ED 211

École Doctorale des Sciences Exactes et leurs Applications

Gestion des risques appliquée aux systèmes d'information distribués

THÈSE

présentée et soutenue publiquement le jeudi 19 décembre 2013

pour l'obtention du

Doctorat de l'Université de Pau et des Pays de l'Adour

(spécialité informatique)

par

Vincent Lalanne

Sous la direction de

Alban Gabillon, Pr - Université de la Polynésie Française
et
Manuel Munier, MCF - Université de Pau et des Pays de l'Adour

Composition du jury

Président :	**Jean-Jacques Lemouland**, Pr	Université de Pau et des Pays de l'Adour
Rapporteurs :	**Hossam Afifi**, Pr	Télécom SudParis
	Jean-Philippe Domenger, Pr	LAboratoire Bordelais de Recherche en Informatique
Examinateurs :	**Philippe Aniorté**, Pr	Université de Pau et des Pays de l'Adour
	Serge Dulucq, Pr	LAboratoire Bordelais de Recherche en Informatique
	Benoît Le Blanc, MCF HDR	École Nationale Supérieure de Cognitique
	Manuel Munier, MCF	Université de Pau et des Pays de l'Adour

Résumé

Dans cette thèse nous abordons la gestion des risques appliquée aux systèmes d'information distribués. Nous traitons des problèmes d'interopérabilité et de sécurisation des échanges dans les systèmes DRM et nous proposons la mise en place de ce système pour l'entreprise : il doit nous permettre de distribuer des contenus auto-protégés. Ensuite nous présentons la participation à la création d'une entreprise innovante qui met en avant la sécurité de l'information, avec en particulier la gestion des risques au travers de la norme ISO/IEC 27005:2011. Nous présentons les risques liés à l'utilisation de services avec un accent tout particulier sur les risques autres que les risques technologiques ; nous abordons les risques inhérents au cloud (défaillance d'un provider, etc...) mais également les aspects plus sournois d'espionnage et d'intrusion dans les données personnelles (Affaire PRISM en juin 2013). Dans la dernière partie nous présentons un concept de DRM d'Entreprise qui utilise les métadonnées pour deployer des contextes dans les modèles de contrôle d'usage. Nous proposons une ébauche de formalisation des métadonnées nécessaires à la mise en œuvre de la politique de sécurité et nous garantissons le respect de la réglementation et de la loi en vigueur.

Mots-clés: ISO/IEC 27005:2011, Web Services, WS security, SOA, Cloud, métadonnées, données personnelles, sécurité de l'information, entreprise innovante, investigation numérique

Abstract

In this thesis we discuss the application of risk management to distributed information systems. We handle problems of interoperability and securisation of the exchanges within DRM systems and we propose the implementation of this system for the company : it needs to permit the distribution of self-protected contents. We then present the (our) participation in the creation of an innovative company which emphasizes on the security of information, in particular the management of risks through the ISO/IEC 27005:2011 standard. We present risks related to the use of services, highlighting in particular the ones which are not technological : we approach inheritent risks in clouds (provider failure, etc ...) but also the more insidious aspects of espionage and intrusion in personal data (Case PRISM in June 2013). In the last section, we present a concept of a DRM company which uses metadata to deploy settings in usage control models. We propose a draft formalization of metadata necessary for the implementation of a security policy and guarantee respect of regulations and legislation.

Keywords: ISO/IEC 27005:2011, Web Services, WS security, SOA, Cloud, metadata, personal data, information security, innovative, digital forensics

Remerciements

Je tiens à remercier ma famille Magali, Clément et Julien, pour la patience et la compréhension dont ils ont fait preuve pendant ces années de recherche.

Je voudrais également remercier les personnes qui m'ont soutenu, encouragé et accompagné pendant ce travail de thèse et tout particulièrement M. Manuel Munier qui a co-dirigé mon travail.

Enfin je tiens à dire un grand merci à l'Université de Pau et des Pays de l'Adour qui m'a permis de mener à terme ce projet.

Aux compañeros qui m'ont indiqué le chemin

La Habana, 2002

Table des matières

Introduction générale

Sommaire

1 Contexte et motivations

La sécurité des systèmes d'information est un problème sensible et préoccupant qui nécessite la mise en place de plusieurs outils et mécanismes de la sécurité informatique. Ces systèmes, autrefois enfermés dans l'entreprise et parfaitement hermétiques au monde extérieur, se sont peu à peu ouverts, d'abord par nécessité de connexions distantes et ensuite pour collecter des informations d'ailleurs. Les mécanismes utilisés pour la sécurisation de ces systèmes font partie de la panoplie maintenant classique dont dispose un ingénieur sécurité.

Dans une architecture orientée services c'est la même chose : une somme d'éléments interconnectés qui ont chacun une fonction bien définie, génèrent des services pour les autres entités. L'assemblage de ses fonctions a pour objet de réaliser un système. Dans une telle structure, il est fondamental que ces services soient fiables et peu sensibles aux attaques extérieures. De la même manière que pour un système d'information, il faudra sécuriser ces services.

C'est le fondement même des systèmes de DRM [1] : système de gestion de contenus. Afin de pouvoir continuer un modèle économique identique au modèle précédent, les éditeurs de contenus musicaux ont fait développer des plates-formes capables de distribuer des morceaux de musiques numériques. Or ces systèmes peuvent également gérer les droits d'utilisation et d'exploitation de ces contenus numériques. Étant développés par des marques différentes, ils n'ont pas réussi à être normalisés et ont donc développé des stratégies différentes avec des technologies différentes. Malgré une obligation légale de rendre ces outils parfaitement interopérables, il n'en a rien été. Nous nous sommes donc penchés sur la possibilité de rendre le système interopérable, nous avons pointé les écueils à contourner ou à franchir afin d'obtenir le résultat. Nous avons également étudié la possibilité de développer la transmission d'informations sur ces services.

Quand un système s'ouvre sur l'extérieur afin de collecter ou de stocker des informations, il est fondamental d'avoir confiance aux fournisseurs de ces informations. Un architecte système, toujours frileux à l'idée d'ouvrir au monde son système d'information, doit pouvoir être sûr de ses prestataires. Il est donc nécessaire de mesurer les probabilités de panne, les dysfonctionnements possibles : rapidement une analyse de risque s'impose. La norme ISO/IEC 27005:2011, apparue en 2008, correspond tout à fait à ce besoin. C'est une norme de gestion des risques liés à la sécurité de l'information, c'est-à-dire qu'elle permet de quantifier un risque par rapport à une menace possible. Il nous a paru essentiel

1. Digital Rights Management

d'étendre cette norme à la notion de service : dans le développement de ce standard la notion de service n'est pas prise en compte. De même, à côté de l'aspect technique de la perte d'un service, il faut prendre en compte l'aspect économique lié à la fourniture de ce service : nous sommes dans une relation client fournisseur. Une relation client fournisseur est matérialisée par un contrat, qui, s'il est rompu, interrompt le service, ou du moins le dégrade. Nous nous sommes donc penchés sur cet aspect, au-delà du critère technique, le résultat sera le même si le service est interrompu. Il nous a donc paru fondamental de l'intégrer dans notre gestion des risques. De la même manière le risque d'espionnage est évident, ou sans aller à la notion d'espionnage on peut se contenter de pertes de confidentialité. Nos propositions en ce sens ont été magistralement démontrées par les révélations d'Edward Snowden et l'affaire PRISM : quelque soit le mécanisme de sécurité employée, il existe un accès direct à l'intérieur du serveur ! La notion de documents auto protégé telle que nous la développons ici paraît être une solution satisfaisante.

Dans une dernière partie nous développons l'utilisation des métadonnées comme élément de gestion et de sécurité dans un environnement de DRM d'entreprise. Les métadonnées, ces « données de données » sont un élément à prendre en compte dans une future gestion des risques, en effet doit-on les considérer comme des données ou simplement comme des indicateurs techniques ?

2 Organisation du mémoire

Ce mémoire est organisé sous la forme d'une thèse par articles. Cette présentation, plutôt atypique dans le monde de l'informatique, présente l'avantage de pouvoir, à travers la chronologie des articles, exposer les problématiques qui se sont présentées et les propositions que nous avons faites pour y rémédier.

Dans un premier chapitre nous allons aborder les problématiques rencontrées dans les domaines qui mettent en œuvre la sécurité des systèmes. Sous la forme d'un état de l'art nous allons aborder l'organisation d'un système DRM en pointant les éléments qui le constituent, en particulier le point crucial que nous avons mis en évidence qui est celui de l'interopérabilité des systèmes. Dans une seconde partie nous allons développer les notions fondamentales qui constituent une architecture orientée services. Ensuite à travers la norme ISO/IEC 27005:2011 nous allons expliquer ce qu'est le risque en sécurité de l'information et comment le gérer. Enfin, nous allons traiter des métadonnées, ces données nous sont créées « sur les données », sont souvent traitées et stockées à l'insu

de leurs utilisateurs. Nous conclurons ce chapitre par une synthèse des problématiques rencontrées. Dans le chapitre trois, à travers un exemple concret qui est celui de la création de l'entreprise BackPlan™ à laquelle j'ai participé, nous présenterons l'intérêt d'une certification à la gestion des risques en sécurité de l'information dans le monde industriel, garantissant ainsi la confiance pour les clients. Dans le chapitre quatre (TRUSTCOM 2012) nous proposerons la mise en place une technologie DRM dédiée au monde de l'entreprise en apportant une solution au problème de la sécurisation des données lorsqu'elles sont transportés sur des systèmes mobiles ou exposées dans des serveurs peu sûrs. Le chapitre cinq (PASSAT 2013) présentera une solution à la prise en compte des services dans la gestion des risques des systèmes d'information. En effet dans le développement de la norme ISO/IEC 27005:2011 il n'est jamais fait mention de l'élément service : à travers cet article nous proposerons de prendre en compte la notion même de service dans la gestion du risque en sécurité de l'information. Le chapitre six concerne l'article présenté à la conférence SARSSI 2013 ou nous poserons les bases de notre méthode d'évaluation des risques que nous avons développé dans le chapitre précédent (la date de soumission dans cette conférence était bien antérieure à celle de PASSAT 2013). Dans l'article (DPM 2013) nous mettrons en évidence les notions de métadonnées qui peuvent être collectées, utilisées, stockées à l'insu des utilisateurs. Nous avancerons des propositions pour garantir le respect de ces informations par l'anonymisation mais aussi grâce au cadre législatif qui doit prendre en compte ces métadonnées dans le document. Le chapitre huit présente l'article présenté à la convention nationale de la recherche en IUT (CNR IUT 2013). Il permettra de présenter les axes de recherche communs pouvant lier le monde des juristes au monde de l'informatique. Enfin le dernier chapitre résume l'apport de la thèse et donne les perpectives à venir.

Table des figures

Chapitre 1

Problématique

Sommaire

Dans ce chapitre je vais aborder les différentes voies que j'ai explorées pendant mon travail de recherche. Toutes ces orientations ont pour point commun les notions de service et de sécurité : du système de gestion de droits système DRM, véritable architecture orientée services dédiée à la distribution et à l'utilisation de contenu numérique, à la gestion des risques dans les systèmes d'information mettant en œuvre les services Web, aux métadonnées qui sont générées pendant les nombreux échanges et dialogues entre les différents éléments de ces systèmes.

Ces problématiques ont bien sûr évolué au cours de mes travaux : en fait de nouvelles pistes se sont ouvertes au fûr et à mesure de ma progression. Du problème initial de la mise en place d'une sécurité des services Web au sein d'un système DRM nous terminons ce tour d'horizon sur la gestion des métadonnées et les conséquences en terme de sécurité de l'information.

Dans tous les cas, le fil rouge que nous allons suivre tout au long de ce chapitre est celui de la sécurité. En premier lieu celle que l'on déploie sur une architecture informatique afin de garantir un bon fonctionnement des systèmes et une bonne tenue aux attaques et en second celle qui concerne la sécurité des données et de l'information.

1.1 Le contrôle d'usage

1.1.1 Les DRM : Digital Rights Managements

Le contenu numérique (audio, vidéo, graphiques et images) peut être facilement copié, transmis et diffusé sur les réseaux. L'expansion de l'Internet ainsi que les nombreux outils de partage de fichiers ont fait que la distribution de médias numériques protégés n'est plus du tout contrôlée. Des copies exactes de données numériques ne sont absolument pas distinguées des originaux, ce qui entrave la croissance des services multimédias, en effet cet aspect à un impact négatif sur la volonté des auteurs, des éditeurs et des fournisseurs de produire de nouvelles œuvres.

Bien que le contenu en ligne soit protégé par la loi du copyright, par des outils de police en ligne (HADOPI), le volume de données contrefaites n'a pas cessé d'augmenter. La nécessité d'une technologie qui puisse mettre l'accent sur la protection des données numériques est devenue de plus en plus évidente.

La gestion des droits numériques

Typiquement, un système de DRM protège la propriété intellectuelle :

— par le cryptage des données de sorte qu'il ne peut être accessible que par des utilisateurs autorisés ou par le marquage du contenu,

— avec un filigrane numérique, de sorte que le contenu ne peut être distribué librement.

Quelle que soit la méthode, le DRM assure que le contenu numérique est utilisé uniquement par ceux qui ont payé ou qui ont le droit de l'utiliser. En ce sens, le DRM représente un concept de gestion et de contrôle de l'accès et de l'utilisation de la ressource. Comme son nom l'indique, il est uniquement applicable aux médias numériques.

La gestion des droits numériques est parfois appelée gestion de restriction numérique. En général, cependant, l'objectif du DRM est perçue comme un moyen d'empêcher les gens, par des moyens technologiques, de l'utilisation de matériel protégé par des moyens qui sont inacceptables pour l'éditeur ou l'auteur dans le cas de DRM d'entreprise.

Il est important de prendre en considération les points de vue de l'utilisateur afin de comprendre les avantages, ainsi que les différents problèmes et les conflits que peuvent mener les systèmes de DRM.

L'architecture DRM

Il n'y a pas d'architecture type, mais il y a plusieurs types de systèmes qui sont proposés par les fabricants.

Dans tous les cas, un système de DRM peut être modélisé par trois blocs tels qu'ils sont définis par Park et al. [PSS00]**.**

Le premier **Capture et création de contenu** se réfère au processus de gestion et de création de contenu. Cela comprend la gestion de droits lorsque le contenu est créé pour la première fois (ou réutilisé et étendu avec des droits appropriés pour le faire) par différents créateurs de contenu / fournisseurs. Ce module prend en charge en particulier la validation des droits (s'assurer que le contenu est créé à partir de contenu existant incluant des droits pour le faire), la création des droits (les droits sont affectés à de nouveaux contenus tels que le propriétaire des droits et les autorisations d'utilisation) et le Workflow de droits (permet au contenu d'être traité par une série d'étapes de processus d'examen et / ou d'approbation de droits).

Le second **Gestion des contenus** désigne le processus de gestion et la facilitation du commerce de contenu. Cela comprend la réception du contenu des créateurs et l'utilisation

comme une entrée pour un système de gestion d'actifs. Les données gérées dans ce système comprennent les métadonnées descriptives et les métadonnées de droits (par exemple : les morceaux, les usages, les paiements, etc.). Ce module prend en charge les fonctions de dépôt (permettre la consultation / extraction de contenu) dans des bases de données distribuées et la consultation / extraction de métadonnées. Les métadonnées couvrent les morceaux, les droits (et les descriptions des travaux) et les fonctions de négociation qui permettent la délivrance de licences aux parties qui ont négocié les accords pour les droits sur le contenu.

Le troisième **Utilisation des contenus** se réfère au processus de l'utilisation du contenu après la transaction. Cela inclut les contraintes liées au contenu échangé. Il prend en charge la gestion des droits (permettre l'application des droits associés au contenu dans l'environnement d'utilisation : par exemple, si l'utilisateur a seulement le droit de consulter le document, l'impression ne sera pas autorisée). Il s'occupe également de la gestion du suivi (permettre le suivi de l'utilisation du contenu lorsque ce suivi fait partie de l'accord de licence : par exemple , l'utilisateur dispose d'une licence pour lire une vidéo dix fois).

Mis ensemble, ces trois parties composent les fonctionnalités de base d'un système DRM.

