (trad. française : La révolution scientifique de l'enseignement, Bruxelles : Dessart, 1968.). New York: Appleton Century Crofts: Appleton-Century-Crofts New York. Retrouvé de <http://www.tcrecord.org/library/abstract.asp?contentid=1942>
- Skinner, B. F., et Richelle, A. (Aut). (1969). *La révolution scientifique de l'enseignement*. Liège: Mardaga.
- Smith, M. K., Welty, C., et McGuinness, D. (Aut). (2004, février 10). OWL Web Ontology Language Guide. Retrouvé 1 octobre 2013, de <http://www.w3.org/TR/owl-guide/>
- Snow, R. E., et Jackson, D. N. (Aut). (1997, octobre). Individual Differences in Conation: Selected Constructs and Measures. Retrouvé 7 avril 2013, de <http://www.cse.ucla.edu/products/Reports/TECH447.pdf>
- Snow, R., et Farr, M. (Aut). (1987). *Aptitude, learning, and instruction: Vol. 3. Cognitive and affective process analyses*. Hillsdale NJ: Lawrence Erlbaum.
- Soloman, B. A., et Felder, R. M. (Aut). (2004). Index of learning styles. *www2.ncsu.edu/unity/lockers/users/f/felder/public/ILSdir/ILS-a.htm*. Retrouvé de <http://southce.org/bwebb/wp-content/uploads/2011/09/Webb-ILC-Handouts.pdf>
- Soualah-Alila, F., Mendes, F., et Nicolle, C. (Aut). (2013). A Context-Based Adaptation In Mobile Learning. *IEEE Computer Society Technical Committee on Learning Technology (TCLT)*, 15(4), 5 pages. Retrouvé de <http://hal.archives-ouvertes.fr/hal-00910239>
- Spivack, N. (Aut). (2004, avril 21). Metaweb Graph. Retrouvé 17 août 2013, de http://novaspivack.typepad.com/nova_spivacks_weblog/2004/04/new_version_of_.html

- Spivack, N. (Aut). (2007, février 9). How the WebOS Evolves? Retrouvé 19 août 2013, de http://novaspivack.typepad.com/nova_spivacks_weblog/2007/02/steps_towards_a.html
- Staab, S., Gröner, G., Parreiras, F. S., et Walter, T. (Aut). (2010). Semantic Model-driven Engineering. *Quelle: http://reasoning web. org/2010/[Zugegriffen Januar 20, 2011]*. Retrouvé de <http://www.uni-koblenz.de/~staab/Presentations/staab-ReasoningWebSummerSchool2010-SemanticModelDrivenEngineering-final.pdf>
- State of the LOD Cloud. (s. d.). Retrouvé 21 août 2013, de <http://lod-cloud.net/state/>
- Stenger, T., et Coutant, A. (Aut). (2011). *Ces réseaux numériques dits sociaux*. Paris: CNRS Editions.
- Stevens, V. (Aut). (2013). What's with the MOOCs?, 16(4). Retrouvé de <http://www.tesl-ej.org/pdf/ej64/int.pdf>
- Surowiecki, J. (Aut). (2004). *The wisdom of crowds: why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations*. New York: Doubleday.
- SYNTAXE : Définition de SYNTAXE. (s. d.). Retrouvé 25 octobre 2013, de <http://cnrtl.fr/definition/syntaxe>
- Szilagyi, I., Greffier, F., et Domenget, J.-C. (Aut). (2011). Apprentissage personnalisé via le web sémantique. *Atelier« Personnalisation de l'apprentissage: quelles approches pour quels besoins? », EIAH 2011*. Retrouvé de http://liris.cnrs.fr/~mlfevre/ActesAtelierEIAH2011/EIAH2011_AtelierPersonnalisationApprentissage_Approche_WebSemantique.pdf
- Szilagyi, I., et Roxin, I. (Aut). (2012). Learner Ontology for the Active Semantic Learning System. Dans *Advanced Learning Technologies (ICALT), 2012 IEEE 12th International Conference on* (p. 393–394). Retrouvé de http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6268130
- TAG - Fiche du terme : société de l'information. (s. d.). Retrouvé 23 avril 2013, de <http://www.thesaurus.gouv.qc.ca/tag/terme.do?id=11844>
- Tardif, J., et Presseau, A. (Aut). (1998). *Intégrer les nouvelles technologies de l'information : quel cadre pédagogique?* Paris: ESF.
- Tarski, A., Woodger, J. H., et Corcoran, J. (Aut). (1956). *Logic, semantics, metamathematics: papers from 1923 to 1938*. Clarendon Press Oxford. Retrouvé de http://www.ontology4.us/download/dot/ontopage/sub_Tars1983a.pdf
- Ted Nelson demonstrates Xanadu Space*. (2008). Retrouvé de http://www.youtube.com/watch?v=En_2T7KH6RA&feature=youtube_gdata_player
- Teguiak, H. V. (Aut). (2012). *Construction d'ontologies à partir de textes: une approche basée sur les transformations de modèles*. Ecole Nationale Supérieure de Mécanique et d'Aérotechnique-Poitiers. Retrouvé de <http://hal.archives-ouvertes.fr/tel-00786260/>
- The document that officially put the World Wide Web into the public domain on 30 April 1993. (1993, avril 30). *CERN Document Server*. Retrouvé 2 juillet 2013, de <https://cds.cern.ch/record/1164399>
- The origins of the Internet in Europe - Google Cultural Institute. (s. d.). Retrouvé 14 août 2013, de <http://www.google.com/culturalinstitute/exhibit/the-origins-of-the-internet-in-europe/QQ-RRh0A?position=8%2C12>

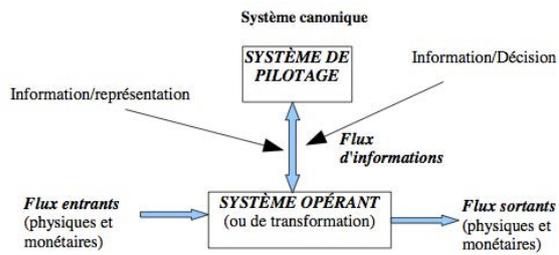
- The Ultimate Goal: Active Learning | Center for Teaching and Learning. (s. d.). Retrouvé 24 mars 2013, de <http://ctl.stanford.edu/handbook/ultimate-goal-active-learning.html>
- THE XANADU MODEL. (s. d.). Retrouvé 16 août 2013, de <http://xanadu.com/xuTheModel/>
- Tisseau, G. (Aut). (1996). *Intelligence artificielle : problèmes et méthodes*. Paris: Presses universitaires de France.
- Tobies, S. (Aut). (2001). Complexity results and practical algorithms for logics in knowledge representation. *arXiv preprint cs/0106031*. Retrouvé de <http://arxiv.org/abs/cs/0106031>
- Tomasello, M. (Aut). (2004). *Aux origines de la cognition humaine*. Paris: Retz.
- Turing, A. M. (Aut). (1950). Computing machinery and intelligence. *Mind*, 59(236), 433–460. Retrouvé de <http://www.jstor.org/stable/10.2307/2251299>
- Tyler, R. W. (Aut). (1942). General statement on evaluation. *The Journal of Educational Research*, 35(7), 492–501. Retrouvé de <http://www.jstor.org/stable/10.2307/27528279>
- Tyler, R. W. (Aut). (1949). *Basic principles of curriculum and instruction*. Chicago; London: University of Chicago Press.
- UNESCO (Aut). (2002, juin 12). Des experts évaluent l'impact des « didacticiels » ouverts pour l'enseignement supérieur. Retrouvé 1 août 2013, de http://portal.unesco.org/ci/fr/ev.php-URL_ID=2492&URL_DO=DO_TOPIC&URL_SECTION=201.html
- UNESCO, O. des N. U. pour l'éducation, la science et la culture (Aut). (2012, juin). Déclaration de Paris sur les REL 2012. Retrouvé 1 août 2013, de http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/Events/French_Paris_OER_Declaration.pdf
- UNESCO, O. des N. U. pour l'éducation, la science et la culture (Aut). (s. d.). Ressources éducatives libres. Retrouvé 1 août 2013, de <http://www.unesco.org/new/fr/communication-and-information/access-to-knowledge/open-educational-resources/>
- United Nations, U. or. (Aut). (s. d.). Competencies for the Future. Retrouvé 30 août 2013, de https://careers.un.org/lbw/attachments/competencies_booklet_en.pdf
- Uschold, M., et Gruninger, M. (Aut). (1996). Ontologies: Principles, methods and applications. *Knowledge engineering review*, 11(2), 93–136. Retrouvé de <http://journals.cambridge.org/production/action/cjoGetFulltext?fulltextid=4071856>
- Vassileva, J., McCalla, G., et Greer, J. (Aut). (2003). Multi-Agent Multi-User Modelling. *User Modelling and User-Adapted Interaction*, 28, 179-210. Retrouvé de <http://julita.usask.ca/Texte/UMUAI-draft-7.doc>
- Vaughan-Nichols, S. (Aut). (2012, juin 6). IPv6: When do you really need to switch? | ZDNet. *ZDNet*. Retrouvé 26 juin 2013, de <http://www.zdnet.com/ipv6-when-do-you-really-need-to-switch-3040155336/>
- Viau, R. (Aut). (2005). *La motivation en contexte scolaire*. De Boeck Supérieur.
- Vidal, P., et Broisin, J. (Aut). (2005). Fédération de ressources pédagogiques. *Environnement Informatiques pour l'Apprentissage Humain*, 117-128. Retrouvé de <http://telearn.archives-ouvertes.fr/hal-00005713/>
- Von Bertalanffy, L. (Aut). (1972). The history and status of general systems theory. *Academy of Management Journal*, 15(4), 407–426. Retrouvé de <http://amj.aom.org/content/15/4/407.short>