Techniques employées dans les systèmes DRM

Afin de fonctionner, les systèmes de DRM utilisent principalement les techniques de chiffrement, clés publique / privée, certificats numériques, watermarking (tatouage), contrôle d'accès, protocoles de communication sécurisés, empreinte, langage d'expression de droits, infrastructure de confiance et hachage.

Les objectifs des DRM

L'objectif d'un système DRM est de fournir une plateforme de confiance autant pour le fournisseur de contenu que pour ses clients.

1. Les DRM doivent pouvoir fournir un contenu numérique protégé. Ce type de protection est garanti par des techniques de chiffrement, ce qui permet aux auteurs et aux producteurs d'envoyer du contenu numérique à travers des réseaux non sécurisés.
2. Les DRM doivent permettre une distribution sécurisée du contenu. Les clés de chiffrement utilisées doivent garantir qu'elles ne seront pas cassées, que le contenu

ne sera pas déchiffré.

3. Les DRM doivent garantir l'authenticité du contenu. Ce sont les clés de hashage qui vont garantir cette fonction.
4. Les DRM doivent garantir que la transaction ne sera pas remise en cause. C'est grâce à la signature électronique que nous pourrons valider cette fonction.
5. Les DRM doivent permettre l'identification des ayants droits. Là ce sont les certificats qui vont être utilisés afin de garantir qu'il s'agit de la bonne personne.

1.1.2 L'interopérabilité des systèmes DRM

La loi Françise du 1er août 2006, dite loi DADVSI, relative au droit d'auteur et aux droits voisins dans la société de l'information dispose à son article L. 331-5 que : « *les mesures techniques ne doivent pas avoir pour effet d'empêcher la mise en effective de l'interopérabilité, dans le respect du droit d'auteur* ».

De manière générale, l'onteropérabilité des systèmes DRM dépend :

— des protocoles de communication entre les composants du système,
— du mécanisme de protection du contenu,
— du langage d'expression de droits (REL : Rights expression Language) utilisé dans chacun des systèmes.

Dans le Chapitre 2 je présente l'article « Vers l'interopérabilité des Systèmes de DRM »présenté à SARSSI 2007, Seconde conférence jointe SAR [6ème Conférence sur la Sécurité des Architectures Réseaux - SAR] - SSI [4ème Conférence sur la Sécurité des Systèmes d'Information - SSI] organisée à Annecy du 12 aux 15 juin 2007.

1.1.3 Les documents sécurisés : DRM d'entreprise

Nous abordons ici une application directe du système des DRM mais dédiée à l'entreprise.

En effet, au cours des années, les mécanismes DRM dits "grand public" sont devenus des sources de mécontentement car empêchant la portabilité d'un système vers un autre. Bien souvent l'utilisateur devait contourner ces protections pour bénéficier des produits qu'il avait acquis.

Mais pour autant le principe de DRM avec son mécanisme de contrôle d'usage est un élément important que nous avons voulu adapter au monde de l'entreprise.

Pourquoi ne pas associer au document son propre système de contrôle d'usage ?

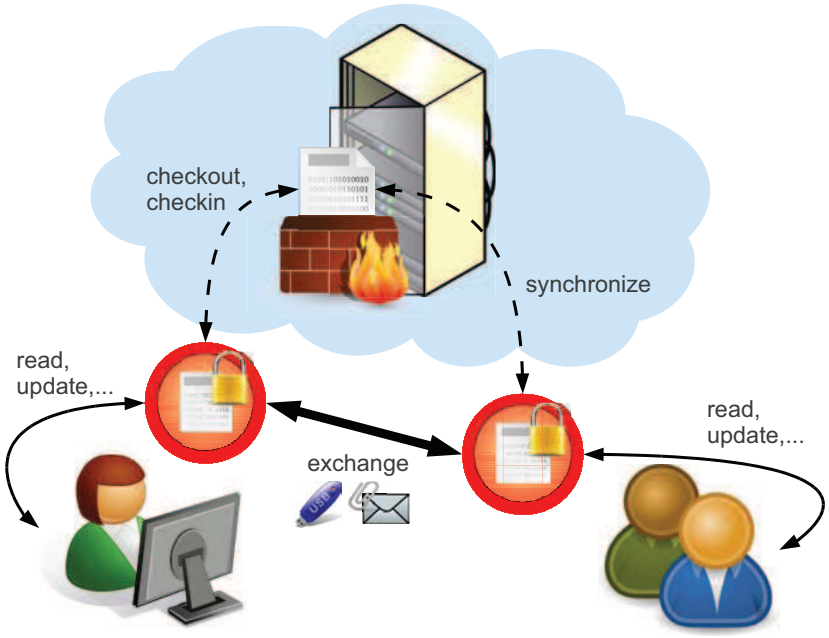


FIGURE 1.1 – Parcours d'un document sécurisé

C'est là qu'apparait la notion de document complexe ou document dit "intelligent".

Nous prenons par exemple le rapport d'activité annuel d'une entreprise : la partie finance est rédigée par le reponsable du service financier, le contrôle de fabrication par le responsable qualité, etc... Or il apparait tout à fait logique que dans ce type de document chacun des collaborateurs ait ses propres droits : lecture, écriture sur tout ou partie du document.

Ainsi un document intelligent pourrait s'articuler de la manière suivante :

- Le document encapsulé : il représente à la fois le contenu et tout le mécanisme de contrôle, c'est une sorte de Système d'Information embarqué.
- Les données : il s'agit du contenu du document (cette partie peut bien sûr être chiffrée.
- Les composants de sécurité intégrés qui doivent permettre :
  - Les contrôles d'usage et d'accès
  - La traçabilité du document : Qui a fait quoi ? À quelle heure ? Où ? ...
  - Le travail collaboratif entre différents participants

Dans ce type d'organisation il est indispensable, comme dans tout système d'information, de mettre en place des balises necessaires au suivi du système : ces balises sont décrites comme faisant parie de la famille des métadonnées.

Dans l'article présenté à TrustCom 2012 (Chapitre 4)

Self-Protecting Documents for Cloud Storage Security - *Manuel Munier, Vincent Lalanne, Magali Ricarde*[2]

11ème conférence IEEE TrustCom 2012 organisée à Liverpool (Grande Bretagne) du 25 au 27 juin 2012

Dans cet article nous présentons notre vision du document sécurisé appliqué au domaine Oil & Gas.

1.1.4 Les métadonnées

Définition :

In extenso ce terme signifie des données par-dessus les données. Dans la pratique ce sont des données qui servent à décrire d'autres données.

La notion de métadonnée est apparue dans les domaines des catalogue d'ouvrages numériques [CML09], plus précisément dans le métier des bibliothèques en ligne. Quand on parcourt l'Internet, nous sommes en face d'une bibliothèque qui ne dispose pas de catalogue, or ce sont les métadonnées qui permettent de retrouver des renseignements sur les ressources numériques : documents, images, sons, vidéo etc.

Les métadonnées sont partout, elles sont générées et véhiculées à travers tous les services du Web. Quand vous prenez une photo avec votre téléphone, les coordonnées GPS sont accolées à votre image. Lorsque vous accédez à un service en ligne, tous vos faits et gestes sont enregistrés, vous ne pouvez pas réfuter le fait que vous avez consulté ce document.

Ces informations ne sont pas considérées comme des informations liées aux documents que vous transportez, mais peuvent par récollement générer de l'information et avoir dans le domaine industriel une valeur contradictoire.

Des informations « sensibles »

Necessaires au bon fonction du système d'information

Dans tout système d'information les métadonnées sont nécessaires d'un point de vue de la sécurité. En effet elles permettent le contrôle du système et son bon fonctionnement.

À travers le contrôle d'usage il est possible de déterminer les dates, lieux d'accès aux données.

Ceci rejoint le domaine de la traçabilité et nous le verrons ci-après la notion de preuve probante.

Le revers de la médaille : une utilisation détournée

D'autre part il ne faut pas exclure que ce type d'information peut également être détourné et

2. Société BackPlan™

servir à des fins qui n'étaient pas celles prévues à l'origine :

— Falsification des métadonnées
— Surveillance des utilisateurs : géolocalisation
— Extraction et déduction d'informations personnelles
— Calcul automatisé d'indicateurs : calcul de performance de l'individu, perte de confiance
— Entrepôt de données - Big Data

Une exposition de son patrimoine informationnel

Inexorablement nous nous apercevons que ces métadonnées sont parfois plus importantes que les données auxquelles elles sont associées.

En effet elle permettent de rentrer dans la vie privée des individus, mais peuvent également dévoiler des données stratégiques de l'entreprise et réveler un patrimoine informationnel non désiré.

Une autre conséquence peut être la mise à jour des relations inter organisationnelle des individus de ces entreprises.

A-t-on le droit de tout collecter ?

Au cours de nos recherche il nous a semblé naturel de se poser cette question.

En effet d'un point de vue technique et pour le bon fonctionnement, certains indicateurs sont absoluement necessaires. Mais d'un point de vue du droit, et en particulier sous l'angle de la vue privée, peut-on le faire ?

C'est pour toutes raisons que nous nous sommes rapprochés du Laboratoi de droit Privé.

Collaboration avec le Laboratoire de droit privé de l'université de Pau et des Pays de l'Adour

Il nous a donc paru tout à fait naturel de nous rapprocher des juristes de droit privé de notre université. Nous avons donc commencé à collaborer avec le **Centre de Recherche et d'Analyse Juridiques**[3] de l'université de Pau et des Pays de l'Adour. Nous avons participé à un colloque sur « Internet et droit d'expression des salariés » (12 janvier 2012 - organisé par le CRAJ). Nous y avons présenté une sensibilisation technique aux risques de la surveillance de l'utilisation d'Internet par les employés dans une entreprise.

Ensuite nous avons mis en place quelques réunions de présentation de nos activités communes, puis rapidement nous nous sommes mis d'accord sur deux axes : la publication d'un article en commun dans le domaine informatique et le soutien technique dans des domaines encore inconnus des juristes. Nous avons réalisé ce premier point avec DPM 2013 (Chapitre 7)

3. C.R.A.J. - EA 1929

et et CNR IUT 2013 (Chapitre 8). Ce partenariat a mené une étudiante de Master deuxième année recherche en droit privé général et appliqué, Camille Drouiller, a présenter un rapport sur « La preuve par les métadonnées », dirigée par M. Pierre-Yves Ardoy.

1.2 Architecture Orientée Services et Web Services

Le concept d'architecture orientée service (SOA) a pris plusieurs formes au cours du temps, mais il est devenu vraiment utilisé avec l'apparition de la technologie Web, des services Web et des standards développés autour de cette technologie. Cependant le concept de SOA n'est pas lié à une technologie spécifique, mais dans de nombreux cas il se réfère à des systèmes distribués utilisant des services Web pour communiquer. D'autres exemples de SOA sont basés sur des appels de procédure qui utilisent des standards binaires propriétaires causant ainsi de sérieux problèmes d'interopérabilité. Les services Web permettent de solutionner les problèmes d'interopérabilité parce qu'ils sont basés sur le langage XML qui est un standard naturellement interopérable. Des efforts significatifs ont été mis en place pour développer la sécurité des services Web afin de leur conforter les critères d'intégrité, de confidentialité, d'authentification, de confiance.

1.2.1 Introduction

SOA est un modèle d'architecture basé sur des services indépendants, avec des interfaces bien définies qui permettront une réutilisation future.

L'idée sous-jacente est de fournir des services avec des interfaces bien pensées qui peuvent être combinées ensemble de différentes manières pour construire différents types d'applications.

Les services présentent leurs fonctionnalités à travers un standard qui utilise un langage de description du service Web (WSDL : Web Services Description Language).

À côté de l'organisation des différentes briques necessaires à l'élaboration d'une application utilisant les services Web il sera nécessaire de penser également à la sécurité du système mis en place :

— sécurité du réseau, de l'application.
— Sécurité du transport : de manière classique ou utilisant SSL (Secure Sockets Layer).
— La sécurité des services Web, à travers l'utilisation de la famille Web Service-Security (WS-*) de standards destinés à sécuriser les messages de services Web.
— L'utilisation d'autres extensions peuvent être employées afin de garantir la confiance entre la compagnie et les services de paiement et d'envoi.

1.2.2 Les avantages d'une SOA

Une architecture orientée services permet d'obtenir tous les avantages d'une architecture client-serveur et notamment la modularité qui permet de remplacer facilement un composant (service) par un autre. Elle offre ainsi une possibilité de réutiliser des composants (par opposition à une système tout-en-un fait sur mesure pour une organisation), de meilleures possibilités d'évolution (il suffit de faire évoluer un service ou d'ajouter un nouveau service), une plus grande tolérance aux pannes et une maintenance facilitée.

1.2.3 Un style d'architecture

Un style d'architecture n'est pas un standard, il s'appuie sur des standards. Lorsqu'on parle de services Web, les standards habituellement employés sont :

— URI pour le nommage des ressources ou des services ;
— HTTP ou HTTPS pour le protocole de communication ;
— XML et XHTML pour les données et les documents ;
— liens hypermédia avec des types MIME pour la navigation et l'accès aux ressources ;
— Le DOM (Document Object Model) comme plate-forme et interface indépendante d'un langage pour lire et modifier le contenu, la structure et le style des documents ;
— Javascript comme langage de scripting pour manipuler le DOM ;
— PKI (Public Key Infrastructure) et
— X509 comme standards pour assurer la sécurité avec des certificats à clés publiques.

Dans le domaine des styles d'architecture, REST et SOAP sont les styles dominants.

Les services Web ainsi que leur sécurité sont basés sur plusieurs standards qui vont être présentés dans les paragraphes qui suivent.

1.2.4 XML : eXtensible Markup Language

XML est la pierre angulaire de la technologie des services Web et de leur sécurité. Les services Web ainsi que les standards qui assurent leurs sécurité utilisent fortement le langage XML. XML a révolutionné les échanges de données grâces à sa simplicité et sa puissance, facilitant ainsi les échanges de données entre les applications à travers des systèmes différents. C'est une méthode d'échange basée sur de simples messages textes : ce standard est ouvert et recommandé par le consortium World Wide Web. Le fait que XML soit simple, facilement lisible par l'homme, permettant de travailler sur des systèmes hétérogènes, le rend parfaitement adapté aux échanges entre plates-formes.

1.2.5 REST : REpresentational State Tranfert

REST est l'acronyme de Representational State Transfer inventé par Roy T. Fielding [FT00].

REST est un modèle de conception (ou *design pattern*) pour l'implémentation de systèmes connectés. Ce n'est ni une technologie ni un standard ; c'est un type d'architecture pour l'exposition de ressources sur le web.

L'architecture RESTful adhère à plusieurs principes :

— les requêtes sont client-serveur ;

— les requêtes sont sans état ;

— clients et serveurs adhèrent à une interface uniforme ; toutes les ressources sont accédées via une interface générique : les méthodes HTTP GET, POST, PUT, DELETE, HEAD ou OPTION ;

— les clients accèdent à des ressources nommées en utilisant des URL, comme des URL HTTP (mais ne se limitant pas à des URL HTTP) ; pour faire simple une ressource est représentée par une URL.

REST n'est pas un standard, il n'existe donc pas de spécifications de REST. Il faut comprendre le style REST et ensuite concevoir des applications ou des services Web selon ce style.

Bien que REST ne soit pas un standard, il utilise des standards. En particulier :

— URI comme syntaxe universelle pour adresser les ressources ;

— HTTP ou HTTPS, protocole sans état [stateless] avec un nombre très limité d'opérations ;

— des liens hypermédia dans des documents XML et XHTML pour représenter à la fois le contenu des informations et la transition entre états de l'application ;

— les types MIME comme text/xml, text/html, image/ jpeg, application/pdf, video/mpeg pour la représentation des ressources.

— REST concerne l'architecture globale d'un système. Il ne définit pas la manière de réaliser dans les détails. En particulier, des services REST peuvent être réalisés en.NET, JAVA, CGI etc.

La sécurité des informations transmises est assurée par le protocole de transport HTTPS.

1.2.6 SOAP : Simple Object Access Protocol

SOAP est un protocole de message qui constitue la base de la pile de protocole des services Web. Les messages SOAP sont conçus pour être indépendants du protocole de transport, mais sont le plus souvent transmis via le protocole HTTP ou HTTPS lorsqu'il est utilisé avec des services Web. Les messages SOAP ne sont pas liés au protocole HTTP, cependant, ils peuvent être utilisés dans des systèmes de files d'attente de messages, envoyés par e-mail ou par d'autres mécanismes de transport.

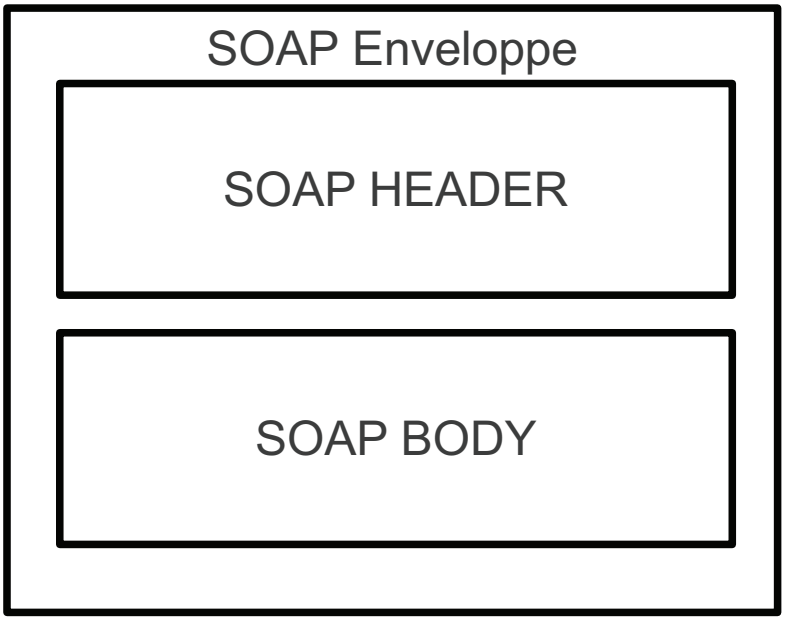


FIGURE 1.2 – Message SOAP

Le standard SOAP est basé sur XML, il définit la structure des messages qui peuvent être transmis entre les systèmes. Les messages SOAP sont constitués par une enveloppe, une en-tête et un corps comme représenté sur la figure 1.2. L'en-tête SOAP permet l'insertion d'éléments de sécurité tels que les signatures numériques et le chiffrement dans le message. Bien que les éléments de sécurité ne se limitent pas seulement à l'en-tête, il est largement utilisé avec les normes WS-S pour transmettre des informations de sécurité avec le message.

Il existe deux modes de messagerie primaires utilisées : le mode SOAP-document et appel de procédure distante (RPC) [Use99]. Le mode document est bien pour la transmission unidirectionnelle de messages, dans lequel l'expéditeur envoie le message SOAP, mais pour lequel on attend pas une réponse. Le mode RPC est plus couramment utilisé, c'est un modèle de question-réponse dans laquelle l'expéditeur soumet la requête SOAP et attend une réponse SOAP.

Au-delà du protocole SOAP nous pouvons identifier un certain nombre de mesures de sécurité existantes. Des spécifications supplémentaires ont été définies au dessus de la pile XML/SOAP pour renforcer la sécurité des infrastructures utilisant des services web (voir figure 1.3).

Les technologies de base autour d'XML sont toutes définies par le World Wide Web Consortium (W3C). C'est le cas par exemple pour XML Encryption [ER02] et XML Signature [BBF+08], permettant de régler respectivement la question de la confidentialité (chiffrement des données) et de l'intégrité des données (authentification du message et/ou du signataire).

À partir de ces standards, des organismes de normalisation tels que l'OASIS ont développé des spécifications telles que Web Service Security [OAS06]. C'est un lot d'extensions SOAP qui garantit au message son intégrité et sa confidentialité. Cette spécification est flexible et peut être

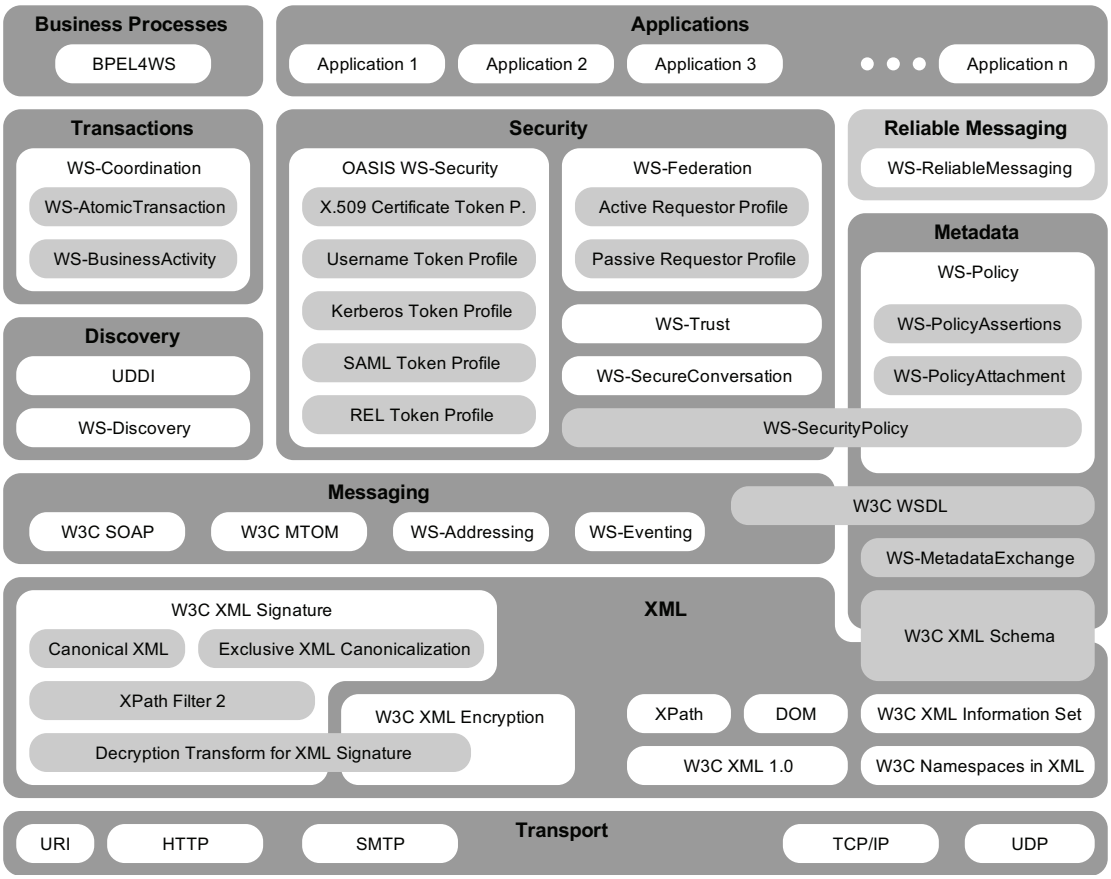


FIGURE 1.3 – La pile de spécifications des Services Web (WS-*) [GPC05]

accommodée de modèles de sécurité aussi variés que PKI, Kerberos et SSL. Un certain nombre de spécifications sont associées à WS-Security :

- **WS-Trust** : spécification permettant la génération, le renouvellement et la validation de *security tokens*, elles assurent la confiance dans le service.
- **WS-SecureConversation** : création et partage de contextes de sécurité (via un *security context token*)
- **WS-Federation** : définition de mécanismes de fédération d'espaces de confiance hétérogènes
- **WS-Authorization** : expression des autorisations
- **WS-Policy** : grammaire flexible et extensible permettant d'exprimer les possibilités, exigences et caractéristiques générales des entités (consommateurs ou fournisseurs)
- **WS-Privacy** : modèle pour indiquer comment les besoins de confidentialité et les pratiques liées aux données privées sont transmises entre organisations

D'autres spécifications concernent plus particulièrement les phases d'authentification et d'autorisation.

L'authentification est le processus pour valider les identités, tandis que l'autorisation est un processus visant à déterminer qu'une partie authentifiée peut accéder à ce type de ressources ou d'effectuer ce type d'actions. La spécification SAML [CKPM05] définit un cadre de travail pour échanger des informations d'authentification et d'autorisation entre les partenaires d'affaires. SAML prend en charge l'authentification unique (SSO) pour les sites affiliés. La spécification XACML [LPL+03] [MG03] définit quant à elle un langage pour le contrôle d'accès, la circulation des règles et l'administration de la politique de sécurité des systèmes d'information.

SOAP (Simple Object Access Protocol) est un mécanisme qui permet d'échanger des messages XML entre applications, principalement en utilisant le protocole HTTP.

1.2.7 WSDL et UDDI

WSDL [4] est un langage de description de services Web. C'est un document XML qui permet de décrire chaque service SOAP sous la forme d'une documentation en ligne. Ce document décrit les messages échangés, habituellement sous la forme d'un schéma XML, ainsi que les détails des interfaces publiques de ce service.

Un document WSDL est donc un simple document XML avec quatre éléments principaux :

- **portType** définit le service Web, les opérations qui peuvent être effectuées et les messages. C'est l'équivalent d'un nom de fonction classique ;
- **message** définit les données de chaque opération. C'est l'équivalent des paramètres d'une fonction classique ;
- **type** définit les types de données en utilisant la notation des schémas XML ;
- **binding** définit le format des messages et les détails du protocole de chaque port.

Cette documentation n'est pas réservée à SOAP. Elle est disponible en ligne. Elle permet de simplifier l'utilisation des services car elle est exploitable directement par les outils de développement et de programmation.

1.2.8 La sécurité des services Web

La sécurisation d'une architecture SOA s'avère beaucoup plus complexe que celle des environnements traditionnels. La somme de services implique la nécessité de pouvoir proposer un modèle d'abstraction qui découple la sécurité de celle des modèles liés aux différents fournisseurs de services. Ces architectures doivent faire face à un certain nombre de menaces et de vulnérabilités [OWA] [CVE] comme l'**Usurpation d'identité** (un utilisateur non accrédité usurpe l'identité d'un utilisateur valide de l'application), la **Modification des données** (un utilisateur détruit ou modifie les informations sans autorisation), la **Répudiabilité (**la possibilité qu'un

4. Web Services Description Language

utilisateur puisse nier avoir effectué telle ou telle opération), la **Divulgation d'information (**des données confidentielles sont rendues visibles à des utilisateurs non accrédités), le **Déni de service (**l'application est rendue indisponible) et l'**Élevation de privilege :** un utilisateur dispose un niveau d'accès à l'application supérieur à celui qui devrait lui être accessible.

Pour se prémunir de ces menaces, différentes techniques liées à la mise en œuvre de la sécurité sont envisageables :

— **Authentification :** identification des applications clientes d'un service Web. Les mécanismes d'authentification permettent de se prémunir contre l'usurpation d'identité (signature).
— **Autorisation :** définition des droits d'un client authentifié vis-à-vis d'une application. Les mécanismes de gestion des autorisations permettent d'éviter l'altération des données et la diffusion d'informations sensibles.
— **Identité :** il est parfois nécessaire de véhiculer le contexte de sécurité de l'utilisateur authentifié via de multiples ressources du système d'informa- tion.
— **Sécurité des communications :** il faut faire en sorte que les messages circulant sur le réseau restent privés (chiffrement) et non altérés (signature du message).
— **Audit :** enregistrement des actions de l'application cliente, qu'elles soient autorisées ou non, pour garantir la non-répudiation.

1.2.9 Une technologie maîtrisée mais non dénuée de risques

À côté de ces risques technologiques que l'on peut qualifier de », l'on se doit maintenant de prendre en compte des risques ayant une portée socio-économique. En effet, il faut bien se rendre à l'évidence que malgré les mécanismes sophistiqués de sécurisation des services Web, il est un critère fondamental qui est sous-entendu : votre fournisseur de services doit être de confiance ! Or depuis le mois de juin 2013 et les révélations d'Edward Snowden, le monde des services a basculé dans la méfiance.

Lorsque l'on apprend que les plus grands fournisseurs de services en ligne (Amazon, Google, Microsoft) ou de réseaux sociaux (Facebook, Twitter) ont une entrée dérobée directement accessible par l'agence de sécurité nationale américaine (NSA), et ceux, même à des fins de lutte contre le terrorisme (Patriot Act [Lee03]), il est normal d'en tenir compte dans une future analyse des risques.

1.3 La gestion des risques en sécurité de l'information

La norme ISO/IEC 27005:2011 [ISO11] s'intègre dans la suite des normes ISO 2700x, qui traitent de la sécurité de l'information. Celle-ci, à la différence de la norme ISO 27001 :2005 [ISO05b], ne fournit pas une liste d'exigences mais des recommandations pour la mise en œuvre d'un processus de gestion des risques en sécurité de l'information. Elle n'est donc pas un référentiel de certification, mais un guide de mise en œuvre. Par conséquent, elle donne des lignes directrices qu'il convient de suivre, d'adapter et parfois d'ignorer dans la conception, le développement et l'exploitation de son système de gestion de risques.

Elle s'inscrit dans la politique de la société BackPlan™ de mettre en avant la sécurité de l'information : c'est un gage de confiance que doivent ressentir les clients envers cette société. La démarche a été initiée par les certifications à la norme ISO 27001 des co-gérantes : Lead Auditor ISO 27 001 et Lead Implementor ISO 27 001.

1.3.1 Présentation

La norme ISO/IEC 27005:2011 décrit le processus de gestion des risques en sécurité de l'information.

La norme ISO 27005 explique en détails comment conduire l'appréciation des risques et le traitement des risques, dans le cadre de la sécurité de l'information.

La norme ISO 27005 demeure utilisable dans toutes les situations, de manière autonome, par exemple pour l'appréciations des risques dans un projet sans lien avec l'ISO 27001.

La norme ISO 27005 applique à la gestion de risques le cycle d'amélioration continue PDCA [5] [Dem50] utilisé dans les normes de systèmes de management, comme l'ISO 27001 en sécurité de l'information. Cela lui permet une souplesse et un pragmatisme pour être utilisée en toutes circonstances et notamment dans des entreprises où tout change sans arrêt. Elle constitue un guide qui s'adapte à tous types d'organismes et de situations.

1.3.2 ISO 27005 : une norme synthétique et consensuelle

La norme ISO/IEC 27005:2011 est un très rare exemple de véritable norme.

Une norme se doit d'être le consensus entre les points de vues des différents acteurs du marché et des différents pays. Elle se doit d'offrir des garanties aux utilisateurs, notamment aux consommateurs, et de permettre une rationalisation et une multiplication des échanges. Or en SSI, une grande majorité de normes ne servent à rien, sont de piètre qualité, et existent car

5. Plan-Do-Check-Act : Cycle de Demming

un consultant trouve du financement quand il est volontaire pour être rédacteur en chef d'une norme. La norme ISO 27005 est l'exemple même de la norme synthétique et consensuelle qui répond à ce que devrait être une norme. Elle reprend tout ce qui avait été fait sur le sujet.

L'ISO 27005 est principalement issue :

- De la norme internationale ISO 13335 qui a modélisé la gestion de risque en sécurité des systèmes d'information dès le lancement de la normalisation en SSI [6] en 1992.
- De la méthode EBIOS [7] [EBI10] de la DCSSI [8] en France, dont notamment le principe du découpage du processus d'appréciation des risques en activités et sous-activités élémentaires avec des entrées, un travail à accomplir et un résultat à obtenir, a été repris dans l'ISO 27005 ;
- De la norme britanique CRAMM [9] [CRA] ;
- La norme américaine OCTAVE [10] [AD01] [OCT]
- De la norme britannique BS7799-3 [BS795], qui avait été éditée par le BSI dans la lignée de la norme BS7799-2 qui est devenue l'ISO 27001.

La norme ISO 27005 présente la gestion de risque en utilisant un vocabulaire courant, que l'on retrouve dans les autres métiers, ce qui facilite sa compréhension et sa lecture. La norme ISO 27005 applique les critères de risques (appelés aussi critères de sécurité) : confidentialité, intégrité, disponibilité, sur l'ensemble du patrimoine informationnel de l'organisme, selon :

- les objectifs de sécurité du processus métier,
- les obligations légales,
- les contraintes réglementaires,
- les aspects financiers et opérationnels,
- la technologie et les facteurs sociaux et humains.