- Voorhees, R. A. (Aut). (2001). Competency-Based Learning Models: A Necessary Future. *New Directions for Institutional Research*, 2001(110), 5–13. Retrouvé de <http://onlinelibrary.wiley.com/doi/10.1002/ir.7/abstract>
- Vygotski, L. S. (Aut). (1997). *Pensée et langage* (3e (édition originale en russe publiée en 1934)). Paris: La Dispute.
- W3.org (Aut). (s. d.). What is Hypertext? Retrouvé 15 août 2013, de <http://www.w3.org/WhatIs.html>
- W3.org, R. W. G. (Aut). (2004, février 10). RDF - Semantic Web Standards. Retrouvé 18 septembre 2013, de <http://www.w3.org/RDF/>
- W3C Document Object Model. (s. d.). Retrouvé 23 septembre 2013, de <http://www.w3.org/DOM/>
- W3C Semantic Web Activity Homepage. (s. d.). Retrouvé 2 juillet 2013, de <http://www.w3.org/2001/sw/>
- Wahlster, W., et Kobsa, A. (Aut). (1989). User Models in Dialog Systems. Dans A. Kobsa et W. Wahlster (Éd), *User Models in Dialog Systems* (p. 4-34). Springer Berlin Heidelberg. Retrouvé de http://link.springer.com/chapter/10.1007/978-3-642-83230-7_1
- Waterschoot, C. (Aut). (2012, février 1). Du web 1.0 au web 4.0: l'évolution du web depuis 1990. | C-Marketing. Retrouvé 28 juin 2013, de <http://c-marketing.eu/du-web-1-0-au-web-4-0/>
- Watson, J. B. (Aut). (1913). Psychology as the Behaviorist Views it. *Psychological Review*, 20, 158-177. Retrouvé de <http://psychclassics.yorku.ca/Watson/views.htm>
- Weick, K. E. (Aut). (1979). The social psychology of organizing (Topics in social psychology series). Retrouvé de <http://www.citeulike.org/group/5521/article/2875270>
- Weide, K. (Aut). (2010, octobre). Worldwide New Media Market Model 1H12 Highlights: Internet Becomes Ever More Mobile, Ever Less PC Based. Retrouvé 4 juillet 2013, de <http://www.idc.com/getdoc.jsp?containerId=237459>
- Welcome to the Brave New World of MOOCs (Massive Open Online Courses)*. (2013). Retrouvé de http://www.youtube.com/watch?v=KqQNvmQH_YM&feature=youtube_gdata_player
- What is a MOOC?* (2010). Retrouvé de http://www.youtube.com/watch?v=eW3gMGqcZQc&feature=youtube_gdata_player
- Wiener, N. (Aut). (1961). *Cybernetics: or the Control and Communication in the Animal and the Machine*. MIT Press.
- Wiley, D. A. (Aut). (1999). The post-LEGO learning object. Retrouvé de <http://www.citeulike.org/group/2668/article/769279>
- Wiley, D. A. (Aut). (2003). *Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy*.
- Wolton, D. (Aut). (2012). *Indiscipliné : 35 ans de recherches*. Paris: Odile Jacob.
- Woolf, B. P. (Aut). (2010). *Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning*. Morgan Kaufmann. Retrouvé de http://books.google.com/books?hl=en&lr=&id=MnrUj3J_VuEC&oi=fnd&pg=PP2&dq=designing+Intelligent+Interactive+Tutors.&ots=lvy2i1pC4i&sig=MimM-T1bWsvtP3TvzjWsMF0uMJ0
- Xanadu FAQ. (s. d.). Retrouvé 27 juin 2013, de <http://xanadu.com.au/general/faq.html>

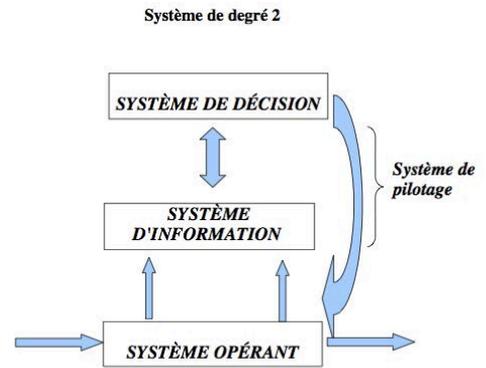
- XHTML 1.0: The Extensible HyperText Markup Language (Second Edition). (s. d.).
Retrouvé 23 septembre 2013, de <http://www.w3.org/TR/xhtml1/>
- XSL Transformations (XSLT). (s. d.). Retrouvé 23 septembre 2013, de
<http://www.w3.org/TR/xslt>
- Yang, S. J., Chen, I. Y.-L., et Shao, N. W. (Aut). (2004). Ontology Enabled Annotation and Knowledge Management for Collaborative Learning in Virtual Learning Community. *Educational Technology & Society*, 7(4), 70–81. Retrouvé de http://www.ifets.info/journals/7_4/ets_7_4.pdf#page=75
- Zacklad, M. (Aut). (2005). Processus de documentarisation dans les Documents pour l'Action (DopA): statut des annotations et technologies de la coopération associées. *Le numérique: Impact sur le cycle de vie du document pour une analyse interdisciplinaire, 13-15 Octobre 2004, Montréal*. Retrouvé de http://archivesic.ccsd.cnrs.fr/sic_00001072/
- Zeleny, M. (Aut). (1987). Management support systems: towards integrated knowledge management. *Human Systems Management*, 1(1), 59-70. Retrouvé de <http://www.milanzeleny.com/documents/publications/mss.pdf>
- Zhang, D. (Aut). (2005). Interactive multimedia-based e-learning: A study of effectiveness. *The American Journal of Distance Education*, 19(3), 149–162. Retrouvé de http://www.tandfonline.com/doi/abs/10.1207/s15389286ajde1903_3
- Zikopoulos, P., Eaton, C., DeRoos, D., Deutsch, T., et Lapis, G. (Aut). (2012). *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data* (1 edition.). The McGraw-Hill Companies, Inc.

Annexes

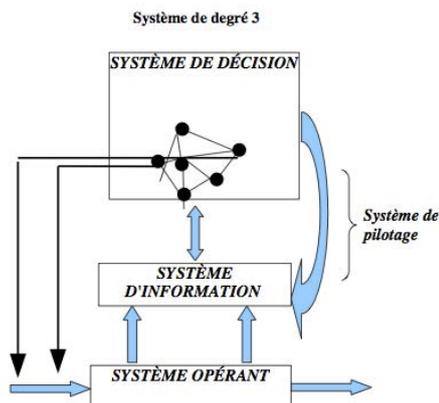
ANNEXE A. Organisation et degrés des systèmes



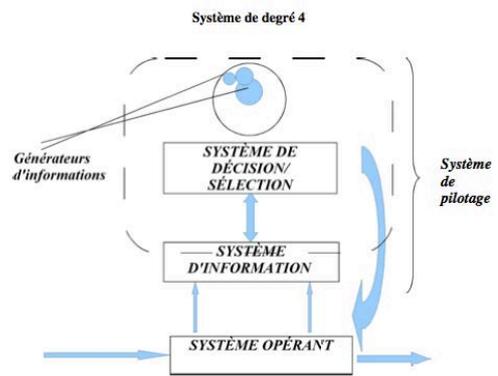
Source : Karsky, Donnadieu, 2002



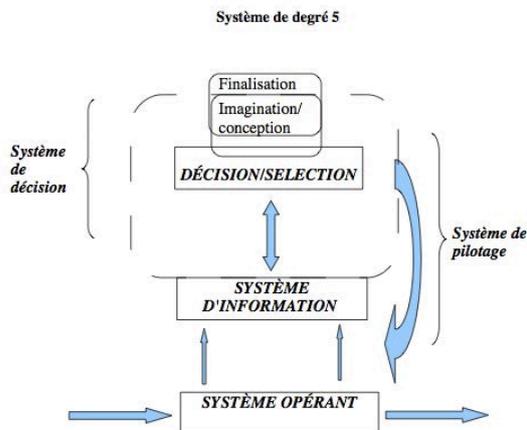
Source : Karsky, Donnadieu, 2002



Source : Karsky, Donnadieu, 2002



Source : Karsky, Donnadieu, 2002



Source : Karsky, Donnadieu, 2002

ANNEXE B. L'ensemble des métadonnées *Dublin Core*

Élément	Élément (anglais)	Commentaire
1. Titre (métadonnée)	Title	Titre principal du document
2. Créateur (métadonnée)	Creator	Nom de la personne, de l'organisation ou du service à l'origine de la rédaction du document
3. Sujet (métadonnée) ou mots clés	Subject	Mots-clefs, phrases de résumé, ou codes de classement
4. Description (métadonnée)	Description	Résumé, table des matières, ou texte libre. Raffinements : table des matières, résumé
5. Éditeur	Publisher	Nom de la personne, de l'organisation ou du service à l'origine de la publication du document
6. Contributeur	Contributor	Nom d'une personne, d'une organisation ou d'un service qui contribue ou a contribué à l'élaboration du document. Chaque contributeur fait l'objet d'un élément Contributor séparé
7. Date (métadonnée)	Date	Date d'un évènement dans le cycle de vie du document
8. Type de ressource	Type	Genre du contenu
9. Format	Format	Type MIME, ou format physique du document
10. Identifiant de la ressource	Identifier	Identificateur non ambigu : il est recommandé d'utiliser un système de référencement précis, afin que l'identifiant soit unique au sein du site, par exemple les URI ou les numéros ISBN. Raffinement : Is Available At
11. Source	Source	Ressource dont dérive le document : le document peut découler en totalité ou en partie de la ressource en question. Il est recommandé d'utiliser une dénomination formelle des ressources, par exemple leur URI
12. Langue (métadonnée)	Language	
13. Relation (métadonnée)	Relation	Lien avec d'autres ressources. De nombreux raffinements permettent d'établir des liens précis, par exemple de version, de chapitres, de standard, etc.
14. Couverture (métadonnée)	Coverage	Couverture spatiale (point géographique, pays, régions, noms de lieux) ou temporelle
15. Droits (métadonnée)	Rights	Droits de propriété intellectuelle, Copyright, droits de propriété divers

ANNEXE C. Correspondance entre les éléments Dublin Core et les éléments du LOM

The Dublin Core defines 15 data elements. These data elements map directly to data elements defined in this Standard, as illustrated in the table below.

Table B.1 –Mapping to unqualified Dublin Core Metadata Element Set

DC.Identifier	1.1.2:General.Identifier.Entry
DC.Title	1.2:General.Title
DC.Language	1.3:General.Language
DC.Description	1.4:General.Description
DC.Subject	1.5:General.Keyword or 9:Classification with 9.1:Classification.Purpose equals "Discipline" or "Idea".
DC.Coverage	1.6:General.Coverage
DC.Type	5.2:Educational.LearningResourceType
DC.Date	2.3.3:LifeCycle.Contribute.Date when 2.3.1:LifeCycle.Contribute.Role has a value of "Publisher".
DC.Creator	2.3.2:LifeCycle.Contribute.Entity when 2.3.1:LifeCycle.Contribute.Role has a value of "Author".
DC.OtherContributor	2.3.2:LifeCycle.Contribute.Entity with the type of contribution specified in 2.3.1:LifeCycle.Contribute.Role.
DC.Publisher	2.3.2:LifeCycle.Contribute.Entity when 2.3.1:LifeCycle.Contribute.Role has a value of "Publisher".
DC.Format	4.1:Technical.Format
DC.Rights	6.3:Rights.Description
DC.Relation	7.2.2:Relation.Resource.Description
DC.Source	7.2:Relation.Resource when the value of 7.1:Relation.Kind is "IsBasedOn".

NOTE 1:--The Dublin Core Metadata Initiative is also developing data element qualifiers to further refine the semantics of the Dublin Core data elements⁹. A further refinement of the mapping in table B.1 can be based on these qualifiers.

NOTE 2:--The LOM working group is committed to working with the Dublin Core Metadata Initiative (DCMI) to develop interoperable metadata, as outlined in the Memorandum of Understanding between the IEEE LTSC LOM WG and the DCMI¹⁰.

Source : (IEEE, 2002, p. 44)

ANNEXE D. Interface de MITx (le MOOC pilot)

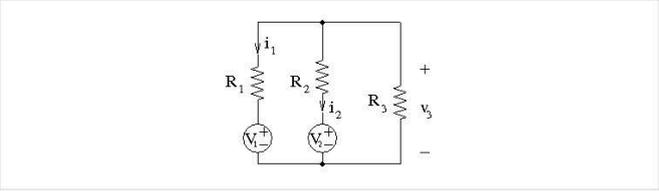
MITx Circuits and Electronics Courseware Course Info Textbook Discussion Wiki Profile

Courseware Index

- Overview
- Week 1
- Week 2**
 - Linearity and Superposition
 - Lecture Sequence
 - Static Discipline and Boolean Logic
 - Lecture Sequence
 - Linearity, Thevenin and Norton; Digital
 - Homework due March 25
 - Superposition experiment
 - Lab due March 25
 - Lab2b: Mixing Two Signals
 - Lab due March 25
 - Week 2 Tutorials
 - Tutorial Index
- Week 3
- Week 4
- Week 5
- Week 6
- Week 7
- Week 8
- Midterm Exam
- Week 9
- Week 10
- Week 11
- Week 12
- Week 13
- Week 14
- Final Exam

S3E3: CIRCUIT VARIABLES ARE SUPERPOSITIONS OF VALUES DUE TO EACH SOURCE SEPARATELY

Consider the familiar network containing two voltage sources and three resistors. You are given that $R_1 = 7.0\Omega$, $R_2 = 3.0\Omega$, $R_3 = 5.0\Omega$, $V_1 = 2.0V$, and $V_2 = 8.0V$.