Elle va ainsi plus loin que les méthodes basées sur les menaces et vulnérabilités visant les ressources informatiques.

L'ISO 27005 met les ressources informatiques comme actifs de soutien aux actifs primordiaux que sont l'information et les métiers de son organisme.

La norme ISO 27005 est fondamentale car trop d'organismes ont pris pour argent comptant la norme ISO 27002 (anciennement ISO 17799 [ISO05a]), malheureusement disponible avant l'ISO 27001, qui définit le processus par lequel la SSI est gérée dans le temps. Ainsi, de nombreux organismes utilisent l'ISO 27002 dans une approche conformité. Cependant, développer une approche conformité par rapport à l'ISO 27002 en SSI donne des résultats catastrophiques. Elle

6. Sécurité des Systèmes d'Information
7. EBIOS : Expression des Besoins et Identification des Objectifs de Sécurité
8. DCSSI : Direction centrale de la sécurité des systèmes d'information - http ://www.ssi.gouv.fr
9. **C**CTA(Central Computer and Telecommunications Agency) **R**isk **A**nalysis and **M**anagement **M**ethod
10. **O**perationally **C**ritical **T**hreat, **A**sset, and **V**ulnerability **E**valuation

engendre des actions inutiles et coûteuses, vous fait vous mettre à dos les services de production informatique qui doivent être vos alliés, oblige les gens à prendre l'habitude de mentir en cochant des cases alors que les actions ne sont pas faites et dégoûte l'ensemble des utilisateurs de la sécurité. Il n'y a que les auditeurs peu compétents qui y trouvent leur compte.

1.3.3 Le risque : au cœur du processus décrit par cette norme

Le risque est la vraisemblance (plus ou moins grande) de voir un événement ciblant un actif et exploitant une ou plusieurs vulnérabilités se réaliser en entraînant des impacts sur un ou plusieurs actifs et des conséquences pour l'organisme.

Définition du risque : « Un risque est une combinaison de la probabilité d'un dommage et de sa gravité » [11]

Définition du risque en sécurité de l'information : « Possibilité qu'une menace donnée exploite les vulnérabilités d'un actif ou d'un groupe d'actifs et nuise donc à l'organisation. Note : le risque est mesuré en termes de combinaison entre la vraisemblance d'un événement et ses conséquences. » [12]

Le risque est la possibilité qu'une ou plusieurs menaces exploitant une ou plusieurs vulnérabilités d'un ou plusieurs actifs entraînent des con- séquences pour l'actif, un groupe d'actifs, un processus voir l'organisme comme le représente le schéma ci-après.

1.3.4 Une approche par la gestion de risque

La norme ISO/IEC 27005:2011 est un guide complet et autonome de mise en œuvre de la gestion des risques de la sécurité de l'information. Or seule une approche de ce type permet de justifier ce qui est fait ou non dans le déploiement des dispositifs de sécurité. L'ISO 27005 utilise une évaluation des risques par scénarios, mais pas une analyse des risques par scénarios, cependant elle recommande l'utilisation de scénarios pour expliquer et justifier dès l'identification des risques. Avant de parler de lanorme ISO 27002, parlez d'abord de l'ISO 27005. Elle vous permettra de choisir, en amont, les mesures de sécurité appropriées de l'ISO 27002, là où c'est justifié.

La norme ISO 27005 détaille le processus de gestion de risque dans les chapitres 6 à 12.

Elle est complétée de 6 annexes de référence A à F, nécessaires à la mise en oeuvre de la méthode.

Ces annexes permettent d'alléger le corps de la norme tout en permettant d'avoir les approfondissements, les explications détaillées, les tableaux et les listes.

11. Source : Guide ISO/CEl 57 7999
12. Référence normative : ISO/IEC 27005:2011

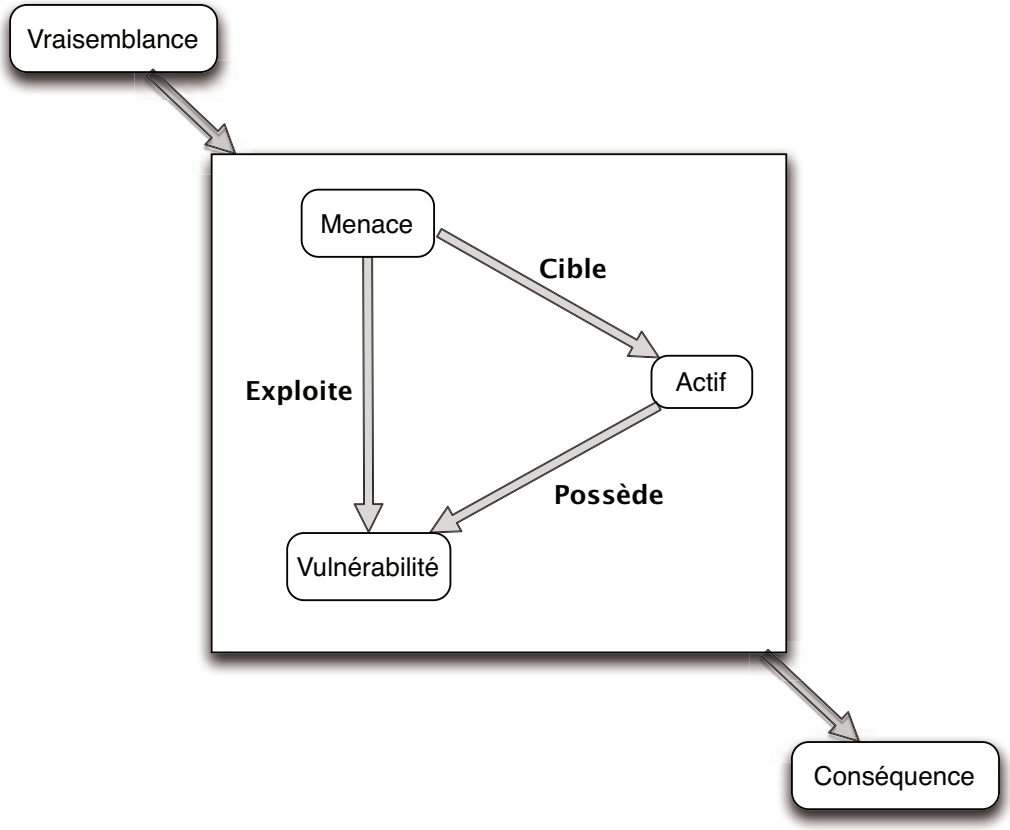


FIGURE 1.4 – Modélisation du risque

— Le chapitre 6 explique le processus de gestion de risque dans son ensemble et la manière dont il se positionne dans un cycle PDCA.
— Le chapitre 7 précise l'établissement du contexte, afin que l'on spécifie son périmètre de l'appréciation des risques, que l'on établisse tous les critères de base, que l'on décrive son environnement, et organise ses processus.
— Le chapitre 8 définit l'appréciation des risques, qui se décompose en l'analyse des risques et l'évaluation des risques.
— Le paragraphe 8.2 définit l'analyse de risques, qui consiste, d'une part, à identifier et valoriser ses actifs sensibles, identifier les menaces et les vulnérabilités, ainsi que les mesures de sécurité existantes, et, d'autre part, à estimer la probabilité d'occurrence des risques identifiés, quantifier les conséquences potentielles, au final, calculer le risque. L'estimation des risques peut être qualitative ou/et quantitative. La norme n'impose aucune méthode de calcul particulière et en propose 3 en annexe.
— Le paragraphe 8.3 décrit l'évaluation des risques afin de prioritiser et ordonnancer les risques par rapport aux critères d'évaluation des risques. Ces critères préalablement établis par les objectifs de sécurité imposés par les métiers de l'organisme et identifiés

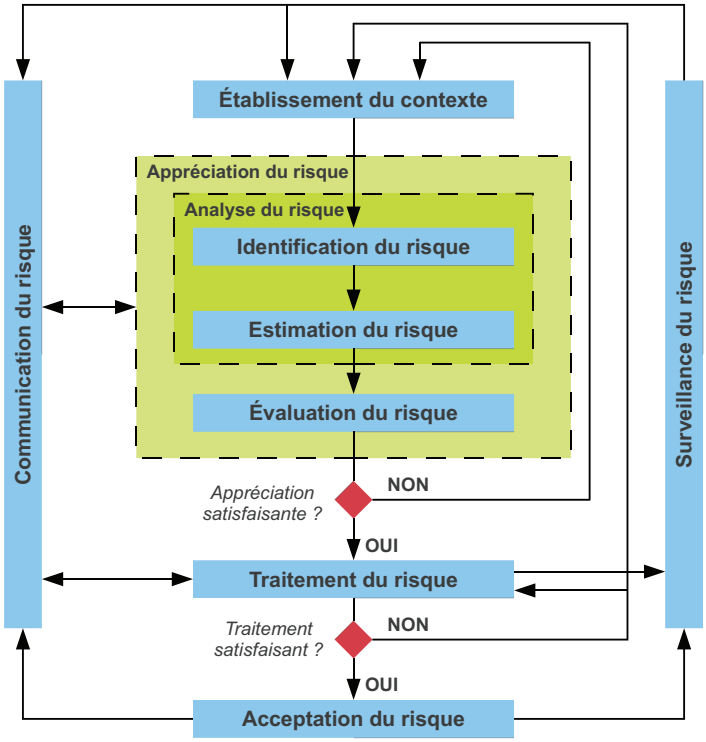


FIGURE 1.5 – Processus de gestion du risque en sécurité de l'information

par les parties prenantes permettent de déterminer le seuil au-delà duquel le risque devra être traité.

— Le chapitre 9 spécifie le traitement du risque. Les quatre choix du traitement du risque sont :
  — **Le refus ou évitement :** le risque est trop élevé, il n'y a pas de mesure de sécurité réaliste pour le réduire, l'activité est supprimée ;
  — **Le transfert** vers un assureur ou un sous-traitant qui saura mieux le gérer ;
  — **La réduction** par l'application des mesures de sécurité ;
  — Et **la prise de risque** : le risque est accepté tel quel sans qu'aucune action soit prise. A ce stade, la norme dit d'intégrer les coûts notamment dans le choix des mesures de sécurité.
— Le chapitre 10 détermine l'acceptation du risque. Il faut calculer le risque résiduel qui sera obtenu une fois que le traitement du risque sera mis en oeuvre. La direction générale doit accepter les risques résiduels, donc accepter le plan de traitement du risque dans son ensemble. Cette décision est formellement documentée et enregistrée, et si par exemple des contraintes de budget ou de temps ne permettent pas à la direction de déployer, c'est elle qui en prend la responsabilité, pas la RSSI.

— Le chapitre 11 décrit la communication du risque. Cette communication est un partage régulier d'informations entre le gestionnaire des risques SSI (en général, le RSSI), les décisionnaires, et les parties prenantes concernant la gestion du risque. Ses buts sont de donner confiance à la direction générale et aux parties prenantes, collectionner les informations concernant les risques encourus, faire connaître les plans de traitement du risque, obtenir le support et les moyens pour la mise en oeuvre du traitement du risque, impliquer la responsabilité des décisionnaires, améliorer les compétences de gestion du risque, et sensibiliser l'organisme à la prévention du risque.
— Le chapitre 12 décrit la surveillance et le réexamen des risques. La surveillance constante du processus de gestion des risques est nécessaire pour s'assurer que le processus reste pertinent et adapté aux objectifs de sécurité des métiers de l'organisme, que chaque risque traité n'est pas surestimé ou sous-estimé, et que ses coûts de gestion sont adaptés à la dimension du risque et aux besoins de sécurité. Il faut aussi identifier les changements nécessitant une réévaluation du risque ainsi que les nouvelles menaces et vulnérabilités.
— L'annexe A liste toutes les contraintes qui peuvent affecter votre processus de gestion des risques.
— L'annexe B aide à identifier et valoriser les actifs à considérer dans son appréciation des risques, et aide à estimer les impacts.
— L'annexe C répertorie les menaces classées par type : dommages physiques, pertes de service essentiel, altération d'informations, etc. La liste n'est pas exhaustive mais complète.
— L'annexe D liste les vulnérabilités, les moyens de recherche de vulnérabilités et des exemples de menaces qui pourraient exploiter ces vulnérabilités.
— L'annexe E décrit trois méthodes de calcul de risque, qui permettent de réaliser l'appréciation des risques, dans l'activité estimation des risques. Les méthodes estiment l'impact potentiel d'un risque par rapport à la valeur de l'actif, à la facilité d'exploitation des vulnérabilités par les menaces, à la probabilité d'occurrence, etc.
— L'annexe F précise toutes les contraintes à intégrer lors de la réduction des risques, notamment lorsque l'on sélectionne des dispositifs de sécurité.

1.3.5 La gestion de risque dans l'ISO 27005 est un processus continu et itératif.

Un RSSI [13] qui démarre son appréciation à partir d'une feuille blanche ne commence pas par décider des échelles alors qu'il ne maîtrise pas encore tous les principes.

Les critères de valorisation des actifs, d'évaluation des risques, éventuellement les critères

13. Responsable de la sécurité des systèmes d'information

d'impact, et les critères d'acceptation des risques, sont explicités dans l'établissement du contexte (chapitre 6) de la norme mais il sont construits au fur et à mesure des processus itératifs et pas séquentiellement au début. Cette approche itérative améliore la finesse de l'analyse à chaque itération, fournit une bonne répartition entre le temps et l'effort fourni pour identifier les mesures de sécurité, permet de traiter les risques en fonction des ressources et moyens disponibles, facilite les liens entre les risques et les conséquences sur les métiers, permet d'avancer lorsque les interlocuteurs sont absents, en facilitant la gestion des susceptibilités et des aspects politiques entre interviewés. L'approche itérative de l'ISO/IEC 27005 permet à la gestion de risques de tendre progressivement vers une maîtrise des risques de haut niveau et conforme aux besoins de l'organisme.

L'ISO/IEC 27005 fait entrer la gestion des risques pour la SSI dans la gestion de risques en général. Toute direction générale doit mettre en œeuvre une gestion globale des risques pour atteindre les objectifs de l'organisme.

Avec l'ISO/IEC 27005, la gestion des risques en sécurité de l'information se rapproche des risques opérationnels, industriels, financiers, etc. Cette nouvelle norme est une méthode complète de la gestion du risque de la sécurité de l'information. Elle est cohérente avec les autres normes ISO de la famille 2700x. Elle permet une gestion simple, pragmatique et adaptée aux besoins de sécurité de tous les métiers ; il n'y a pas de processus linéaire à suivre et peu d'obligations dans le formalisme. Elle est neutre sur les méthodes d'évaluation qualitative et quantitative. Cette norme est supportée par de nombreux produits commerciaux, déjà plus d'une dizaine sur le papier, cependant le tableur demeure l'outil le plus performant et le plus utilisé

1.4 Synthèse

La problématique initiale qui se présentait à nous était la notion de sécurité des services dans les architectures distribuées. En effet, par essence même, une architecture orientée services est ouverte sur l'extérieur. On ne peut donc pas se retrancher derrière un mur, une enceinte, une protection pour protéger son système. Il faut donc implémenter des mécanismes afin de garantir une sûreté maximale des échanges entre les briques de l'infrastructure.

Rapidement on se rend compte, lors de l'utilisation de services externes, que l'on ne peut pas tout maîtriser ! En effet ces services sont générés par des fournisseurs qui vous donnent des recommandations techniques pour vous permettre de pouvoir les exploiter, les mécanismes de sécurité sont mis en place et normalement dans le meilleur des mondes notre infrastructure est sure.

Malheureusement, il faut également envisager la défaillance de ces fournisseurs de services (panne, faillite, attaques, etc.). Dans une infrastructure on ne peut se permettre qu'un service

soit défaillant. D'où la mise en place d'une seconde brique pour ce service (redondance) afin de garantir son bon fonctionnement. Bien entendu il n'est pas question de doubler toutes les mesures de protection, et c'est là qu'intervient la notion de risque. Dans la gestion du risque il faut nécessairement bâtir des scénarios impliquant tous les aspects possibles d'une défaillance. L'étude de de la norme ISO/IEC 27005:2011 est parfaite lorsque l'on envisage le côté purement technique et procédural d'une gestion des risques en sécurité de l'information dans une entreprise. Cette norme n'a pas encore prévu les risques économiques, par exemple la faillite d'un fournisseur de services, les risques législatifs où une loi pour favoriser la consultation de données sur des serveurs installés sur son territoire.

Quand on envisage le problème sous l'angle de la sécurité de l'information, c'est-à-dire en terme macroscopique par rapport à la notion de sécurité des donnés, de nouvelles pistes s'offrent à nous : les métadonnées en sont une. La prise en compte de ces métadonnées dans les échanges revêt une nouvelle approche du traitement de l'information. Bien souvent non sécurisées, collectées à l'insu de leurs propriétaires, ces informations initialement nécessaire à la bonne gestion des échanges, devient un risque lorsque l'on se situe au niveau de la sécurité de l'information. Il nous a paru nécessaire à ce niveau-là d'envisager les aspects juridiques à l'utilisation de ces métadonnées : a-t-on le droit d'exploiter ces données à d'autres titres que les raisons initiales de gestion du système ?

Nous le voyons, nous sommes passés d'une problématique purement dédiée à l'étude de la sécurisation des services Web à la a prise en compte de la sécurité de l'information dans des infrastructures qui exploitent "services" et "métadonnées".

Forts de ce constat, nous avons donc orienté notre problématique de la sécurité des systèmes informatiques vers la sécurité de l'information.

Chapitre 2

SARSSI 2007
Vers l'interopérabilité des systèmes de DRM

Sommaire

2.1 Présentation

Cet article a été présenté lors de la seconde conférence jointe SAR [14] - SSI [15] organisée à Annecy du 12 aux 15 juin 2007. Cette conférence française a pour vocation de présenter les recherches, les résultats, cas pratiques et toutes idées innovantes sur le thème de la sécurité des réseaux et des systèmes d'information.

2.2 Introduction

La distribution de contenus numériques via Internet offre de nouveaux services aux consommateurs et fournit de nouvelles opportunités commerciales aux artistes et créateurs.

Cependant le succès de ces nouveaux canaux de distribution dépend de l'effectivité des mécanismes de protection des intérêts des différents contributeurs de la chaine. Ces mécanismes sont assurés par les systèmes de gestion des droits numériques ou systèmes DRM (Digital Right Management). Les systèmes DRM ont pour but de contrôler et de protéger la propriété intellectuelle des contenus numériques : ainsi ils restreignent au travers de licences d'utilisation ce que l'utilisateur peut faire du contenu qu'il a acheté.

A ce jour plusieurs systèmes de gestion de droits numériques existent sur le marché. La majorité de ces systèmes utilise une solution de type logicielle (ex : Itunes) bien qu'à côté Il existe également des solutions de type matériel (DVD-CSS : Content Scramble System).

Dans cet article nous nous intéressons aux solutions logicielles.

Chacun de ces systèmes emploie une technique propriétaire de protection de contenus. Les licences délivrées par un système ne sont comprises que par lui seul. Cette situation a conduit a la mise sur le marché de produits qui ont pour but de faciliter l'interopérabilité des systèmes existants. Nous pouvons prendre comme exemples les Systèmes DMDfusion et OPERA. DMDfusion incorpore les systèmes Microsoft, Adobe et Real Networks. Le Système OPERA quant a lui définit une architecture pour l'interopérabilité des systèmes DRM. Cependant cette architecture ne permet pas d'échanges entre systèmes de DRM puisque les licences sont éditées pour un système particulier et ne sont pas interprétées par les autres systèmes.

Sur le plan législatif le besoin d'interopérabilité se fait aussi ressentir. C'est dans ce sens que la loi Francaise n 2006-961 du 1er août 2006, dite loi DADVSI, relative au droit d'auteur et aux droits voisins dans la société de l'information dispose a son article L. 331-5 que : « les mesures techniques ne doivent pas avoir pour effet d'empêcher la mise en œuvre effective de l'interopérabilité, dans le respect du droit d'auteur ».

14. 6ème Conférence sur la Sécurité des Architectures Réseaux - SAR
15. 4ème Conférence sur la Sécurité des Systèmes d'Information - SSI

Les fournisseurs de mesures techniques donnent l'accès aux informations essentielles a l'interopérabilité.

2.3 Les thèmes abordés

Le but de cet article est d'identifier les obstacles a l'interopérabilité des systèmes DRM et de proposer des solutions alternatives. Nos propositions sont destinées a faciliter l'interopérabilité des différents composants qui interviennent dans l'architecture DRM. Il s'agit pour nous d'exploiter les standards existants pour permettre a ces composants de communiquer entre eux afin de rendre de facon efficace et sécurisée le service de DRM. Concrètement l'interopérabilité des systèmes de DRM dépend :

— du langage d'expression de droits (REL - Rights Expression Language) utilisé dans chacun des systèmes. En effet, la licence d'utilisation est éditée dans un langage et n'est comprise que par les systèmes qui utilisent le même langage.

— Des protocoles de communication entre les composants du système. Les systèmes ne peuvent pas communiquer si chacun utilise des protocoles propriétaires et compréhensibles que par lui seul.

— Du mécanisme de protection du contenu. Chaque système doit véhiculer les informations nécessaires au rendu du contenu sécurisé par l'utilisateur ayant acquis une licence. Par exemple, le logiciel client doit connaitre l'algorithme utilisé pour protéger le contenu afin de le décrypter.

La section 2 présente les notions préliminaires dont nous aurons besoin pour le développement de nos propositions :

— le protocole SOAP,

— les services Web

La section 3 expose les principaux acteurs d'un système DRM basique : son architecture de sécurité, ainsi que le processus d'utilisation des ressources protégées. Nous présentons l'état de l'art a la section 6.

2.4 Les propositions

Les principales contributions de ce papier sont les suivantes :

— Nous montrons l'intérêt de l'utilisation des langages ODRL et MPEG-21 REL pour l'édition des licences (section 4).

— Nous tirons parti du standard OASIS Web Services Security : SOAP Message security pour l'échange de messages entre les acteurs d'une architecture DRM (section 5). Nous

proposons alors une interface de services web (Web Services) pour chaque composant du système au travers des squelettes de messages émis et attendus. Chaque élément de l'ensemble peut retrouver un service dans le système a partir de la définition d'interface du composant qui le fournit. Les différents composants communiquent par échanges de messages SOAP (Service Oriented Architecture Protocol).

— Nous exploitons la propriété d'extensibilité de XML (eXtensible Markup Language), pour convoyer avec le contenu sécurisé les informations nécessaires a son rendu (section 5).

2.5 Conclusions

Le manque d'interopérabilité entre les différents systèmes DRM a souvent été avancé pour expliquer l'échec des DRMs notamment dans le cadre de la distribution de contenu musical. Dans ce papier, nous identifions les obstacles a ce besoin d'interopérabilité et proposons des solutions pour les contourner. Ainsi, nous utilisons les concepts véhiculés par les services web pour définir l'architecture générale d'un système DRM interopérable. Les différents acteurs de l'architecture communiquent par échanges de messages SOAP.

Nous avons établi qu'il était difficile de faire le choix entre les deux langages d'expression de droits les plus utilisés de la littérature que sont ODRL et MPEG-21 REL. Nous envisageons alors la définition d'un langage générique qui engloberait ODRL et MPEG-21 REL. Autre perspective de ce travail consiste a définir une ontologie des DRMs afin de spécifier de facon standard les interfaces des différents acteurs de l'architecture proposée.

2.6 Texte de l'article

Vers l'interopérabilité des Systèmes de DRM (Digital Rights Management)[*]

Majirus Fansi Vincent Lalanne Alban Gabillon
Université de Pau et des Pays de l'Adour.IUT des Pays de l'Adour, 371 Rue du Ruisseau, 40004 Mont de Marsan, France.
E-Mails : janvier-majirus.fansi@etud.univ-pau.fr, { vincent.lalanne, alban.gabillon} @univ-pau.fr

Un des principaux inconvénients des systèmes de gestion des droits numériques actuels est le manque d'interopérabilité entre ces systèmes. Dans cet article, nous identifions les obstacles à l'interopérabilité des solutions DRM et proposons des solutions alternatives. Ainsi, nous préconisons l'usage des langages d'expression de droits ODRL et MPEG-21 REL. Aussi, nous proposons une interface de service web pour chaque acteur de l'architecture DRM afin de faciliter les échanges entre différents systèmes. Les différents composants du système DRM communiquent alors par échanges de messages SOAP. Nous tirons parti de la propriété d'extensibilité de XML pour véhiculer avec le contenu protégé, les informations nécessaires à son rendu par tout logiciel client sous réserve que l'utilisateur détienne une licence.

Mots Clés: DRM, XML, XrML, MPEG-REL, ODRL, Web Services, Contrôle d'Usage, Sécurité.

1. Introduction

La distribution de contenus numériques via Internet offre de nouveaux services aux consommateurs et fournit de nouvelles opportunités commerciales aux créateurs de contenu. Cependant le succès de ces nouveaux canaux de distribution dépend de l'effectivité des mécanismes de protection des intérêts des différents intervenants de la chaîne. Ces mécanismes sont assurés par les systèmes de gestion des droits numériques ou systèmes DRM (Digital Right Management). Les systèmes DRM ont donc pour but de contrôler et de protéger la propriété intellectuelle des contenus numériques tels que les documents, les images, les vidéos et les sons. Ainsi ils restreignent au travers de licences d'utilisation, ce que l'utilisateur peut faire du contenu qu'il a acheté. Une licence spécifie un règlement énonçant les conditions d'utilisation d'un contenu numérique. A ce jour plusieurs systèmes de gestion de droits numériques existent sur le marché. La majorité de ces systèmes, comme `Apple iTunes`, utilise une solution de type logiciel. Il existe aussi des solutions de type matériel telles que le système `DVD-CSS` (`Content Scramble System`).

Dans cet article nous nous intéressons aux solutions logicielles. Chacun de ces systèmes emploie une technique propriétaire de protection de contenus. Les licences délivrées par un système ne sont interprétables que par lui seul. Cette situation a conduit à la mise sur le marché de produits qui ont pour but de faciliter l'interopérabilité des systèmes existants. Nous pouvons prendre comme exemples les Systèmes DMDfusion et OPERA. DMDfusion incorpore les systèmes Microsoft, Adobe et Real Networks. Le Système OPERA quant à lui définit une architecture pour l'interopérabilité des systèmes DRM. Cependant cette architecture ne permet pas d'échanges entre systèmes de DRM puisque les licences sont éditées pour un système particulier et ne sont pas interprétables par les autres systèmes. Sur le plan législatif le besoin d'interopérabilité se fait aussi ressentir. C'est dans ce sens que la loi Française n° 2006-961 du 1er août 2006, dite loi DADVSI, relative au droit d'auteur et aux droits voisins dans la société de l'information dispose à son article L. 331-5 que: les mesures

[*] Réalisé dans le cadre du projet CASC (ACI Sécurité Informatique 2003-2006), avec le soutien du Conseil Général des Landes

techniques ne doivent pas avoir pour effet d'empêcher la mise en œuvre effective de l'interopérabilité, dans le respect du droit d'auteur. Les fournisseurs de mesures techniques donnent l'accès aux informations essentielles à l'interopérabilité.

Le but de cet article est d'identifier les obstacles à l'interopérabilité des systèmes DRM et de proposer des solutions alternatives. Nos propositions sont destinées à faciliter l'interopérabilité des différents composants qui interviennent dans l'architecture DRM. Il s'agit pour nous d'exploiter les standards existants pour permettre à ces composants de communiquer entre eux afin de rendre de façon efficace et sécurisée le service de DRM.

Concrètement l'interopérabilité des systèmes de DRM dépend:

- du langage d'expression de droits (REL - Rights Expression Language) utilisé dans chacun des systèmes. En effet, la licence d'utilisation est éditée dans un langage et n'est interprétable que par les systèmes qui utilisent le même langage.
- des protocoles de communication entre les composants du système. Les systèmes ne peuvent pas communiquer si chacun utilise des protocoles propriétaires et compréhensibles que par lui seul.
- du mécanisme de protection du contenu. Chaque système doit véhiculer les informations nécessaires au rendu du contenu sécurisé par l'utilisateur ayant acquis une licence. Par exemple, le logiciel client doit connaître l'algorithme utilisé pour protéger le contenu afin de le décrypter.

Les principales contributions de ce papier sont les suivantes:

- Nous montrons l'intérêt de l'utilisation des langages ODRL [4] et MPEG-21 REL [5] pour l'édition des licences (section 4).
- Nous tirons parti du standard OASIS Web Services Security: SOAP Message security pour l'échange de messages entre les acteurs d'une architecture DRM (section 5). Nous proposons alors une interface de services web (Web Services) pour chaque composant du système au travers des squelettes de messages émis et attendus. Chaque élément de l'ensemble peut retrouver un service dans le système à partir de la définition d'interface du composant qui le fournit. Les différents composants communiquent par échanges de messages SOAP (Service Oriented Architecture Protocol).
- Nous exploitons la propriété d'extensibilité de XML (eXtensible Markup Language), pour convoyer avec le contenu sécurisé les informations nécessaires à son rendu (section 5).

La section 2 présente les notions préliminaires dont nous aurons besoin. La section 3 expose les principaux acteurs d'un système DRM basique, son architecture de sécurité, ainsi que le processus d'utilisation des ressources protégées. Nous présentons l'état de l'art à la section 6. Finalement la section 7 conclut ce papier.

2. Notions préliminaires

Dans cette section nous présentons les notions qui seront employées tout au long de cet article. Nous nous intéressons en particulier aux recommandations W3C (World Wide Web Consortium) SOAP (Service Oriented Architecture Protocol) et Service Web (Web service).

2.1. Service Oriented Architecture Protocol (SOAP)[6]

SOAP est un protocole de communication basé sur XML (eXtensible Markup Language) qui permet aux applications d'échanger des informations à travers Internet. Avant l'avènement de SOAP les applications communiquaient essentiellement par des appels de procédures à distance RPC (Remote Procedure Call) entre objets telles que DCOM (Distributed Component Object Model), CORBA (Common Object Request Broker Architecture), Java RMI (Remote Method Invocation). Cependant les RPCs posent un problème de sécurité puisque ces types de trafics sont souvent bloqués par les pare feux (firewall) et les serveurs proxy. Un meilleur moyen de communication entre applications consiste à utiliser http (Hyper Text Transfert Protocol). SOAP a été créé pour permettre cela. Il permet aux applications installées sur différents systèmes d'exploitation (SE), avec des

technologies et langages de programmation différents de communiquer en s'échangeant des messages SOAP (codés en XML) et en utilisant le protocole http. Un message SOAP est un document XML ordinaire contenant les éléments suivants:

- Un élément enveloppe (Envelope) obligatoire qui identifie le document XML comme un message SOAP
- Une entête optionnelle (Head) qui contient des informations liées au message. Cet élément est généralement utilisé pour convoyer les informations de sécurité liées au message.
- Un élément corps du message (Body) qui contient les informations d'appel et de réponse des procédures distantes.
- Optionnellement un élément Fault qui fournit les informations à propos des éventuelles erreurs survenues lors du traitement du message.

La figure 1 (a) montre le squelette d'un message SOAP; pour des raisons de simplicité, nous ne faisons pas apparaître les définitions d'espaces nominaux [7].

2.2. Service Web (Web service)

Un service web[2] est un système logiciel conçu pour permettre l'interopérabilité des interactions entre machines à travers le réseau. Il a une interface décrite dans un format compréhensible par la machine (en l'occurrence WSDL – Web Service Description Language). Les autres systèmes communiquent avec le service web conformément à sa description en utilisant les messages SOAP typiquement convoyés via http. Un service web est une notion abstraite qui doit être réalisé par des agents concrets. Un agent est le dispositif concret (logiciel ou matériel) qui envoie et reçoit les messages, alors que le service est la ressource caractérisée par l'ensemble abstrait des fonctionnalités qui est fourni. Cette distinction s'illustre par le fait qu'un même service web peut être implémenté dans plusieurs langages de programmation différents. Chaque implémentation représente alors un agent, le service rendu restant le même.

Hormis SOAP, les deux autres technologies qui sous-tendent les services web sont UDDI (Universal, Description, Discovery and Integration) et WSDL.

Un document WSDL est au format XML. Il décrit un service web. Il définit la syntaxe et la sémantique des messages et spécifie les méthodes que le service expose, ainsi que les emplacements réseaux où les agents fournisseurs[3] peuvent être invoqués.