The voltage v_3 across resistor R_3 can be expressed as the sum of the voltage x_1 due to V_1 acting alone and the voltage x_2 due to V_2 acting alone.

The value (in Volts) of x_1 is: ✓

The value (in Volts) of x_2 is: ✓

Similarly, the current i_1 into the resistor R_1 can be expressed as the sum of the current y_1 due to V_1 acting alone and the current y_2 due to V_2 acting alone.

The value (in Amperes) of y_1 is: ✓

The value (in Amperes) of y_2 is: ✓

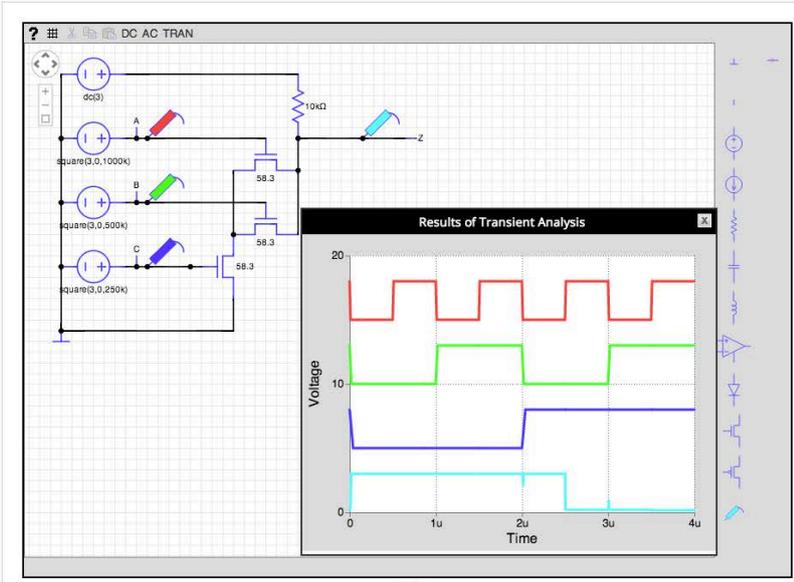
MITx Circuits and Electronics Courseware Course Info Textbook Discussion Wiki Profile

Courseware Index

- Overview**
- Welcome
 - Video
- System Usage Sequence
 - Lecture Sequence
- Lab0: Using the tools
 - Lab
- Circuit Sandbox**
 - Lab
- Week 1
- Week 2
- Week 3
- Week 4
- Week 5
- Week 6
- Week 7
- Week 8
- Midterm Exam
- Week 9
- Week 10
- Week 11
- Week 12
- Week 13
- Week 14
- Final Exam

CIRCUIT SANDBOX

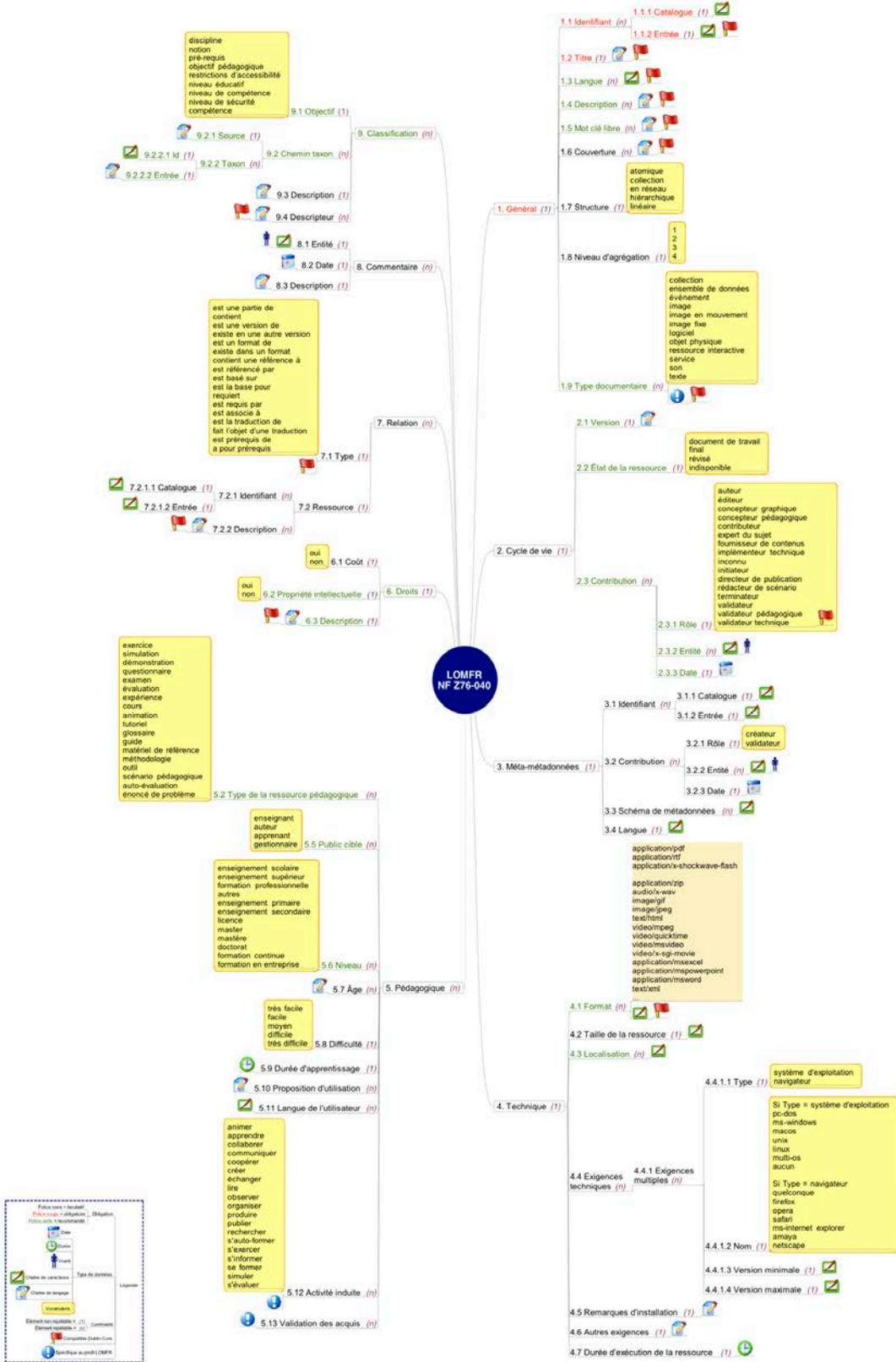
Here's a sandbox where you can experiment with all the components we'll discuss in 6.002x. If you click on CHECK below, your diagram will be saved on the server and you can return at some later time.



Results of Transient Analysis

Check Show Answer

ANNEXE E. Le standard LOM-f



ANNEXE F. Les triplets axiomatiques d’RDF et RDFS

Les triples axiomatiques d’RDF sont les suivants :

```
rdf:type rdf:type rdf:Property .
rdf:subject rdf:type rdf:Property .
rdf:predicate rdf:type rdf:Property .
rdf:object rdf:type rdf:Property .
rdf:first rdf:type rdf:Property .
rdf:rest rdf:type rdf:Property .
rdf:value rdf:type rdf:Property .
rdf:_1 rdf:type rdf:Property .
rdf:_2 rdf:type rdf:Property .
...
rdf:nil rdf:type rdf:List .
```

Les triplets axiomatiques d’RDFS sont les suivants :

```
rdf:type rdfs:domain rdfs:Resource .
rdfs:domain rdfs:domain rdf:Property .
rdfs:range rdfs:domain rdf:Property .
rdfs:subPropertyOf rdfs:domain rdf:Property .
rdfs:subClassOf rdfs:domain rdfs:Class .
rdf:subject rdfs:domain rdf:Statement .
rdf:predicate rdfs:domain rdf:Statement .
rdf:object rdfs:domain rdf:Statement .
rdfs:member rdfs:domain rdfs:Resource .
rdf:first rdfs:domain rdf:List .
rdf:rest rdfs:domain rdf:List .
rdfs:seeAlso rdfs:domain rdfs:Resource .
rdfs:isDefinedBy rdfs:domain rdfs:Resource .
rdfs:comment rdfs:domain rdfs:Resource .
rdfs:label rdfs:domain rdfs:Resource .
rdf:value rdfs:domain rdfs:Resource .

rdf:type rdfs:range rdfs:Class .
rdfs:domain rdfs:range rdfs:Class .
rdfs:range rdfs:range rdfs:Class .
rdfs:subPropertyOf rdfs:range rdf:Property .
rdfs:subClassOf rdfs:range rdfs:Class .
rdf:subject rdfs:range rdfs:Resource .
rdf:predicate rdfs:range rdfs:Resource .
rdf:object rdfs:range rdfs:Resource .
rdfs:member rdfs:range rdfs:Resource .
rdf:first rdfs:range rdfs:Resource .
rdf:rest rdfs:range rdf:List .
rdfs:seeAlso rdfs:range rdfs:Resource .
rdfs:isDefinedBy rdfs:range rdfs:Resource .
rdfs:comment rdfs:range rdfs:Literal .
rdfs:label rdfs:range rdfs:Literal .
rdf:value rdfs:range rdfs:Resource .

rdf:Alt rdfs:subClassOf rdfs:Container .
rdf:Bag rdfs:subClassOf rdfs:Container .
rdf:Seq rdfs:subClassOf rdfs:Container .
rdfs:ContainerMembershipProperty rdfs:subClassOf rdf:Property .

rdfs:isDefinedBy rdfs:subPropertyOf rdfs:seeAlso .

rdf:XMLLiteral rdf:type rdfs:Datatype .
```

```

rdf:XMLLiteral rdfs:subClassOf rdfs:Literal .
rdfs:Datatype rdfs:subClassOf rdfs:Class .

rdf:_1 rdf:type rdfs:ContainerMembershipProperty .
rdf:_1 rdfs:domain rdfs:Resource .
rdf:_1 rdfs:range rdfs:Resource .
rdf:_2 rdf:type rdfs:ContainerMembershipProperty .
rdf:_2 rdfs:domain rdfs:Resource .
rdf:_2 rdfs:range rdfs:Resource .
...

```

Source : <http://www.w3.org/TR/rdf-mt/>

ANNEXE G. Classes de complexité

Dans la théorie de la complexité, les problèmes computationnels sont classifiés dans des *classes de complexité* en fonction des ressources nécessaires à la résolution de ces problèmes (e.g. temps nécessaire donné par le nombre des étapes de l'algorithme de résolution, espace nécessaire en mémoire pour enregistrer les variables de l'algorithme). Nous rappelons les classes de complexité suivantes (source : www.w3.org/TR/owl-profiles/#Computational_Properties).