UDDI est un service de répertoire basé sur XML qui permet à des organisations de publier et de retrouver des services web. UDDI est interrogé par des messages SOAP et fournit l'accès aux documents WSDL.

Un exemple illustratif de ces concepts est le scénario suivant: une organisation met à disposition un standard UDDI pour la distribution de contenus numériques sécurisés. Les entreprises spécialisées dans la protection de contenu (packager) publient leurs services dans le répertoire. Les compagnies de distribution de contenus sécurisés (la FNAC par exemple) peuvent alors chercher dans le répertoire l'interface (document WSDL) des packagers. Quand l'interface est trouvée, le distributeur communique avec le service immédiatement et reçoit les ressources souhaitées.

Les services web se sont imposés comme solution effective pour l'interopérabilité entre les machines du réseau. Cependant il s'est posé le problème de confidentialité et d'intégrité des messages échangés. Le standard OASIS Web Services Security (WSS)[8] a été proposé pour répondre à ce besoin. WSS étend les messages SOAP afin d'assurer l'intégrité et la confidentialité des messages échangés. L'extension consiste à ajouter les informations de sécurité dans l'entête d'un message SOAP classique.

L'intégrité est assurée par la recommandation XML Signature [9]. XML signature est une recommandation du W3C destinée à permettre l'utilisation de signatures numériques dans les

[2] www.w3.org/2002/ws/

[3] Un agent fournisseur d'un service est un ensemble logiciel qui rend ce service

documents XML. Une caractéristique importante apportée par XML est la possibilité de ne signer que des portions spécifiques d'un même document.

```
<?xml version="1.0"?>
<soap:envelope>
 <soap:header>
    ….
 </soap:header>
 <saop:body>
    ….
  <soap:fault>
    ….
  </soap:fault>
 </soap:body>
</soap:envelope>
```

(a)

```
<?xml version="1.0"?>
<soap:envelope>
 <soap:header>
   <wss:security>
     <encryptedKey>...</encryptedKey>
     <signature> …  </signature>
   </wss:security>
 </soap:header>
 <saop:body>
   <encryptedData>...</encryptedData>
    ….
  <soap:fault>
    ….
  </soap:fault>
 </soap:body>
</soap:envelope>
```

(b)

Figure 1 – (a) Squelette des messages SOAP (b) Squelette SOAP + Sécurité

La confidentialité des données transmises est assurée par la recommandation `XML Encryption` [10]. `XML Encryption` est une recommandation W3C qui permet de chiffrer les données et de représenter le résultat du chiffrement en XML. Les données à chiffrer peuvent être de toute sorte (y compris un document, une chanson, une vidéo).

La figure 1 (b) montre un message `SOAP` étendu. L'élément `encryptedData` contient les données chiffrées ainsi que l'algorithme utilisé pour le chiffrement. Il s'agit de cryptographie à clé sécrète. La clé utilisée pour le chiffrement est à son tour chiffrée par la clé publique du destinataire. Le résultat est convoyé par l'élément `encryptedKey` qui renseigne aussi sur l'algorithme utilisé pour protéger la clé sécrète. Enfin l'élément `signature` contient la signature numérique de l'expéditeur. Il renseigne sur la portion du message concernée par la signature, la clé à utiliser pour la valider et sur l'algorithme utilisé.

3. Structure des systèmes DRM

Cette section présente la structure basique des systèmes DRM. Elle montre comment les différents acteurs sont agencés pour accéder aux ressources protégées et définit les types de contrôleurs DRM.

3.1. Principaux Acteurs des systèmes DRM

Bartolini et al. [1] définissent l'ensemble des besoins des systèmes de gestion de contenu en terme d'ensemble d'acteurs. Chacun de ces acteurs jouent différents rôles dans le système DRM. Les rôles décrits par Bartolini et al. sont:

1. l'`auteur` (ou le `créateur`) est responsable de la création du contenu.
2. le `détenteur de droit` est celui qui détient le copyright du contenu. Il est référencé dans la majorité des systèmes DRM comme le `content owner`. L'auteur n'est pas nécessairement le détenteur du copyright.
3. le `producteur` est l'entité responsable de la protection de la ressource. Il est encore appelé `packager` dans les systèmes DRM.

Interopérabilité des systèmes DRM

4. le `distributeur` est l'entité responsable de la distribution du contenu protégé. Il correspond à la composante `content distribution` des systèmes DRM.
5. le `registre IPR` (Intellectual Property Rights) est l'entité qui délivre des licences aux utilisateurs. Il est généralement appelé `license server` dans les systèmes DRM.
6. le `générateur d'identifiant` ou `UNI` (Unique Number Issuer) génère un identifiant unique pour chaque création. Ce service est rendu par le `packager` dans les systèmes DRM.

Ces acteurs constituent avec le `client DRM` l'architecture basique des systèmes DRM que nous représentons à la figure 2.

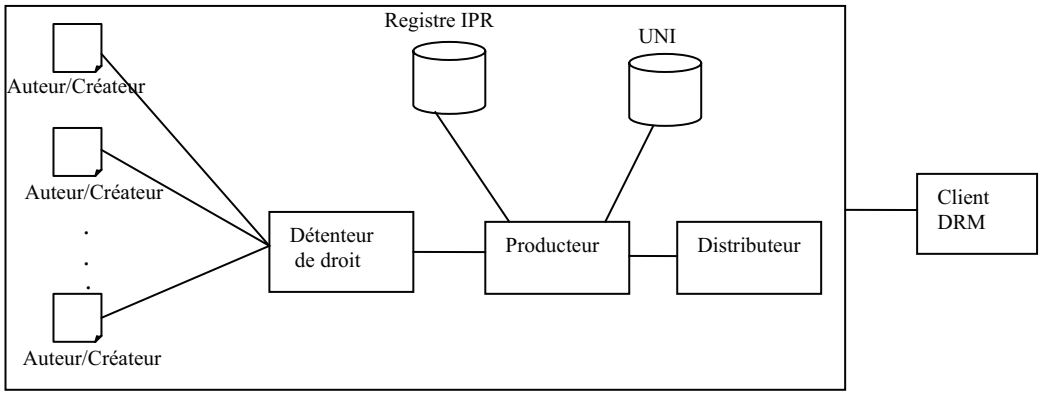


Figure 2 – Acteurs dans un Système DRM

3.2. Architectures de sécurité

Park et al.[2] énoncent trois facteurs qui distinguent les différentes architectures de sécurité impliquées dans la distribution des contenus sécurisés: la présence de la machine virtuelle, l'ensemble des règles de contrôle et le mode de distribution.

Le premier niveau de distinction est la présence de la machine virtuelle. Elle est décrite par Park et al.[2] comme étant "un logiciel qui tourne au dessus d'un environnement informatique vulnérable et emploie des fonctions de contrôle pour fournir les moyens de protéger et de gérer l'accès et l'usage de la ressource numérique". La machine virtuelle est généralement incorporée dans les systèmes DRM sous forme de `plugin`. Les clients qui ne possèdent pas une telle machine virtuelle ne peuvent pas gérer et contrôler l'accès et l'usage aux données sécurisées. Aujourd'hui, chaque système DRM possède sa propre machine virtuelle propriétaire. Par exemple, les morceaux protégés par le système `Apple FairPlay` ne peuvent être écoutés que via `iTunes music player`. De même ceux protégés par `Windows DRM` ne peuvent être écoutés que via la machine virtuelle `Windows Media Rights Manager` pluggée au logiciel `Windows Media Player`.

Le second niveau de distinction concerne l'ensemble des règles de contrôle. Les règles de contrôle d'accès et d'usage de contenus numériques sont exprimées par les langages d'expression de droits (REL). Ces règles ainsi exprimées représentent la licence d'utilisation. Park et al. distinguent trois façons de distribuer la licence:

- licence prédéfinie ou fixe: la machine virtuelle est équipée de la licence d'utilisation pour toutes les ressources qu'elle est censée contrôler. Ce mode de distribution est facile à implanter mais présente l'inconvénient d'être très peu flexible. Le système de chiffrement pour DVDs (DVD-CSS) est un exemple de contrôle par licence prédéfinie.
- Licence encastrée: la licence est encastrée dans la ressource protégée. Ceci peut être fait en insérant la ressource et la licence dans une enveloppe de sécurité.
- Licence séparée: la licence et la ressource sont distribuées séparément.

Plusieurs systèmes DRM existant utilisent les deux derniers modes de distribution de licences. Par exemple, `Apple iTunes Music Store` utilise les licences encastrées (`Apple FairPlay`) combinées à une licence fixe basique incorporée dans `iTunes music player`.

Le troisième et dernier niveau de distinction repose sur le mode de distribution du contenu protégé. En effet, l'utilisateur peut obtenir la ressource par un canal direct tel que la messagerie électronique. Il peut aussi le recevoir par téléchargement via un serveur de distribution.

3.3. Utilisation des ressources protégées

La figure 3 montre comment Erickson [3] décrit l'usage des contenus numériques protégés dans la majorité des systèmes DRM. Ces systèmes utilisent pour la plupart des licences séparées. Le contenu est d'abord protégé par le `packager` et distribué aux utilisateurs. Un serveur de licence (`license server`) crée une licence d'utilisation à partir des droits de l'utilisateur. La machine virtuelle (`DRM Controller`) se sert de la licence pour rendre le contenu à l'utilisateur. Différentes solutions DRM utilisent des techniques différentes pour produire le contenu sécurisé (`content package`), des formats de licences (`license`) différents et des protocoles de communication entre composants (`Content packager`, `License server`, `Client`) différents.

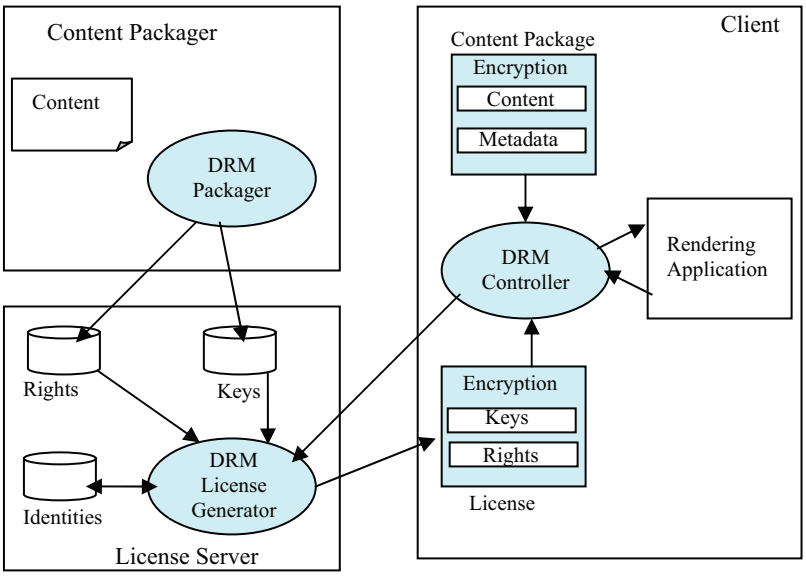


Figure 3 – Usage d'une ressource DRM

4. Langages d'Expression de Droits:REL (Rights Expression Languages)

L'effectivité de l'interopérabilités des systèmes de DRM dépend en partie du langage d'expression des droits utilisé pour créer les licences dans les différentes solutions. En effet, afin que deux systèmes se comprennent il est nécessaire que la licence éditée avec un des systèmes soit compréhensible par l'autre et vice versa. Un standard REL est donc l'une des clés pour l'interopérabilité des systèmes DRM propriétaires.

Plusieurs langages de droits ont été proposés `XrML`[4], `ODRL` [4], `MPEG-21 REL` [5], `FORM` [11]. A ce jour, deux principaux langages permettent de décrire les licences. D'un côté, `ODRL` (Open Digital Rights Language) est une recommandation `W3C`. Un sous ensemble de `ODRL` est utilisé par le forum des industries leaders du mobile (`OMA - Open Mobile Alliance`) dans le cadre du système DRM OMA. De l'autre côté, `MPEG-21 REL` est un standard `ISO/IEC`. Dans cette section,

[4] www.xrml.org

Interopérabilité des systèmes DRM

nous décrivons brièvement les langages MPEG-21 REL et ODRL. Ensuite nous les positionnons par rapport à l'objectif d'interopérabilité.

4.1. MPEG-21 REL

MPEG (Moving Picture Experts Group), est le groupe de travail SC 29/WG 11 de l'ISO/IEC (International Organization for Standardization/International Electrotechnical Commission) pour les technologies de l'information chargé du développement des normes internationales pour la compression, la décompression, le traitement et le codage de la vidéo, de l'audio et de leur combinaison, de façon à satisfaire un large panel d'applications. Un des standards produits par MPEG est MPEG-21. Son but est d'offrir un cadre normatif pour la distribution et l'usage des contenus multimédias à tous les intervenants de la chaîne. MPEG-21 est constitué de plusieurs parties dont la partie 5 qui spécifie le langage d'expression de droits MPEG-21 REL [5].

MPEG-21 REL est basé sur la proposition XrML (eXtensible rights Markup Language). Avec MPEG-21 REL il est possible de spécifier pour une ressource numérique (contenu ou logiciel), qui est autorisé à utiliser cette ressource, les droits disponibles et les conditions nécessaires à l'exercice de ces droits sur la ressource.

Le standard MPEG-21 dans sa partie 6 spécifie un dictionnaire de données qui définit de façon unique les différents termes utilisés par MPEG-21 REL pour exprimer les droits.

Le noyau de MPEG-21 REL est constitué des éléments principal, resource, right et condition (voir figure 4):

- Principal: identifie une entité telle que la personne, l'organisation, ou le dispositif à qui les droits sont attribués.
- Right: spécifie l'action que l'entité identifiée par l'élément principal peut être autorisé à exercer sur une quelconque ressource.
- Resource: identifie l'objet sur lequel est basée l'autorisation. Un URI (Unified Ressource Identifier) peut être utilisé pour identifier la ressource.
- Condition: spécifie une ou plusieurs conditions qui doivent être satisfaites pour que les droits soient exercés. Par exemple, une condition pourrait limiter l'écoute d'une chanson à un nombre de fois ou pendant un intervalle de temps précis.

Ces éléments sont encapsulés dans un élément grant, lui aussi partie d'un élément englobant license. Une licence (élément license) peut contenir une ou plusieurs autorisations (éléments grant), l'émetteur de la licence (élément issuer), accorde les privilèges contenus dans la licence. La figure 4 (a) montre une structure simple de licence MPEG-21 REL. Comme le montre cette figure, l'émetteur de la licence peut y attacher une signature numérique.

L'élément grant est la partie d'une licence MPEG-21 REL qui accorde à l'entité identifiée par l'élément principal le droit d'utiliser la ressource sous certaines conditions. Par exemple, considérons un fichier audio distribué à un utilisateur (Toto). Le document MPEG-21 REL possède une entrée qui exprime le fait que Toto a le droit d'écouter 3 fois ce fichier audio. La figure 4 (b) montre l'élément grant qui spécifie cette règle.

4.2. ODRL

ODRL est une proposition de standard pour l'expression des droits sur les contenus numériques. ODRL est destiné à fournir des mécanismes flexibles et interopérables pour la publication, la distribution et l'usage transparent des ressources numériques telles que la musique, la vidéo, les logiciels et autres créations numériques. Ce langage est une recommandation du W3C. Il ne nécessite pas de licence et est disponible librement.

Les licences ODRL exploitent un dictionnaire de données formé des éléments qui permettent d'exprimer des droits sur les contenus numériques.

ODRL est basé sur un modèle extensible d'expression de droits qui comprend trois principales entités:

- `Party`: comprend les utilisateurs finaux ou les détenteurs de droit. Les détenteurs de droits sont les entités qui participent à la création, à la production, ou à la distribution des ressources numériques. Cet élément est analogue aux éléments `principal` et `issuer` du standard `MPEG-21 REL`.
- `Permission`: représente les autorisations, qui peuvent contenir des contraintes, des obligations et des conditions. Les contraintes sont les limites imposées à l'usage de la ressource (par exemple regarder une vidéo au maximum 5 fois). Les obligations sont des pré requis pour l'obtention des permissions (par exemple payer 5$ chaque fois afin de regarder la vidéo). Les conditions spécifient des exceptions, qui si elles sont satisfaites, révoquent les permissions ou entraînent la renégociation de celles-ci (par exemple si la carte de crédit expire alors toutes les permissions sur la vidéo sont révoquées).
- `Asset`: c'est la ressource à protéger. Elle doit être identifiée de façon unique. C'est l'équivalent de l'élément `ressource` du standard `MPEG-21 REL`.

Considérons l'exemple précédent exprimé avec `MPEG-21 REL` à la figure 4 (b). La figure 5 donne sa représentation en `ODRL` en utilisant les entités `party`, `permission` et `asset`.

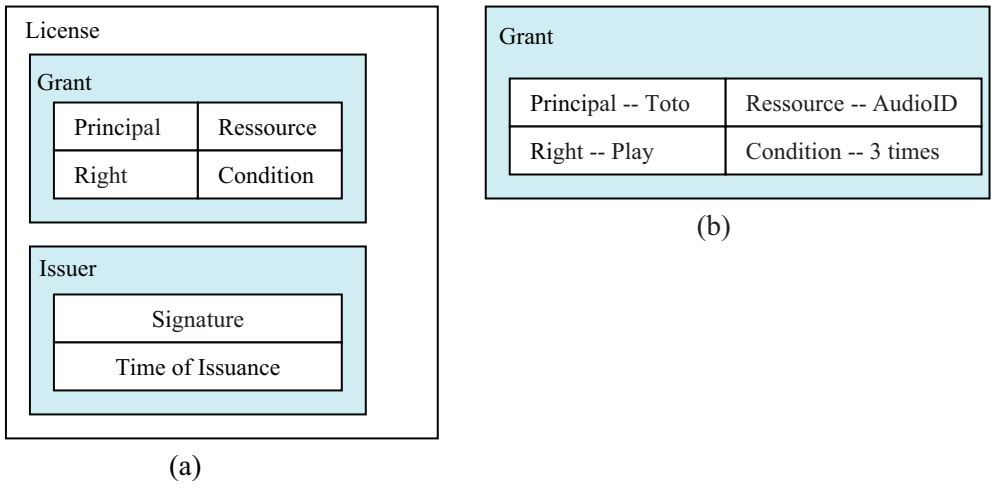


Figure 4 – (a) licence MPEG-21 REL – (b) Autorisation MPEG-21 REL

4.3. Interopérabilité entre MPEG-21 REL et ODRL

`ODRL` et `MPEG-21 REL` ont beaucoup de similarités. Ils sont syntaxiquement basés sur XML. Et bien que les termes utilisés dans ces langages soient différents, ils sont sémantiquement très proches et permettent d'adresser à quelques différences près les mêmes besoins. Cependant, il est difficile de faire le choix d'un langage au détriment de l'autre pour deux raisons au moins:

- Les deux langages `ODRL` et `MPEG-21 REL` bénéficient respectivement du soutien des organisations de standardisation `W3C` et `ISO/IEC`: D'une part, `ODRL` est une recommandation `W3C` dont un sous ensemble est utilisé par le forum des industries leaders du mobile (`OMA` - `Open Mobile Alliance`) dans le cadre du système DRM OMA. D'autre part, `MPEG-21 REL` est un standard `ISO/IEC`. Il est basé sur `XrML` qui a inspiré certains systèmes DRM propriétaires existants (Windows media Right manager de Microsoft en est un exemple).

- Ces deux langages sont au format XML. Ils sont donc extensibles et peuvent chacun de son côté évoluer par enrichissement de fonctionnalités. Nous ne pouvons donc dire que l'un est plus expressif que l'autre.

Ces deux langages d'expression de droits sont largement utilisés. Il est donc important de faciliter l'interopérabilité entre les systèmes qui les utilisent. Toute solution DRM doit donc être capable d'interpréter les licences éditées à partir des langages ODRL et MPEG-21 REL.

Dans ce papier nous préconisons l'interprétation des deux langages ODRL et MPEG-21 REL dans les solutions DRM. Le module d'interprétation de licences des clients DRM comporterait alors deux unités, une pour chacun des deux langages. Ainsi à la réception d'une licence, le client DRM utiliserait l'unité appropriée pour l'interpréter.

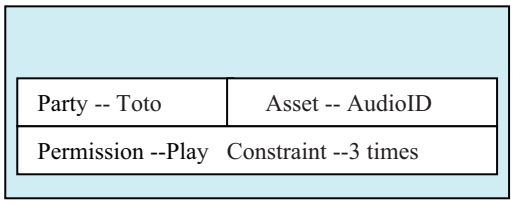


Figure 5 – Autorisation en ODRL

5. Proposition d'architecture pour des systèmes DRM interopérables

Dans cette section, nous proposons une architecture pour faciliter l'interopérabilité des systèmes DRM. Nous décrivons d'abord les différents composants de l'architecture, avant de détailler l'interface et le protocole de communication entre ces composants.

5.1. Composants de l'architecture

Nous proposons une architecture composée des éléments de base communs à la majorité des systèmes DRM propriétaires existants. Nous exploitons, pour ne citer que quelques unes, les architectures des systèmes Windows Media Rights Manager de Microsoft, Electronic Media Management System de IBM, Media Commerce Suite de RealNetworks. La figure 6 présente notre architecture. Le composants (ou acteurs) de cette architecture sont décrit comme suit:

5.1.1. content owner

Ce composant fournit le contenu numérique, ainsi que les contrats et les droits au packager. Les contrats et les droits détermineront l'usage autorisé du contenu par un utilisateur donné.

5.1.2. packager

Le `packager` assure les fonctions suivantes: la compression si possible des données et la protection du contenu. La protection du contenu comprend le chiffrement et le watermarking.

Le `packager` reçoit en entrée la ressource non protégée du `content owner`. Il fournit les clés de décryptage, l'algorithme utilisé pour protéger le contenu ainsi que les droits et les contrats au composant `license/key_distribution`. Enfin le contenu protégé est transmis à l'élément `content_distribution`.

Le watermarking consiste à rajouter, sur un medium (qui peut être une image, une chanson, un film vidéo), une marque (en anglais, watermark signifie filigrane) qui doit être suffisamment imperceptible pour ne pas détériorer le medium et suffisamment robuste pour pouvoir être décelée même après traitement du medium. Le contenu d'une marque, typiquement quelques bits d'information, peut être de différentes natures.

- La marque peut contenir des informations sur les permissions attachées au document. Ces informations reflètent les droits et contrats spécifiés par le `content owner`. C'est au dispositif terminal client que revient la charge de détecter et de respecter les instructions de la marque en fonction de la licence acquise. Aussi, le client ne doit pas être corrompu.
- La marque peut indiquer l'identifiant du `content owner`, celui du contenu et l'`URL` (Uniform Ressource Locator) du serveur de licences.

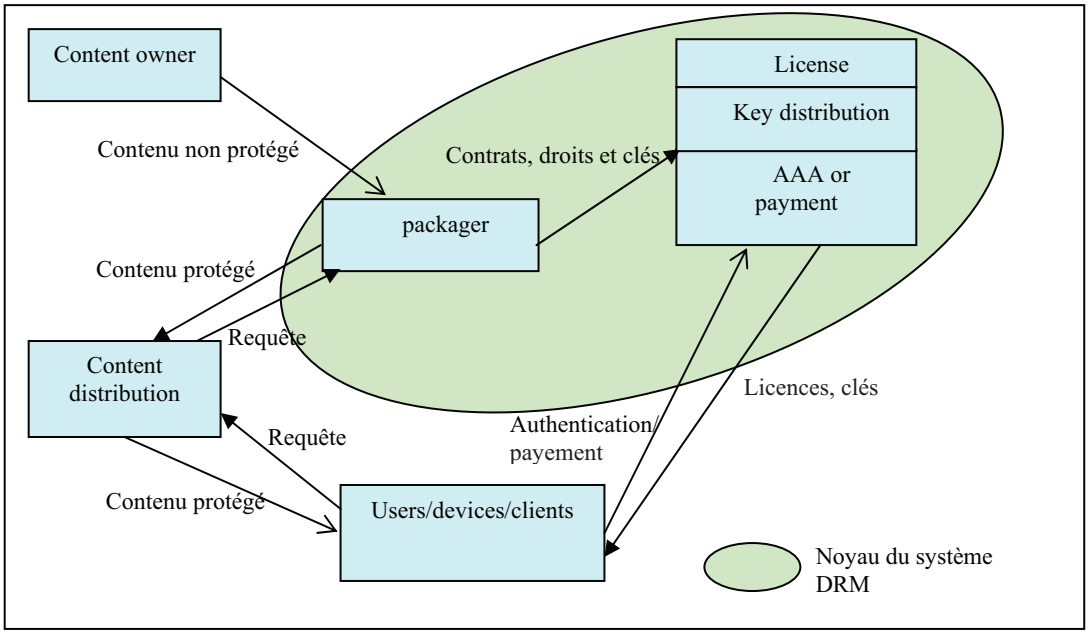


Figure 6 – Architecture DRM

5.1.3. License/key_distribution/AAA_or_Payment

L'entité `AAA_or_Payment` assure l'autorisation, l'authentification et le contrôle d'accès. En complément ou tout simplement à la place de ce processus, les utilisateurs payent les frais de licence. Dès que cette étape est faite, il contacte le module `license/key_distribution` qui fournit la licence à l'utilisateur.

La composante `license/key_distribution` reçoit les droits, contrats et clés de décryptage du `packager`. Il délivre les licences, clés de décryptage ainsi que l'algorithme utilisé pour protéger le contenu aux utilisateurs sur la base:

- des informations fournies par la composante `AAA_or_payment`; Ces informations concernent l'identifiant du contenu et les contraintes d'usages telles que le nombre de fois où l'utilisateur pourra écouter la chanson ou la limite dans le temps de la validité de la licence.
- des contrats et des droits reçus du `packager`.

5.1.4. Content_Distribution

Ce composant peut distribuer le contenu sous plusieurs formes

- La distribution peut se faire par Internet en téléchargement ou en streaming (diffusion des contenus audio ou vidéo en continu, au fur et à mesure du déléchargement du fichier);
- Elle peut aussi se faire sur supports physiques tels que DVDs, CDs.

Interopérabilité des systèmes DRM

Il reçoit les contenus sécurisés du packager. Chaque contenu à un identifiant qui permet d'aller chercher la licence et la clé de décryptage.

5.1.5. Users/devices/clients

Les droits peuvent être associés à l'utilisateur ou au matériel. Une application cliente est responsable du contrôle d'usage et du décryptage du contenu numérique. L'application cliente doit disposer d'une paire de clés (privée, publique) et doit pouvoir interpréter les licences ODRL et MPEG-21 REL.

5.2. Communication entre composants

La figure 6 présente notre architecture sous deux parties. Une partie formée du noyau du système DRM. C'est cette partie qui s'occupe de la protection du contenu ainsi que de la délivrance des licences d'utilisation. Dans l'autre portion interviennent les composants externes (content owner, content_distribution, users/devices/clients) dont la présence est nécessaire au bon fonctionnement du système.

L'effectivité de l'interopérabilité repose sur la communication entre le noyau et les entités externes. Aussi, le noyau met-il en place une interface de services web décrite dans un document WSDL pour faciliter les échanges avec les autres entités. Les communications se font alors par échanges de messages SOAP sécurisés [8].

Dans la suite de cette section nous présentons les communications entre éléments d'une paire de composants incluant à chaque fois un composant externe et le composant du noyau qui lui est directement associé. Pour chaque paire nous présentons le squelette des principaux messages qui sont échangés. Le but de ces exemples n'est pas de contraindre la structure des messages qui sont échangés mais de montrer que les informations requises par chaque composant peuvent bien être transmises par la solution proposée, à savoir les messages SOAP sécurisés. Ainsi, nous ne faisons pas apparaître d'appels à méthodes dans les échanges.

5.2.1. Content owner <- - > packager

Le content owner chiffre le contenu numérique avec une clé générée de façon aléatoire. La clé générée est chiffrée à son tour en utilisant la clé publique du packager. L'ensemble contenu chiffré, clé chiffrée, algorithmes de chiffrement, contrats et droits est alors transmis au packager (voir figure 7 (a)). A la réception du message, le packager obtient la clé sécrète utilisé pour chiffrer le contenu à l'aide de sa clé privée. La ressource protégée est alors déchiffrée à son tour à l'aide de la clé sécrète. Si tout se passe bien alors la packager retourne au content owner un accusé de réception. Sinon un message d'erreur est envoyé (voir figure 7 (b)).

5.2.2. Packager <- -> license server

Le packager chiffre le contenu numérique avec une clé générée de façon aléatoire. La clé générée est chiffrée à son tour en utilisant la clé publique du license server. Le packager marque (watermarking) ensuite le contenu chiffré et envoie l'ensemble contenu protégé, clés chiffrés, algorithmes, contrats et droits au license server. La structure du message envoyé est similaire à celle de la figure 7 (a). Le license server stocke ces informations dans sa base de données et retourne un accusé similaire à celui de la figure 7 (b) au packager.

5.2.3. Packager <- -> content distribution server

L'interface entre le packager et le content_distribution est une interface de requêtes. Le composant content_distribution adresse des requêtes au packager par rapport aux contenus qu'il aimerait acquérir. Le packager lui retourne les contenus protégés demandés s'ils en existent dans sa base.

5.2.4. Content distribution server <- -> users

L'utilisateur télécharge les contenus sécurisés de façon classique. Notons qu'il peut aussi les recevoir sur supports physiques tels que CDs. Il peut aussi les transmettre à d'autres utilisateurs.

5.2.5. Users/devices/clients <- - > license server

Le client DRM et le `license server` communiquent par échanges de messages SOAP sécurisés.

Lorsque le client tente d'exploiter le contenu protégé pour la première fois, il constate que la licence n'est pas disponible. Il extrait l'url du `license server` du contenu et le contacte pour obtenir une licence. La demande de licence est un message soap contenant l'identifiant du contenu et la clé publique du client (voir figure 8 (a)).

Le `license server` renvoie l'utilisateur vers la composante `AAA_or_payment` pour l'achat d'une licence ou pour authentification. Après authentification et/ou payement, Le composant `AAA_or_payment` transmet alors au `license server` les éléments définissant le type de licence sollicité par l'utilisateur.

Le `license server` chiffre alors la clé de décryptage du contenu en se servant de la clé publique du client. Sur la base des contrats et des droits liés à ce contenu, il établit une licence (`ODRL` ou `MPEG-21 REL`) sur laquelle il appose sa signature. Ensuite il transmet l'ensemble clé chiffrée, signature numérique et licence au client DRM (voir figure 8 (b)).

Le client DRM déchiffre la clé reçue à l'aide de sa clé privée. Après vérification de la signature, il peut alors décrypter la ressource et la rendre à l'utilisateur suivant les directives de la licence.

```
<?xml version="1.0"?>
<soap:envelope>
  <soap:header>
    <wss:security>
      <encryptedKey>
       <encryptionMethod Algorithm="rsa-1_5"/>
        <cipherData>
          <cipherValue>donfdle…</cipherValue>
        </cipherData>
      </encryptedKey>
       <signature> --  optionnel --  </signature>
    </wss:security>
  </soap:header>
  <saop:body>
   <encryptedData>
     <encryptedMethod Algorithm="tripledes-cbc"/>
     <cipherData>
      <cipherValue>d2dnqye…</cipherValue>
     </cipherData>
   </encryptedData>
   <rights-and-contracts>…</rights-and-contracts>
  </soap:body>
</soap:envelope>
```

(a)

```
<?xml version="1.0"?>
<soap:envelope>
  <soap:header>
    <wss:security>
     <signature>-- optionnel -- </signature>
    </wss:security>
  </soap:header>
  <saop:body>
    <ack> accusé de réception  </ack>
    <soap:fault>-- optionnel --</soap:fault>
  </soap:body>
</soap:envelope>
```

(b)

Figure 7 – messages échangés entre le content owner et le packager

Interopérabilité des systèmes DRM

```
<?xml version="1.0"?>
<soap:envelope>
 <soap:header>
   <wss:security>
     <binarySecurityToken
                  valueType="X509">
       MIIEZzCCA9CgAwI…..
     </binarySecurityToken>
   </wss:security>
 </soap:header>
 <saop:body>
   <content_id>
       un identifiant de contenu
   </content_id>
 </soap:body>
</soap:envelope>
```

(a)

```
<?xml version="1.0"?>
<soap:envelope>
 <soap:header>
   <wss:security>
     <encryptedKey>
      <encryptionMethod Algorithm="rsa-1_5"/>
      <cipherData>
        <cipherValue>dinnfble…</cipherValue>
      </cipherData>
     </encryptedKey>
     <signature> -- Required -- </signature>
   </wss:security>
 </soap:header>
 <saop:body>
   <license type="MPEG-21 REL">
    <rights> contenu de la licence </rights>
   </license>
 </soap:body>
</soap:envelope>
```

(b)

Figure 8 – messages échangés entre clients DRM et le `license server`

6. Etat de l'art

Dans cette section, nous présentons les travaux existants. Nous comparons à chaque fois notre proposition avec ces travaux.