- N2EXPTIME est la classe des problèmes solubles par un algorithme non-déterministe dans un temps qui est au maximum l'exponentielle double de la taille de l'entrée (e.g., environ 2^{2^n} , n étant la taille de l'entrée) ;
- NEXPTIME est la classe des problèmes solubles par un algorithme non-déterministe dans un temps qui est au maximum l'exponentielle de la taille de l'entrée (e.g. environ 2^n , n pour la taille de l'entrée) ;
- PSPACE est la classe des problèmes solubles par un algorithme déterministe en utilisant un espace qui est au maximum donné par une expression polynomiale en fonction de la taille de l'entrée (e.g. environ $n \cdot c$, n étant la taille de l'entrée et c une constante) ;
- NP est la classe des problèmes solubles par un algorithme non-déterministe utilisant le temps qui est au maximum polynomiale en la taille de l'entrée (e.g. environ n^c , n étant la taille de l'entrée et c une constante) ;
- PTIME est la classe des problèmes solubles par un algorithme déterministe en utilisant le temps qui est au maximum donné par une expression polynomiale de la taille de l'entrée (e.g. environ n^c , n étant la taille de l'entrée et c une constante). PTIME est souvent désigné comme traçable (en anglais *tractable*), alors que les problèmes dans les classes ci-dessus sont souvent désignés comme non-traçable (en anglais *intractable*).
- LOGSPACE est la classe des problèmes solubles par un algorithme déterministe en utilisant l'espace qui est au maximum donné par le logarithmique en fonction de la taille de l'entrée (e.g. environ $c \cdot \log(n)$, n étant la taille de l'entrée et c une constante). NLOGSPACE est la version non-déterministe de cette classe.

AC^0 est une sous-classe de la classe LOGSPACE définie par l'intermédiaire de *circuits* (à la différence des modèles définis par l'intermédiaire de machines Turing, qui permettent de comprendre la complexité des calculs sériels, les circuits,— qui sont des interconnexions fixes entre les portes et qui calculent une seule fonction Booléenne sur un nombre déterminé d'entrés — permettent de modeler les calculs parallèles). AC^0 est la classe des problèmes définissables en utilisant une famille de circuits de profondeur constante et de taille polynômiale, qui peut être générée par une machine Turing déterministe en temps logarithmique (dans la taille de l'entrée). Intuitivement, AC^0 permet d'utiliser un nombre polynômial de processeurs, mais le moment de l'exécution doit être constant.

Source : http://www.w3.org/TR/owl2-profiles/#Computational_Properties

ANNEXE H. La modélisation des utilisateurs dans le cadre du SASA

La modélisation de types d'utilisateurs :

```
:Person rdf:type owl:Class .

:User rdf:type owl:Class ;
      rdfs:subClassOf :Person .

:Learner rdf:type owl:Class ;
         owl:equivalentClass [ rdf:type owl:Class ;
                                owl:intersectionOf ( :User
                                                       [ rdf:type owl:Restriction ;
                                                         owl:onProperty :hasRole ;
                                                         owl:hasValue :Student
                                                       ]
                                                    )
                                ] ;
         rdfs:subClassOf :User .

:Role rdf:type owl:Class ;
      owl:equivalentClass [ rdf:type owl:Class ;
                              owl:oneOf ( :Student
                                             :Professor
                                             :Secretary
                                             :Admin
                                             )
                              ] .

:Admin rdf:type :Role ,
        owl:NamedIndividual .

:Professor rdf:type :Role ,
             owl:NamedIndividual .

:Secretary rdf:type :Role ,
             owl:NamedIndividual .

:Student rdf:type :Role ,
            owl:NamedIndividual .
```

La modélisation des styles d'apprentissage :

```
:FSLSDimension rdf:type owl:Class ;
    owl:equivalentClass [ rdf:type owl:Class ;
        owl:oneOf ( :Participation
            :Format
            :PresentationOrder
            :Perception
        )
    ] ;
    rdfs:subClassOf :LearningStyleDimension .

:Format rdf:type :LearningStyleDimension ,
    owl:NamedIndividual ;
    :hasDirection :Verbal ,
        :Visual .

:Participation rdf:type :LearningStyleDimension ,
    owl:NamedIndividual ;
    :hasDirection :Active ,
        :Reflective .

:Perception rdf:type :LearningStyleDimension ,
    owl:NamedIndividual ;
    :hasDirection :Intuitive ,
        :Sensory .

:PresentationOrder rdf:type :LearningStyleDimension ,
    owl:NamedIndividual ;
    :hasDirection :Global ,
        :Sequential .

:FSLSDimensionDirection rdf:type owl:Class ;
    owl:equivalentClass [ rdf:type owl:Class ;
        owl:oneOf ( :Reflective
            :Sensory
            :Active
            :Global
            :Verbal
            :Visual
            :Sequential
            :Intuitive
        )
    ] ;
    rdfs:subClassOf :LearningStyleDimensionDirection .

:hasDirection rdf:type owl:ObjectProperty ;
    rdfs:domain :LearningStyleDimension ;
    rdfs:range :LearningStyleDimensionDirection .

:hasDirectionValue rdf:type owl:ObjectProperty ;
    rdfs:domain :LearningStyleDimensionDirection ;
    rdfs:range :LearningStyleDimensionDirectionValue .
```

La modélisation des préférences d'apprentissage :

```
:Preference rdf:type owl:Class .
:PreferenceType rdf:type owl:Class ;
    owl:equivalentClass [ rdf:type owl:Class ;
        owl:oneOf ( :Activity
            :Subject
        )
    ] .
:hasMotivation rdf:type owl:DatatypeProperty ;
    rdfs:domain :Preference ;
    rdfs:range rdfs:Literal .
:timeSpent rdf:type owl:DatatypeProperty ;
    rdfs:domain :Preference ;
    rdfs:range rdfs:Literal .
:Activity rdf:type :PreferenceType ,
    owl:NamedIndividual .
:Subject rdf:type :PreferenceType ,
    owl:NamedIndividual .
```

ANNEXE I. Caractéristiques des ressources d'apprentissage

```
:hasPurpose rdf:type owl:ObjectProperty ;
    rdfs:domain :LearningResource ;
    rdfs:range :Purpose ;
    rdfs:subPropertyOf :hasDescription-Classification .
:EducationalPurpose rdf:type owl:Class ;
    rdfs:subClassOf :Purpose .
:Purpose-competence rdf:type :EducationalPurpose ,
    :Purpose ,
    owl:NamedIndividual .
:Purpose-educationalObjective rdf:type :EducationalPurpose ,
    :Purpose ,
    owl:NamedIndividual .
:Purpose-accessibilityRestrictions rdf:type :Purpose ,
    owl:NamedIndividual .
:Purpose-discipline rdf:type :Purpose ,
    owl:NamedIndividual .
:Purpose-educationalLevel rdf:type :Purpose ,
    owl:NamedIndividual .
:Purpose-idea rdf:type :Purpose ,
    owl:NamedIndividual .
:Purpose-prerequisite rdf:type :Purpose ,
    owl:NamedIndividual .
:Purpose-securityLevel rdf:type :Purpose ,
    owl:NamedIndividual .
:Purpose-skillLevel rdf:type :Purpose ,
    owl:NamedIndividual .
```

Les propriétés *pédagogiques* des ressources d'apprentissage :

```
:hasDifficulty rdf:type owl:ObjectProperty ;
  rdfs:range :Difficulty ;
  rdfs:subPropertyOf :hasDescription-Educational .

:hasIntendedContext rdf:type owl:ObjectProperty ;
  rdfs:range :Context ;
  rdfs:subPropertyOf :hasDescription-Educational .

:hasIntendedEndUserRole rdf:type owl:ObjectProperty ;
  rdfs:range :IntendedEndUserRole ;
  rdfs:subPropertyOf :hasDescription-Educational .

:hasIntendedTypicalAgeRange rdf:type owl:ObjectProperty ;
  rdfs:range :TypicalAgeRange ;
  rdfs:subPropertyOf :hasDescription-Educational .

:hasInteractivityLevel rdf:type owl:ObjectProperty ;
  rdfs:range :InteractivityLevel ;
  rdfs:subPropertyOf :hasDescription-Educational .

:hasInteractivityType rdf:type owl:ObjectProperty ;
  rdfs:range :InteractivityType ;
  rdfs:subPropertyOf :hasDescription-Educational .

:hasLearningResourceType rdf:type owl:ObjectProperty ;
  rdfs:range :LearningResourceType ;
  rdfs:subPropertyOf :hasDescription-Educational .

:hasSemanticDensity rdf:type owl:ObjectProperty ;
  rdfs:range :SemanticDensity ;
  rdfs:subPropertyOf :hasDescription-Educational .

:hasTypicalLearningTime rdf:type owl:ObjectProperty ;
  rdfs:range :TypicalLearningTime ;
  rdfs:subPropertyOf :hasDescription-Educational .
```

Valeurs possibles pour la propriété hasInteractivityLevel :

```
:InteractivityLevel-high rdf:type :InteractivityLevel ,
  owl:NamedIndividual .

:InteractivityLevel-low rdf:type :InteractivityLevel ,
  owl:NamedIndividual .

:InteractivityLevel-medium rdf:type :InteractivityLevel ,
  owl:NamedIndividual .

:InteractivityLevel-veryHigh rdf:type :InteractivityLevel ,
  owl:NamedIndividual .

:InteractivityLevel-veryLow rdf:type :InteractivityLevel ,
  owl:NamedIndividual .
```

La représentation des *types* de ressources d'apprentissage :

:LearningResourceType-diagramme rdf:type :LearningResourceType
, owl:NamedIndividual .

:LearningResourceType-exercice rdf:type :LearningResourceType
, owl:NamedIndividual .

:LearningResourceType-narrativeText rdf:type
:LearningResourceType , owl:NamedIndividual .

:LearningResourceType-questionnaire rdf:type
:LearningResourceType , owl:NamedIndividual .

:LearningResourceType-simulation rdf:type
:LearningResourceType , owl:NamedIndividual .

:LearningResourceType-videoLecture rdf:type
:LearningResourceType , owl:NamedIndividual .

:LearningResourceType-graphe rdf:type :LearningResourceType ,
owl:NamedIndividual .

:LearningResourceType-texteNaratif rdf:type
:LearningResourceType , owl:NamedIndividual .

:LearningResourceType-examen rdf:type :LearningResourceType ,
owl:NamedIndividual .

:LearningResourceType-experience rdf:type
:LearningResourceType , owl:NamedIndividual .

:LearningResourceType-problemStatement rdf:type
:LearningResourceType , owl:NamedIndividual .

:LearningResourceType-selfAssessment rdf:type
:LearningResourceType , owl:NamedIndividual .

ANNEXE J. La modélisation des objectifs d'apprentissage

Les principales classes modélisées dans l'ontologie des objectifs d'apprentissage :

```
:GenericLearningObjective rdf:type owl:Class .
:LearningObjective rdf:type owl:Class ;
    rdfs:subClassOf :GenericLearningObjective .
:SpecificLearningObjective rdf:type owl:Class ;
    rdfs:subClassOf :GenericLearningObjective .
:Competence rdf:type owl:Class ;
    rdfs:subClassOf :GenericLearningObjective .

:Level rdf:type owl:Class .
:LevelScale rdf:type owl:Class .
```

Relations possibles entre les objectifs d'apprentissage :

```
:isPrerequisiteFor rdf:type owl:IrreflexiveProperty ,
    owl:ObjectProperty ;
    rdfs:domain :GenericLearningObjective ;
    rdfs:range :GenericLearningObjective .

:hasPrerequisite rdf:type owl:IrreflexiveProperty ,
    owl:ObjectProperty ;
    rdfs:domain :GenericLearningObjective ;
    rdfs:range :GenericLearningObjective ;
    owl:inverseOf :isPrerequisiteFor .

:isPartOf rdf:type owl:IrreflexiveProperty ,
    owl:ObjectProperty ;
    rdfs:domain :LearningObjective ;
    rdfs:range :LearningObjective ;
    owl:inverseOf :hasPart .

:hasPart rdf:type owl:IrreflexiveProperty ,
    owl:ObjectProperty ;
    rdfs:range :LearningObjective ;
    rdfs:domain :LearningObjective .

:isEndorsedBy rdf:type owl:ObjectProperty ;
    rdfs:domain :Competence ;
    rdfs:range :LearningObjective .
```

Représentation des niveaux des objectifs d'apprentissage :

```
:hasCompetenceLevel rdf:type owl:ObjectProperty ;  
  rdfs:domain :LearningObjective ;  
  rdfs:range :LearningObjectiveLevel .
```

```
:hasDefinedLevel rdf:type owl:ObjectProperty ;  
  rdfs:range :Level ;  
  rdfs:domain :LevelScale .
```

```
:hasLevel rdf:type owl:ObjectProperty ;  
  rdfs:domain :LearningObjectiveLevel ;  
  rdfs:range :Level .
```

```
:hasLevelScale rdf:type owl:ObjectProperty ;  
  rdfs:domain :LearningObjectiveLevel ;  
  rdfs:range :LevelScale .
```

Les types d'un objectif spécifique d'apprentissage :

```
:LearningObjectiveType rdf:type owl:Class ;  
  owl:equivalentClass [ rdf:type owl:Class ;  
    owl:oneOf ( :LearningObjectiveType-cognitive  
      :LearningObjectiveType-psychomotor  
      :LearningObjectiveType-attitude)  
  ] .
```

```
:hasLearningObjectiveType rdf:type owl:ObjectProperty ;  
  rdfs:range :LearningObjectiveType ;  
  rdfs:domain :SpecificLearningObjective .
```

ANNEXE K. Extraits de l'ontologie de l'objet d'évaluation

```
:EvaluationObject rdf:type owl:Class ;
    rdfs:subClassOf [ rdf:type owl:Restriction ;
        owl:onProperty :hasType ;
        owl:someValuesFrom :EvaluationObjectType
    ] .

:EvaluationObjectType rdf:type owl:Class .
:EvaluationObjectType-problem rdf:type
    :EvaluationObjectType , owl:NamedIndividual .
:EvaluationObjectType-question rdf:type
    :EvaluationObjectType , owl:NamedIndividual .
:EvaluationObjectType-questionnaire rdf:type
    :EvaluationObjectType , owl:NamedIndividual .

:EvaluationType rdf:type owl:Class .
:EvaluationType-exam rdf:type :EvaluationType ,
    owl:NamedIndividual .
:EvaluationType-exercice rdf:type :EvaluationType ,
    owl:NamedIndividual .
:EvaluationType-selfAssessment rdf:type :EvaluationType ,
    owl:NamedIndividual .

:Solution rdf:type owl:Class .
:CorrectSolution rdf:type owl:Class ;
    rdfs:subClassOf :Solution ;
    owl:disjointWith :IncorrectSolution .
:IncorrectSolution rdf:type owl:Class ;
    rdfs:subClassOf :Solution .

:Statement rdf:type owl:Class .
```

Propriétés dans l'ontologie de l'objet d'évaluation :

```
:hasType rdf:type owl:ObjectProperty ;
    rdfs:domain :EvaluationObject ;
    rdfs:range :EvaluationObjectType .

:hasEvaluationType rdf:type owl:ObjectProperty ;
    rdfs:domain :EvaluationObject ;
    rdfs:range :EvaluationType .

:hasRequiredPassingLevel rdf:type owl:ObjectProperty ;
    rdfs:domain :EvaluationObject ;
    rdfs:range :EvaluationObjectTypeCharacteristics .

:hasStatement rdf:type owl:ObjectProperty ;
    rdfs:domain :EvaluationObject ;
    rdfs:range :Statement .

:hasSolution rdf:type owl:ObjectProperty ;
    rdfs:domain :EvaluationObject ;
    rdfs:range :Solution .
```

ANNEXE L. Extraits de l'ontologie d'annotation

```
:AnnotationObject rdf:type owl:Class ;
    rdfs:subClassOf owl:Thing .

:AnnotationObjectType rdf:type owl:Class ;
    rdfs:subClassOf owl:Thing .

:AnnotationObjectType-Appreciation    rdf:type
    :AnnotationObjectType , owl:NamedIndividual .
:AnnotationObjectType-Note           rdf:type
    :AnnotationObjectType , owl:NamedIndividual .
:AnnotationObjectType-Tag             rdf:type
    :AnnotationObjectType , owl:NamedIndividual .

:Appreciation rdf:type owl:Class ;
    owl:equivalentClass [ rdf:type owl:Class ;
        owl:intersectionOf ( :AnnotationObject
            [ rdf:type owl:Restriction ;
                owl:onProperty :hasType ;
                owl:hasValue :AnnotationObjectType-Appreciation
            ])
        ] ;
    rdfs:subClassOf :AnnotationObject .

:Note rdf:type owl:Class ;
    owl:equivalentClass [ rdf:type owl:Class ;
        owl:intersectionOf ( :AnnotationObject
            [ rdf:type owl:Restriction ;
                owl:onProperty :hasType ;
                owl:hasValue :AnnotationObjectType-Note
            ])
        ] ;
    rdfs:subClassOf :AnnotationObject .

:Tag rdf:type owl:Class ;
    owl:equivalentClass [ rdf:type owl:Class ;
        owl:intersectionOf ( :AnnotationObject
            [ rdf:type owl:Restriction ;
                owl:onProperty :hasType ;
                owl:hasValue :AnnotationObjectType-Tag
            ])
        ] ;
    rdfs:subClassOf :AnnotationObject .
```

Critères d'appréciation d'une ressource d'apprentissage

```
:AppreciationCriterion rdf:type owl:Class .

:AppreciationCriterion-General    rdf:type
    :AppreciationCriterion , owl:NamedIndividual .
:AppreciationCriterion-Content    rdf:type
    :AppreciationCriterion , owl:NamedIndividual .
:AppreciationCriterion-Presentation    rdf:type
    :AppreciationCriterion , owl:NamedIndividual .
:AppreciationScale rdf:type owl:Class ;
    rdfs:subClassOf owl:Thing .

:hasAppreciationCriterion rdf:type owl:ObjectProperty ;
    rdfs:domain :Appreciation ;
    rdfs:range :AppreciationCriterion .
:hasAppreciationScale rdf:type owl:ObjectProperty ;
    rdfs:domain :Appreciation ;
    rdfs:range :AppreciationScale .
:hasType rdf:type owl:ObjectProperty ;
    rdfs:domain :AnnotationObject ;
    rdfs:range :AnnotationObjectType .
```

ANNEXE M. Extraits de l'ontologie de cadre d'enseignement

```
:CNUDomain rdf:type owl:Class .
:CNUGroup rdf:type owl:Class .
:CNUSection rdf:type owl:Class .
:Formation rdf:type owl:Class .
:Specialisation rdf:type owl:Class .
:Mention rdf:type owl:Class .
:Course rdf:type owl:Class ;
    owl:equivalentClass :LearningUnit ;
    rdfs:subClassOf :LearningPath .
:CertificationType rdf:type owl:Class .
:LearningPath rdf:type owl:Class .
:LearningType rdf:type owl:Class .
:LearningUnit rdf:type owl:Class .
:LearningUnitElement rdf:type owl:Class .
:Professor rdf:type owl:Class .
:LearningResource rdf:type owl:Class .
:hasMention rdf:type owl:ObjectProperty ;
    rdfs:range :Mention ;
    rdfs:domain :Specialisation .
:hasSpecialisation rdf:type owl:ObjectProperty ;
    rdfs:domain :Formation ;
    rdfs:range :Specialisation .
:hasFormation rdf:type owl:ObjectProperty ;
    rdfs:range :Formation ;
    rdfs:domain :LearningUnit .
:hasCertificationType rdf:type owl:ObjectProperty ;
    rdfs:range :CertificationType ;
    rdfs:domain :LearningUnit .
:taughtBy rdf:type owl:ObjectProperty ;
    rdfs:domain :LearningUnit ;
    rdfs:range :Professor .
```

ANNEXE N. Relations entre les entités déclarées dans les ontologies

Relations entre le parcours d'apprentissage et la ressource d'apprentissage :

```
:containLearningResource rdf:type owl:ObjectProperty ;
    rdfs:domain :LearningPath ;
    rdfs:range :LearningResource .
:hasNext rdf:type owl:FunctionalProperty ,
    owl:ObjectProperty ;
    rdfs:subPropertyOf :isFollowedBy .
:isFollowedBy rdf:type owl:ObjectProperty ,
    owl:TransitiveProperty ;
    rdfs:range :OWLListElement ;
    rdfs:subPropertyOf :hasListProperty .
```

Relations entre les objets pédagogiques, les objectifs d'apprentissage et l'apprenant :

```
res:learningObject1 lo:hasLearningObjective res:objectif1 .
res:learningObject2 lo:hasLearningObjective res:objectif4 .
res:learningObject3 lo:hasLearningObjective res:objectif5 .
res:learningObject4 lo:hasLearningObjective res:objectif5 .
res:learningObject4 lo:hasLearningObjective res:objectif6 .
res:learningObject5 lo:hasLearningObjective res:objectif7 .
res:learningObject6 lo:hasLearningObjective res:objectif8 .

res:apprenant lo:hasAttended
[
    lo:learningResource res:learningObject1 ;
    lo:startedTime      "2012-10-20T10:00:00Z"^^xsd:dateTime ;
    lo:completedTime   "2012-10-20T10:45:00Z"^^xsd:dateTime ]
,
[
    lo:learningResource res:learningObject5 ;
    lo:startedTime      "2012-10-30T11:00:00Z"^^xsd:dateTime ;
    lo:completedTime   "2012-10-30T12:00:00Z"^^xsd:dateTime ]
.
```

La représentation des relations entre l'objectif d'apprentissage, l'objet d'évaluation et l'apprenant :

```
res:objectif5 eo:isValidatedByEvaluationObject res:evalObj1 .
res:evalObj1 eo:hasEvaluationPointsMax 10 ;
             eo:hasMinValidationPoints 7 .

res:objectif6 eo:isValidatedByEvaluationObject res:evalObj2 .
res:evalObj2 eo:hasEvaluationPointsMax 20 ;
             eo:hasMinValidationPoints 10 .

res:objectif2 eo:isValidatedByEvaluationObject res:evalObj4 .
res:evalObj4 eo:hasEvaluationPointsMax 20 ;
             eo:hasMinValidationPoints 10 .
```

```
res:apprenant eo: hasAttended [
  eo:attendedEvaluationObject res:evalObj1 ;
  eo:startedTime "2013-10-15T10:00:00Z"^^xsd:dateTime ;
  eo:completedTime "2013-10-15T12:00:00Z"^^xsd:dateTime ;
  eo:obtainedPoints 4
] .
```

```
res:apprenant eo: hasAttended [
  eo:attendedEvaluationObject res:evalObj2 ;
  eo:startedTime "2011-10-15T09:00:00Z"^^xsd:dateTime ;
  eo:completedTime "2011-10-15T12:00:00Z"^^xsd:dateTime ;
  eo:obtainedPoints 9
] .
```

```
res:apprenant eo:hasAttended [
  eo:attendedEvaluationObject res:evalObj4 ;
  eo:startedTime "2012-10-15T10:00:00Z"^^xsd:dateTime ;
  eo:completedTime "2012-10-15T12:00:00Z"^^xsd:dateTime ;
  eo:obtainedPoints 16
] .
```

Représentation du style d'apprentissage dans le SASA :

```
res:apprenant up:hasLearningStyleDimensionValue [
  up:hasLearningStyle up:FelderSilvermanLearningStyle ;
  up:hasDimension up:ParticipationFSLs ;
  up:hasDimensionValue -11 ;
  up:atTime "2012-10-15T12:00:00Z"^^xsd:dateTime ] .
```

```
res:apprenant up:hasLearningStyleDimensionValue [
  up:hasLearningStyle up:FelderSilvermanLearningStyle ;
  up:hasDimension up:PerceptionFSLs ;
  up:hasDimensionValue 4 ;
  up:atTime "2012-10-15T12:00:00Z"^^xsd:dateTime ] .
```

```
res:apprenant up:hasLearningStyleDimensionValue [
  up:hasLearningStyle up:FelderSilvermanLearningStyle ;
  up:hasDimension up:FormatFSLs ;
  up:hasDimensionValue -9 ;
  up:atTime "2012-10-15T12:00:00Z"^^xsd:dateTime ] .
```

```
res:apprenant up:hasLearningStyleDimensionValue [
  up:hasLearningStyle up:FelderSilvermanLearningStyle ;
  up:hasDimension up:PresentationOrderFSLs ;
  up:hasDimensionValue 8 ;
  up:atTime "2012-10-15T12:00:00Z"^^xsd:dateTime ] .
```

Représentation de l'adéquation des ressources d'apprentissage sur les différentes dimensions dans le SASA :

```
res:learningObject1
lo:hasSuitabilityForLearningStyleDimensionValue [
    lo:suitabilityDimension up:ParticipationFSLs ;
    lo:suitabilityValue -11 ] .
```

```
res:learningObject1
lo:hasSuitabilityForLearningStyleDimensionValue [
    lo:suitabilityDimension up:PerceptionFSLs ;
    lo:suitabilityValue 0 ] .
```

```
res:learningObject1
lo:hasSuitabilityForLearningStyleDimensionValue [
    lo:suitabilityDimension up:FormatFSLs ;
    lo:suitabilityValue -10 ] .
```

```
res:learningObject1
lo:hasSuitabilityForLearningStyleDimensionValue [
    lo:suitabilityDimension up:PresentationOrderFSLs ;
    lo:suitabilityValue 8 ] .
```

Association des tags aux ressources d'apprentissage :

```
res:evalObj3      ao:tagged
[    ao:withTag "Semantic Web" , "Social Semantic Web" ;
    ao:byAgent res:enseignant ] .
```

```
res:learningObject8 ao:tagged
[    ao:withTag "Semantic" , "Semantic Web", "Web 3.0" ;
    ao:byAgent res:apprenant ] .
```

ANNEXE O. Exemples des requêtes SPARQL utilisées en SASA

La requête suivante nous permet d'obtenir la liste des objets pédagogiques correspondants aux objectifs d'apprentissage dont l'apprenant peut accéder :

```
SELECT DISTINCT ?objectif ?learning_object
WHERE {
  res:apprenant co:hasAcquired ?pre .
  ?pre co:isPrerequisiteFor ?objectif .
  ?learning_object lo:hasLearningObjective ?objectif .

  FILTER NOT EXISTS { res:apprenant co:hasAttended ?objectif . }
}
```

Obtention de la liste des objets d'évaluation dont l'apprenant n'a pas validé :

```
SELECT ?evalObj
WHERE {
  res:apprenant eo:hasAttended ?bObj .
  ?bObj eo:attendedEvaluationObject ?evalObj .
  ?bObj eo:obtainedPoints ?obtainedPoints .

  ?evalObj eo:hasMinValidationPoints ?minValPoints .
  FILTER (?obtainedPoints < ?minValPoints)
}
```

Obtention de la liste des objets des objets pédagogiques associées aux objets d'évaluation non-validés par l'apprenant :

```
SELECT DISTINCT ?objective ?learningObject
WHERE {
  res:apprenant eo:hasAttended ?bObj .
  ?bObj eo:attendedEvaluationObject ?evalObj .
  ?bObj eo:obtainedPoints ?obtainedPoints .

  ?evalObj eo:hasMinValidationPoints ?minValPoints .

  ?objective eo:isValidatedByEvaluationObject ?evalObj .
  ?learningObject lo:hasLearningObjective ?objective .

  FILTER (?obtainedPoints < ?minValPoints)
}
```

Obtention de la liste des objets des objets pédagogiques correspondants aux tags communs avec les objets d'évaluation :

```

SELECT ?learningObject
WHERE {
res:apprenant eo:hasAttended ?bObj .
?bObj eo:attendedEvaluationObject ?evalObj .
?bObj eo:obtainedPoints ?obtainedPoints .
?evalObj eo:hasMinValidationPoints ?minValPoints .

?evalObj          ao:tagged/ao:withTag      ?tag .
?learningObject   rdf:type                  lo:LearningObject ;
                  ao:tagged/ao:withTag      ?tag .

FILTER (?obtainedPoints < ?minValPoints)
}

```

ANNEXE P. Exemple du code d'interrogation du service DBPedia Lookup

```

private static final byte NUMBER_OF_TOPICS_RETURNED = 7;
private static final String DBPEDIA_LOOKUP_SEARCH =
"http://lookup.dbpedia.org/api/search.asmx/KeywordSearch";

    public String searchTerm(String queryTerms, String dbpediaClass)
throws IOException {
    String results = "";
    HttpClient httpClient = new DefaultHttpClient();
    List<NameValuePair> params = new ArrayList<NameValuePair>();
    params.add(new BasicNameValuePair("QueryString", queryTerms));
    params.add(new BasicNameValuePair("MaxHits",
Byte.toString(NUMBER_OF_TOPICS_RETURNED)));
    params.add(new BasicNameValuePair("QueryClass", dbpediaClass));
    String url = DBPEDIA_LOOKUP_SEARCH + "?" +
URLEncodedUtils.format(params, "UTF-8");
    HttpGet httpGet = new HttpGet(url);
    httpGet.setHeader("Accept", "application/json");
    HttpResponse httpResponse = httpClient.execute(httpGet);
    HttpEntity entity = httpResponse.getEntity();
    if (entity != null) {
        InputStream instream = entity.getContent();
        try {
            ObjectMapper mapper = new ObjectMapper();
            JsonNode rootNode = mapper.readValue(instream,
JsonNode.class);
            results = rootNode.get("results").toString();
        } catch (Exception e) {
            results = e.getMessage();
        } finally {
            instream.close();
        }
    }
    return results.toString();
}
}

```

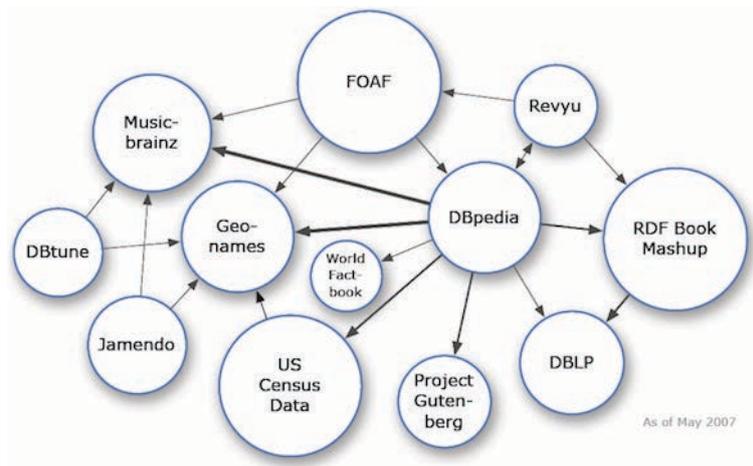
ANNEXE Q. Séquence du fichier de transformation XSLT des métadonnées XML de Youtube en RDF

```
<xsl:template match="/">
  <xsl:element name="rdf:RDF">
    <xsl:apply-templates select="atom:feed"/>
  </xsl:element>
</xsl:template>

<xsl:template match="atom:feed">
  <xsl:for-each select="atom:entry">
    <xsl:element name="rdf:Description">
      <xsl:attribute name="rdf:about">
        <xsl:value-of select="concat('http://semlearn.pu-
pm.univ-fcomte.fr/resource/youtube_id/',atom:id)"/>
      </xsl:attribute>
      <xsl:apply-templates select="atom:link"/>
      <xsl:apply-templates select="atom:category"/>
    </xsl:element>
  </xsl:for-each>
</xsl:template>
```

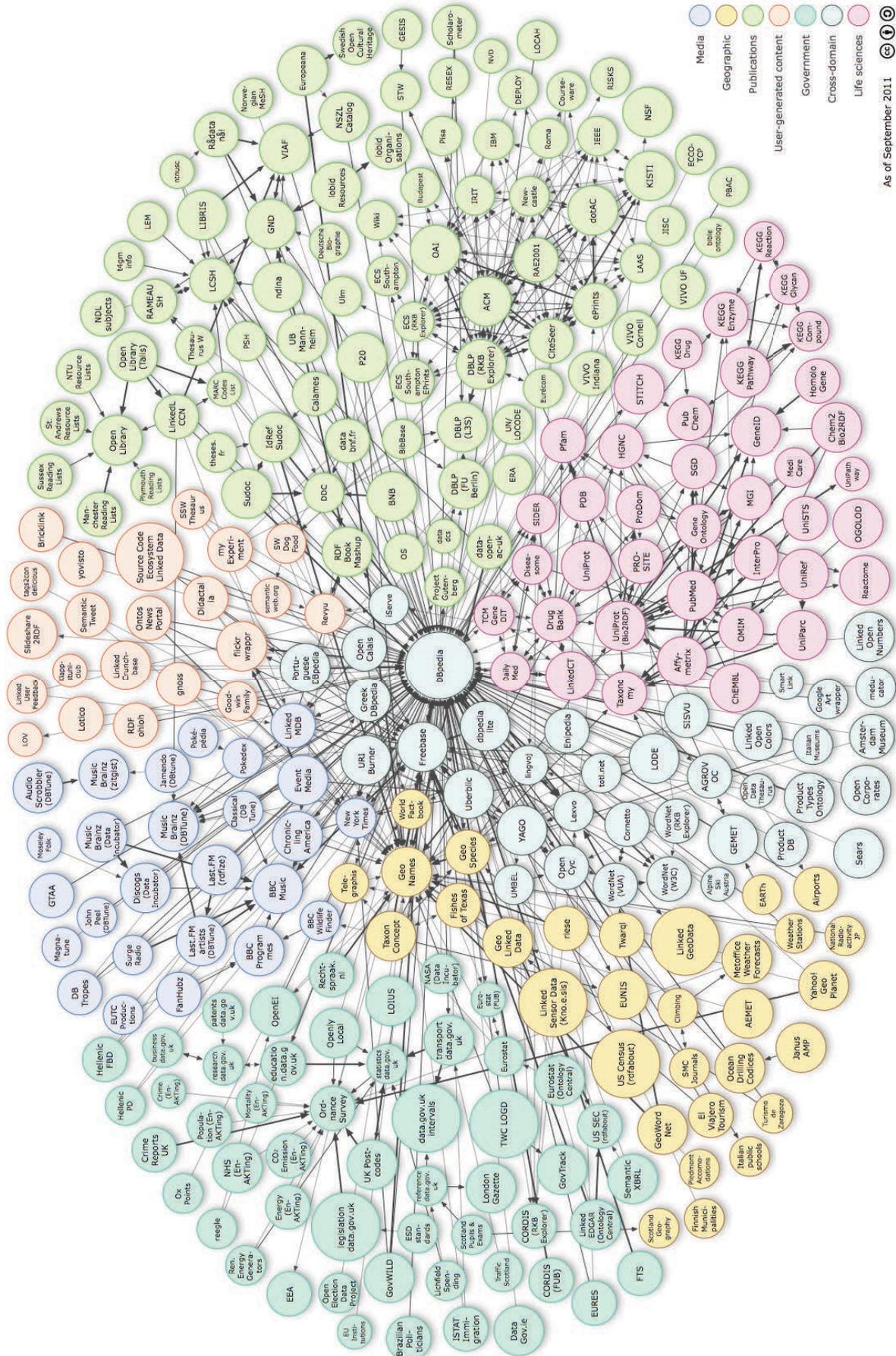
ANNEXE R. Exemples d'applications dans le cadre du Web sémantique : données liées et agents intelligents

Si en 2007 les *données ouvertes et liées* contenaient 12 ensembles de données (*datasets*)



Le diagramme des données ouvertes et liées en 2007 (source : <http://lod-cloud.net/>)

En septembre 2011 le nombre d'ensembles de données est passé à 295, regroupant des données provenant de 7 domaines différents.



La diagramme des données ouvertes et liées en 2011 (source : <http://lod-cloud.net/>)

Ces données sont accessibles sur le portail <http://datahub.io>, qui est construit sur la plateforme *CKAN*³⁰⁹ mise à disposition par la *Fondation pour la connaissance ouverte*³¹⁰.

L'interconnexion des ensembles de données est effectuée grâce aux *liens typés*³¹¹ utilisés dans l'approche *Linked Open Data*. L'existence d'une flèche indique qu'il existe au moins 50 liens entre les ensembles de données reliées. La signification des liens est spécifiée dans le vocabulaire qui le définit. Les vocabulaires les plus utilisés dans LOD sont notés dans le tableau suivant :

Les vocabulaires les plus utilisés dans le *Linked Open Data* (source : « State of the LOD Cloud », s. d.)

Le préfixe du vocabulaire	Le lien du vocabulaire	Utilisation dans les ensembles de données
dc	http://purl.org/dc/elements/1.1/	92 (31,19 %)
foaf	http://xmlns.com/foaf/0.1/	81 (27,46 %)
skos	http://www.w3.org/2004/02/skos/core#	58 (19,66 %)
geo	http://www.w3.org/2003/01/geo/wgs84_pos#	25 (8,47 %)
xhtml	http://www.w3.org/1999/xhtml/vocab#	19 (6,44 %)
akt	http://www.aktors.org/ontology/portal#	17 (5,76 %)
bibo	http://purl.org/ontology/bibo/	14 (4,75 %)
mo	http://purl.org/ontology/mo/	13 (4,41 %)
vcard	http://www.w3.org/2006/vcard/ns#	10 (3,39 %)
sioc	http://rdfs.org/sioc/ns#	10 (3,39 %)

Les tailles des cercles dans la indiquent le volume de données³¹² existant dans l'ensemble de données concerné. Au centre du diagramme, *DBpedia* est constituée par des informations extraites de *Wikipedia* et rendues structurées en utilisant le langage RDF. Les nombre d'instances (e.g. personne, lieu, organisation, œuvre) dans la version

³⁰⁹ CKAN (Comprehensive Knowledge Archive Network - <http://ckan.org/>) est une application web permettant le stockage et la distribution de données ; elle est utilisée par le portail de gouvernance ouverte des institutions nationales, européennes et régionales <http://ckan.org/about>.

³¹⁰ Open Knowledge Foundation a développé la plateforme *CKAN* pour le registre public *TheDatahub.org* afin d'enregistrer les données ouvertes fournies par les différentes institutions (e.g. data.gov.uk, publicdata.eu).

³¹¹ Différents d'un lien générique de type *anchor* utilisé en *html*, les liens employés dans LOD utilisent effectivement le *prédicat* d'un *triplet RDF*. Dans ce cas, le sujet se trouve dans l'ensemble de données source et l'objet se trouve dans l'ensemble de données destination.

³¹² Dans ce cas il s'agit de *triplets RDF* (v. § 3.3.2)..

anglaise du DBpedia est de 3 769 926, pendant que la version française compte 1 197 334 d'instances (DBpedia.org, s. d.).

Les *données ouvertes et liées* utilisent les technologies clés du Web sémantique : les URI pour l'identification des entités ou des concepts, le HTTP pour l'extraction des ressources ou de leur description et le RDF pour structurer et relier les données.

Parmi les agents intelligents les plus connus, nous notons *Deep Blue* et *Watson*. Développées par IBM, ils parviennent à dépasser l'homme dans des tâches appropriées à l'intelligence humaine. Ainsi, le 11 mai 1997, *Deep Blue*, ordinateur spécialisé dans le jeu d'échecs, prend sa revanche et affronte G. Kasparov, le champion mondial d'échecs. En six parties d'échec, on décompte deux victoires de *Deep Blue*, une défaite et trois parties nulles (« IBM100 - Deep Blue », 2012).

Si l'environnement de *Deep Blue* est limité à l'échiquier, IBM monte en compétences avec *Watson* qui, en février 2011, s'attaque au jeu télévisé de culture générale *Jeopardy*. Cette fois, *Watson* peut « comprendre » les questions en langage naturel, buzzer pour prendre la main, trouver les réponses (sans être connecté à Internet), énoncer les réponses à l'aide d'un synthétiseur vocal et choisir le thème et le montant de la prochaine question (*Jeopardy - Watson vs. The Humans - Day 1*, 2011). La « mémoire » de *Watson* est une base de données intégrant des contenus structurés (DBpedia, WordNet, Yago) et non structurés (e.g. le texte complet de Wikipédia, des encyclopédies, dictionnaires, articles) (Heaton, 2011). *Watson* joue contre deux champions de *Jeopardy* et remporte la victoire. Les avancées et les conquêtes des agents intelligents dans l'espace des capacités humaines nous rapprochent du fameux *test de Turing*³¹³ et du *point de singularité*³¹⁴.

³¹³ Proposé par Alain Turing (Turing, 1950), le test qui porte son nom est censé tester les capacités de la machine via des discussions portées en langage naturel avec plusieurs agents humains et artificiels. Si, suite aux réponses obtenues, l'enquêteur humain ne peut pas distinguer l'identité de l'agent artificiel, le test est réussi.

³¹⁴ Dans son livre « The singularity is near », Ray Kurzweil prévoit le passage du test de Turing avant 2029 (Kurzweil, 2005, p. 200), et l'atteinte du point de singularité représentant « une transformation profonde et perturbante de la capacité humaine » en 2045 (ibid., p.136).

À l'heure actuelle, les agents intelligents font leur chemin vers une disponibilité globale. Sur le Web, nous trouvons le *moteur computationnel Wolfram Alpha*³¹⁵ dont l'objectif est de rendre toute connaissance systématique immédiatement calculable et accessible à tous. Ils parviennent à *recueillir et à organiser toutes les données objectives*, implémenter chaque modèle, méthode ou algorithme connu, afin de rendre possible le calcul de tout ce qui peut être calculé sur n'importe quel sujet. L'objectif est de réaliser un *moteur de connaissances* capable d'accepter des interrogations exprimées dans une *forme naturelle*, de *générer des résultats puissants* et de les présenter avec un maximum de clarté (« About Wolfram|Alpha: Making the World's Knowledge Computable », s. d.). Le résultat présenté par Wolfram Alpha permet en outre à l'utilisateur de connaître l'interprétation de la requête effectuée par le moteur.

Plus personnels et « conscients » du contexte de l'utilisateur (espace et temps), sont les agents intelligents intégrés dans les téléphones portables intelligents (smartphones). Parmi les plus connus et utilisés, rappelons *Siri* et *Google Now*. *Siri*, une application basée sur les technologies sémantiques développée au départ par T. Gruber et son équipe, sert à aider ses utilisateurs dans leurs tâches quotidiennes. Il s'agit d'un assistant personnel intelligent pour iOS (présentée le 4 octobre 2011). Possédant les données personnelles de l'utilisateur, *Siri* fournit des réponses adaptées et personnalisées à ses utilisateurs. *Google Now*, intégré dans les dispositifs basés sur le système d'exploitation Android, se situe dans cette même catégorie des agents intelligents. L'exploitation des informations personnelles (e.g. position géographique, liste des contacts) stockées dans le téléphone portable permet ensuite aux agents intelligents d'offrir une variété de services d'adaptation (e.g. liste de points d'intérêt proche géographiquement, calcul des trajets en temps réel).

³¹⁵ <http://www.wolframalpha.com/>

ANNEXE S. Les termes déclarés dans le vocabulaire RDF(S)

Nous énumérons ici les références IRI déclarées dans le *vocabulaire RDF* (V_{RDF}) sans évoquer leur signification ou interprétation. V_{RDF} contient les références suivantes : `rdf:type`, `rdf:Property`, `rdf:XMLLiteral`, `rdf:nil`, `rdf:List`, `rdf:Statement`, `rdf:subject`, `rdf:predicate`, `rdf:object`, `rdf:first`, `rdf:rest`, `rdf:Seq`, `rdf:Bag`, `rdf:Alt`, `rdf:_1`, `rdf:_2` ... `rdf:value` (*Brickley et Guha, 2004; Hayes et McBride, 2004*). Certaines de ces références (e.g. `rdf:type`, `rdf:Property`) sont reprises et explicitées dans la section suivante.

Nous énumérons ici les références IRI appartenant au vocabulaire de RDFS (V_{RDFS}) sans insister sur leur signification et interprétation : `rdfs:domain`, `rdfs:range`, `rdfs:Resource`, `rdfs:Literal`, `rdfs:Datatype`, `rdfs:Class`, `rdfs:subClassOf`, `rdfs:subPropertyOf`, `rdfs:member`, `rdfs:Container`, `rdfs:ContainerMembershipProperty`, `rdfs:comment`, `rdfs:seeAlso`, `rdfs:isDefinedBy`, `rdfs:label` (*ibid.*). À travers ce vocabulaire, l'objectif de RDFS est d'offrir des outils de base pour la construction de son propre vocabulaire composé de classes et de propriétés ainsi que de leur structure hiérarchique.

Liste des figures

Figure 1-1. Ancrage historique et développement des courants de pensée de l'apprentissage	20
Figure 1-2. L'apprenant dans son environnement d'apprentissage	23
Figure 1-3. Fonctions de l'environnement d'apprentissage	25
Figure 1-4. Le triangle pédagogique de J. Houssaye	26
Figure 1-5. Image d'ensemble sur l'apprentissage actif	32
Figure 1-6. Démarche générale d'apprentissage	33
Figure 1-7. Caractéristiques individuelles de l'apprenant	37
Figure 1-8. Parcours d'apprentissage personnalisé	39
Figure 1-9. Vue générale des logiciels éducatifs	42
Figure 1-10. Hybridation des modes d'apprentissage	44
Figure 1-11. Hiérarchie <i>donnée-information-connaissance-compréhension-sagesse</i> d'Ackoff.....	49
Figure 1-12. Hiérarchie DICS.....	50
Figure 1-13. Donnée, information et connaissance.....	51
Figure 1-14. Relations entre les types de connaissances	53
Figure 1-15. La trahison des images, de René Magritte.....	58
Figure 1-16. Terminologie de l'objet pédagogique, adaptée selon	60
Figure 1-17. Niveau de granularité – réutilisabilité	62
Figure 1-18. La pyramide des contenus d'apprentissage	63
Figure 1-19. La structure générale d'un objet pédagogique.....	65
Figure 2-1. Bases de l'ingénierie pédagogique	68
Figure 2-2. Système - ensemble d'éléments en interaction	70
Figure 2-3. Système – boîte noire : fonction de transformation	71
Figure 2-4. Système fermé.....	72
Figure 2-5. La boucle de rétroaction d'un système.....	72
Figure 2-6. Typologie des systèmes.....	78
Figure 2-7. Système canonique.....	79
Figure 2-8. Système de degré quatre	80
Figure 2-9. Schéma d'un système général de communication.....	87
Figure 2-10. Schéma d'un système général de communication avec les apports de Weaver.....	87
Figure 2-11. Le modèle OSI – l'Internet et le Web.....	90
Figure 2-13. La chronologies des MOOCs.....	107
Figure 2-14. Un modèle systémique de l'apprentissage.....	112
Figure 3-1. L'organisation de la documentation selon P. Otlet.....	118
Figure 3-2. L'accès à l'information selon P. Otlet.....	118
Figure 3-3. Le système d'hypertexte proposé par T. Berners-Lee	123
Figure 3-4. Les technologies du Web	124
Figure 3-5. La participation dans le Web 1.0 et le Web 2.0	128
Figure 3-6. Tag Cloud: les mots clés du Web 2.0	130
Figure 3-7. La disjonction des réseaux sociaux	131
Figure 3-8. Caractérisation de la Big Data selon IBM.....	132

Figure 3-9. Ce que "voit" l'ordinateur avec HTML (a), dans le Web sémantique avec les métadonnées (b) vs ce que voit l'homme (c).....	133
Figure 3-10. L'évolution des plateformes dans le cadre de l'évolution du Web	134
Figure 3-11. La sémantique dans l'évolution du Web	135
Figure 3-12. Le triangle sémiotique.....	137
Figure 3-13. La transmission de la sémantique dans la communication (humaine et technique)	138
Figure 3-14. Schéma de la feuille de route du Web sémantique	141
Figure 3-15. L'évolution du Web	142
Figure 3-16. L'architecture en couche des standards du Web sémantique	145
Figure 3-17. Relation entre l'URI, l'URL et l'URN.....	148
Figure 3-18. La syntaxe générique de l'URI	148
Figure 3-19. Exemples d'URI	149
Figure 3-20. Triangle sémiotique dans l'espace d'information du Web.....	149
Figure 3-21. Négociation du contenu RDF à travers le HTTP.....	151
Figure 3-22. Un triplet RDF	154
Figure 3-23. Exemple d'un énoncé sous forme de graphe.....	154
Figure 3-24. L'attribution des IRI dans un graphe RDF.....	155
Figure 3-25. Graph RDF contenant un enchaînement des plusieurs triplets RDF, des nœuds littéraux et un nœud blanc.....	156
Figure 3-26. Le prédicat <code>rdf:type</code> et la métaclasse <code>rdfs:Class</code>	158
Figure 3-27. La définition d'une propriété RDF	158
Figure 3-28. Exemple d'un graphe RDF avec la séparation de la connaissance terminologique.....	159
Figure 4-1. Hiérarchie des modèles.....	172
Figure 4-2. Les éléments d'une base de connaissances	173
Figure 4-3. Sous-langages pour le langage OWL.....	177
Figure 4-4. Familles de logique.....	187
Figure 4-5. Les profils d'OWL 2 DL et leurs classes de complexité.....	189
Figure 4-6. Représentation d'une interprétation DL	191
Figure 4-7. Types d'ontologies.....	199
Figure 4-8. Spectre ontologique	200
Figure 4-9. L'articulation entre les ontologies FOAF, SKOS et SIOC.....	202
Figure 5-1. Objets dans le système d'apprentissage	209
Figure 5-2. Objets dans le modèle du SASA.....	210
Figure 5-3. Dimensions et caractéristiques de l'apprenant	213
Figure 5-4. Niveau d'acquisition des connaissances	214
Figure 5-5. Représentation du style d'apprentissage	214
Figure 5-6. Le style d'apprentissage Felder-Felderman.....	215
Figure 5-7. Préférences personnelles de l'apprenant.....	216
Figure 5-8. Représentation de la catégorie <i>Général</i> de LOM.....	217
Figure 5-9. Représentation de la catégorie <i>Pédagogique</i> de LOM	218
Figure 5-10. Représentation de la catégorie <i>Classification</i> de LOM.....	220
Figure 5-11. Relation <i>objet pédagogique - ressource d'apprentissage</i>	221
Figure 5-12. Modélisation des objectifs d'apprentissage.....	224
Figure 5-13. La modélisation de l'objet d'évaluation.....	227
Figure 5-14. Modélisation d'objet d'annotation.....	228
Figure 5-15. Modélisation du cadre d'enseignement.....	229
Figure 5-16. Relations entre les systèmes et les ontologies du SASA.....	239

Figure 6-1. Le SASA à l'intersection de l'apprentissage, de la systémique et du Web sémantique	241
Figure 6-2. L'architecture sur niveaux du SASA.....	243
Figure 6-3. Les modules du SASA	245
Figure 6-4. Le noyau sémantique	246
Figure 6-5. Système actif d'apprentissage	248
Figure 6-6. Un exemple de graphe RDF avec d'objectifs d'apprentissage	249
Figure 6-7. Dimensions du style d'apprentissage Felder-Silverman	254
Figure 6-8. Métadonnées d'une ressource de type vidéo sur YouTube	264
Figure 6-9. L'importation des ontologies en Protégé	266
Figure 6-10. L'intégration des ontologies réalisées au niveau de l'ontologie d'application	266
Figure 6-11. Les contrôleurs dans le SASA.....	267
Figure 6-12. L'architecture des classes sur le serveur	268
Figure 6-13. L'architecture des classes sur le client.....	268
Figure 7-1. Mesures sur les ontologies du SASA	271
Figure 7-2. Parcours d'apprentissage sans adaptation.....	273
Figure 7-3. Parcours d'apprentissage adapté.....	273
Figure 7-4. Identification des ressources externes au SASA.....	274

Liste des tableaux

Tableau 1-1. Classification de M. Develay	10
Tableau 1-2. Sommaire de modèles de contenu et niveaux de granularité de l'objet pédagogique	62
Tableau 2-1. Les approches analytique et systémique.....	76
Tableau 3-1. Sommaire des inférences suite à la déréréfencement d'un IRI	150
Tableau 3-2. Espaces de nommage des vocabulaires de base du Web sémantique	152
Tableau 3-3. Règles de déduction dans le vocabulaire RDF.....	160
Tableau 3-4. Règles de déduction dans le vocabulaire RDFS	160
Tableau 4-1. Représentation en OWL des constructeurs des logiques de description et leur syntaxe en FOL (First-Order Logics).....	186
Tableau 5-1. Association entre les éléments du LOM et LRMI.....	219
Tableau 6-1. La « distance » entre les ressources pédagogiques et le style d'apprentissage.....	255
Tableau 6-2. Caractéristiques des objets pédagogiques associées aux dimensions du style d'apprentissage.....	259

Technologies sémantiques pour un système actif d'apprentissage

Mots-clés : apprentissage en ligne, système actif d'apprentissage, ingénierie des connaissances, modélisation sémantique, ontologie

Résumé de la thèse :

Les méthodes d'apprentissage évoluent et aux modèles classiques d'enseignement viennent s'ajouter de nouveaux paradigmes, dont les systèmes d'information et de communication, notamment le Web, sont une partie essentielle. Afin d'améliorer la capacité de traitement de l'information de ces systèmes, le Web sémantique définit un modèle de description de ressources (Resource Description Framework - RDF), ainsi qu'un langage pour la définition d'ontologies (Web Ontology Language - OWL).

Partant des concepts, des méthodes, des théories d'apprentissage et en suivant une approche systémique, nous avons utilisé les technologies du Web sémantique pour réaliser une plateforme d'apprentissage capable d'enrichir et de personnaliser l'expérience de l'apprenant. Les résultats de nos travaux sont concrétisés dans la proposition d'un prototype pour un Système Actif et Sémantique d'Apprentissage (SASA). Suite à l'identification et la modélisation des entités participant à l'apprentissage, nous avons construit six ontologies, englobant les caractéristiques de ces entités. Elles sont les suivantes : (1) ontologie de l'apprenant, (2) ontologie de l'objet pédagogique, (3) ontologie de l'objectif d'apprentissage, (4) ontologie de l'objet d'évaluation, (5) ontologie de l'objet d'annotation, (6) ontologie du cadre d'enseignement.

L'intégration des règles au niveau des ontologies déclarées, cumulée avec les capacités de raisonnement des moteurs d'inférences incorporés au niveau du noyau sémantique du SASA, permettent l'adaptation du contenu d'apprentissage aux particularités des apprenants. L'utilisation des technologies sémantiques facilite l'identification des ressources d'apprentissage existant sur le Web ainsi que l'interprétation et l'agrégation de ces ressources dans le cadre du SASA.

Semantic Technologies for an Active Learning System

Keywords: e-Learning, active learning system, knowledge engineering, semantic modeling, ontology

Abstract:

Learning methods keep evolving and new paradigms are added to traditional teaching models where the information and communication systems, particularly the Web, are an essential part. In order to improve the processing capacity of information systems, the Semantic Web defines a model for describing resources (Resource Description Framework - RDF), and a language for defining ontologies (Web Ontology Language - OWL).

Based on concepts, methods, learning theories, and following a systemic approach we have used Semantic Web technologies in order to provide a learning system that is able to enrich and personalize the experience of the learner. As a result of our work we are proposing a prototype for an Active Semantic Learning System (SASA). Following the identification and modeling of entities involved in the learning process, we created the following six ontologies that summarize the characteristics of these entities: (1) learner ontology, (2) learning object ontology, (3) learning objective ontology, (4) evaluation object ontology, (5) annotation object ontology and (6) learning framework ontology.

Integrating certain rules in the declared ontologies combined with reasoning capacities of the inference engines embedded in the kernel of the SASA allow the adaptation of learning content to the characteristics of learners. The use of semantic technologies facilitates the identification of existing learning resources on the web as well as the interpretation and aggregation of these resources within the context of SASA.