A notre connaissance, `Opera` et `DMDFusion` sont les seuls systèmes existants qui facilitent l'interopérabilité des solutions DRM. `DMDFusion` [13] incorpore les systèmes `Microsoft`, `Adobe` et `Real Networks`. Il fournit une interface commune pour ces trois solutions. Le contenu est protégé en utilisant les techniques spécifiques à chacune des solutions. Bien qu'intéressante, cette proposition ne prend pas en compte les autres systèmes existants. `Opera` [12] comble ce manquement en proposant une architecture ouverte pour l'interopérabilité des systèmes DRM. Il s'agit ici aussi d'un interfaçage de solutions existantes. Lorsque le système reçoit une demande de licence, il édite une licence pour le système de DRM d'où provient la requête. Ainsi, en réponse à une requête `Real` sera éditée une licence `Real Network`. Nous en déduisons 2 inconvénients majeurs:

- Non seulement toutes les requêtes en vue d'obtenir une licence sont gérées par le client `Opera`, mais le processus de délivrance de licences est aussi centralisé sur son serveur.
- Si un constructeur change sa technologie ou met sur le marché un nouveau produit, le noyau du système `Opera` devra être modifié pour prendre en compte les nouvelles données.

Notre proposition n'est pas liée à une solution particulière. Elle utilise les standards existants pour faciliter la communication entre les composants d'une architecture basique de DRM. Notre suggestion est évolutive. En effet il est possible sans modifier les composants présents dans l'architecture d'en ajouter un. Il suffirait de définir l'interface du nouveau composant en prenant en compte l'interface (`WSDL`) de ceux avec lesquels il échangera (en consultant par exemple un répertoire `UDDI`). Mieux encore, notre proposition donne la possibilité aux intervenants de l'architecture de définir des interfaces standardisées.

Polo et al. [14] constatent que les langages `ODRL` et `MPEG-21 REL` ont des entités différentes mais qui représentent les mêmes informations. En conséquence, ils proposent un mécanisme basé

XSLT [16] (eXtensible Stylesheet Language Transformation) pour la transformation d'une licence exprimée en MPEG-21 REL en une licence équivalente en ODRL et vice versa. Cette idée est ingénieuse, mais manque de sémantique. Elle nécessite donc la définition d'un modèle générique (à l'aide de RDF - Ressource Descriptive Framework [15] par exemple) et le développement d'une ontologie pour ce modèle. Nous proposons dans ce papier l'usage des langages MPEG-21 REL et ODRL comme langages d'expression de droits.

7. Conlusion

Le manque d'interopérabilité entre les différents systèmes DRM a souvent été avancé pour expliquer l'échec des DRMs notamment dans le cadre de la distribution de contenu musical. Dans ce papier, nous identifions les obstacles à ce besoin d'interopérabilité et proposons des solutions pour les contourner. Ainsi, nous utilisons les concepts véhiculés par les services web pour définir l'architecture générale d'un système DRM inter opérable. Les différents acteurs de l'architecture communiquent par échanges de messages SOAP.

Nous avons établi qu'il était difficile de faire le choix entre les deux langages d'expression de droits les plus utilisés de la littérature que sont ODRL et MPEG-21 REL. Nous envisageons alors la définition d'un langage générique qui engloberait ODRL et MPEG-21 REL. Autre perspective de ce travail consiste à définir une ontologie des DRMs afin de spécifier de façon standard les interfaces des différents acteurs de l'architecture proposée.

8. References

[1] *Electronic copyright management systems: Requirements*, players and technologies, F. Bartolini, Cappellini, A. Piva, A. Fringuelli, In proceedings of the 10th IEEE International Workshop on Database and Expert Systems Applications, 1999.

[2] *Security architectures for controlled digital information dissemination*, J. Park, R. Sandhu, J. Schifalacqua, *In proceedings of the 16th Annual Computer Security Applications Conference*, 2000.

[3] *Fair use, drm and trusted computing*, J. Erickson, Communications of the ACM Vol. 46, No. 4 (2003), 34-39.

[4] *ODRL-Open Digital Rights Language*, R. Iannella, W3C Note, http://www.w3.org/TR/odrl/ 2002.

[5] *ISO/IEC 21000:2004 Information technology – Multimedia framework (MPEG-21)*, International Organization for Standardization (ISO), 2004.

[6] SOAP - *Service Oriented Architecture Protocol,* M. Gudgin, M. Hadley, N. Mendelsohn, J. J. Moreau, H. F. Nielsen, W3C Recommendation, http://www.w3.org/TR/soap/, 2003.

[7] Namespace in XML 1.0, T. Bray, D. Hollander, A. Layman, R. Tobin, W3C Recommendation, http://www.w3.org/TR/REC-xml-names/, 2006.

[8] Web services security: SOAP Message Security 1.1, OASIS Standard Specification, 2006.

[9] *XML Signature Syntax and Processing*, M. Bartel, J. Boyer, B. Fox, B. LaMachia, E. Simon, W3C Recommendation, http://www.w3.org/TR/xmldsig-core/, 2002.

[10] *XML Encryption Syntax and Processing*, T. Imamura, B. Dillaway, E. Simon, W3C Recommendation, http://www.w3.org/TR/xmlenc-core/, 2002.

[11] *FORM: a federated rights expression model for open DRM frameworks*, T. Sans, F. Cuppens, N. Cuppens-Boulahia, In proceedings of the 11th Annual Asian Computer Science Conference, 2006.

Interopérabilité des systèmes DRM

[12] OPERA-Interoperability of Digital Rights Management (DRM) Technologies, S. Wegner, Eurescom Technical Information, http://www.eurescom.de, 2003.

[13] DMD Fusion, DMDSecure/SafeNet, http://www.safenet-inc.com/.

[14] *Interoperability between ODRL and MPEG-21 REL, J. Polo*, J. Prados, J. Delgado, In proceedings of the 1st International ODRL Workshop, 2004.

[15] *RDF-Ressource Descriptive Framework,* W3C Semantic Web Activity, http://www.w3.org/RDF/.

[16] XSLT, J. Clark, W3C Recommendation, http://www.w3.org/TR/xslt, 1999.

Chapitre 3

BackPlan^{TM} entreprise innovante

Sommaire

3.1 Contexte initial : du besoin au projet

3.1.1 Le besoin

Les projets industriels évoluent dans une organisation particulièrement complexe. Les acteurs sont nombreux, multidisciplinaires, multilingues et géographiquement répartis sur plusieurs pays.

Ces projets nécessitent la coordination des différents services de l'entreprise maître d'ouvrage avec les intervenants : les ingénieries, les entreprises réalisatrices, les organismes de contrôle, les institutionnels : en un mot, la coordination de l'ensemble des interfaces.

Actuellement, aucune solution technique ne permet de gérer cet aspect qui n'est absolument pas pris en compte par les outils traditionnels tels que les ERP (Enterprise Resource Planning), les outils de collaboration (groupware,...).

Par exemple les ERP peuvent être présents dans l'entreprise alors qu'ils devraient l'être pour tous les facteurs du projet : une entreprise tierce ne peut pas se permettre d'ouvrir son outil interne, le risque de perte de confidentialité liée à une mauvaise configuration par exemple est trop grand pour la sécurité des données de la société.

Dans la pratique, la coordination repose donc principalement sur l'expérience des chefs de projet et se base uniquement sur le suivi de la "courbe en S planning / coût ", les échanges se résument essentiellement à l'utilisation du courrier électronique, du téléphone et pour les échanges de données massifs, un CD-ROM envoyé par le service postal !

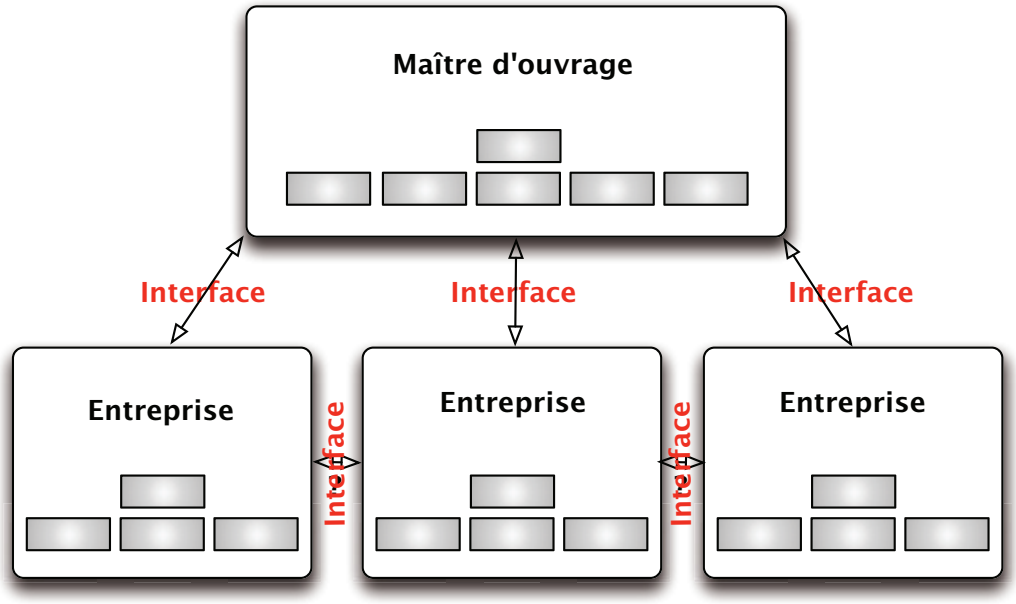


FIGURE 3.1 – La gestion des interfaces

Cette coordination d'interface reste particulièrement sensible pour différentes raisons.

Tout d'abord, il existe un manque de visibilité globale sur l'organisation en mode projet :

1. Le maître d'ouvrage et sa direction ont peu ou pas d'indicateurs concernant les échanges entre les entreprises impliquées : les entreprises impliquées peuvent s'échanger des informations cruciales pour le projet alors que le responsable n'est pas au courant.
2. Chaque entreprise concentre d'abord ses efforts sur ses réalisations propres, avant d'identifier et de clarifier les problèmes d'interface.
3. Peu d'entreprises ont la connaissance de l'organisation complète du projet.
4. Certaines entreprises justifient ainsi des retards car elles sont en attente d'informations ou de choix techniques.
5. Concernant les interfaces, les problèmes de contrôle et de responsabilité sont particulièrement délicats.
6. Le manque d'anticipation en gestion d'interface provoque de nombreux problèmes, retards et surcoûts qui se révèlent durant les phases finales du projet : installation et commissioning.
7. Il n'existe pas d'outil d'analyse des risques spécifiques au suivi des interfaces.

Expérimentation : un exemple de manque de coordination d'interface
Lors d'un projet : Tous les indicateurs et comptes rendus de réunion d'un projet ont été transmis à un chef de projet expérimenté mais extérieur au projet concerné. Après un temps suffisant pour la consultation et la compréhension de ces informations, il a été demandé à ce chef de projet son analyse, sa hiérarchisation des difficultés et des risques, puis sa proposition de plan d'action.
Résultat : Ses conclusions sont très décalées de la réalité car les méthodes et outils actuels ne permettent pas à la maîtrise d'ouvrage d'avoir une vision claire du terrain !

3.1.2 Description du projet : méthodologie, outil informatique et service.

Afin de structurer la coordination d'interface et d'améliorer l'efficacité des pratiques actuelles, le projet a consisté en trois phases :

1. rédiger une méthodologie basée sur la formalisation de la communication et l'analyse des risques,
2. développer un outil informatique facilitant l'application de la méthodologie,
3. constituer une société spécialisée en coordination d'interface et management des risques associés qui propose ses services aux maîtres d'ouvrage en appliquant la méthodologie et en la mettant en œuvre avec l'outil informatique.

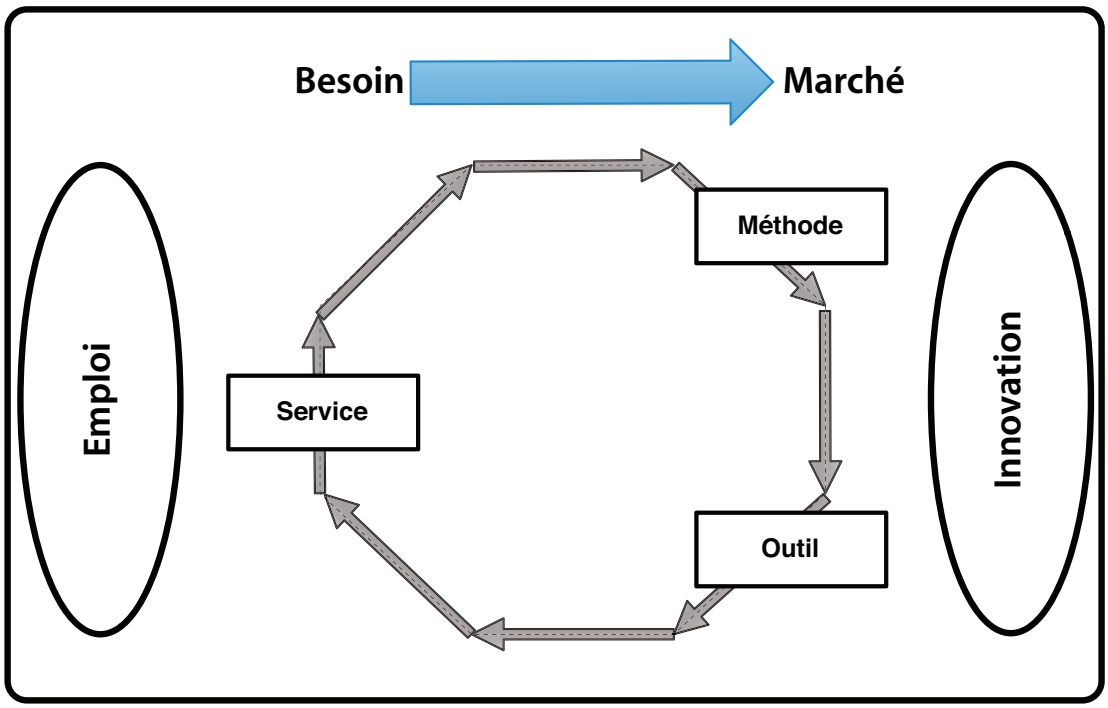


FIGURE 3.2 – Méthode - Outil - Service selon BackPlan™

La méthodologie :

La méthodologie formalise sous forme de procédures les pratiques d'échange, de questionnement, de prise de décision et d'approbation. Elle constitue la base des bonnes pratiques, afin d'uniformiser les démarches et de les optimiser. Ainsi, les informations sont structurées et constituent la base de l'analyse des risques.

L'outil :

— À partir de retours d'expérience sur des projets similaires et de concertations préliminaires avec les entreprises du projet, l'outil permet d'établir un modèle prévisionnel des échanges entre entreprises et d'en pondérer l'importance.
— Durant toute la durée du projet, l'outil permet d'appliquer les procédures avec une utilisation simple (semblable à une messagerie) et depuis tout poste relié à internet.
— Il permet de centraliser les informations, les échanges et commentaires est les participants au projet, tout en permettant les recherches.
— Cet outil est complémentaire aux logiciels actuels de planning et de gestion électronique des documents (GED).
— Il permet d'établir des indicateurs, de vérifier l'avancement des entreprises et de fournir à la direction du projet des informations particulièrement pertinentes dans le cadre de

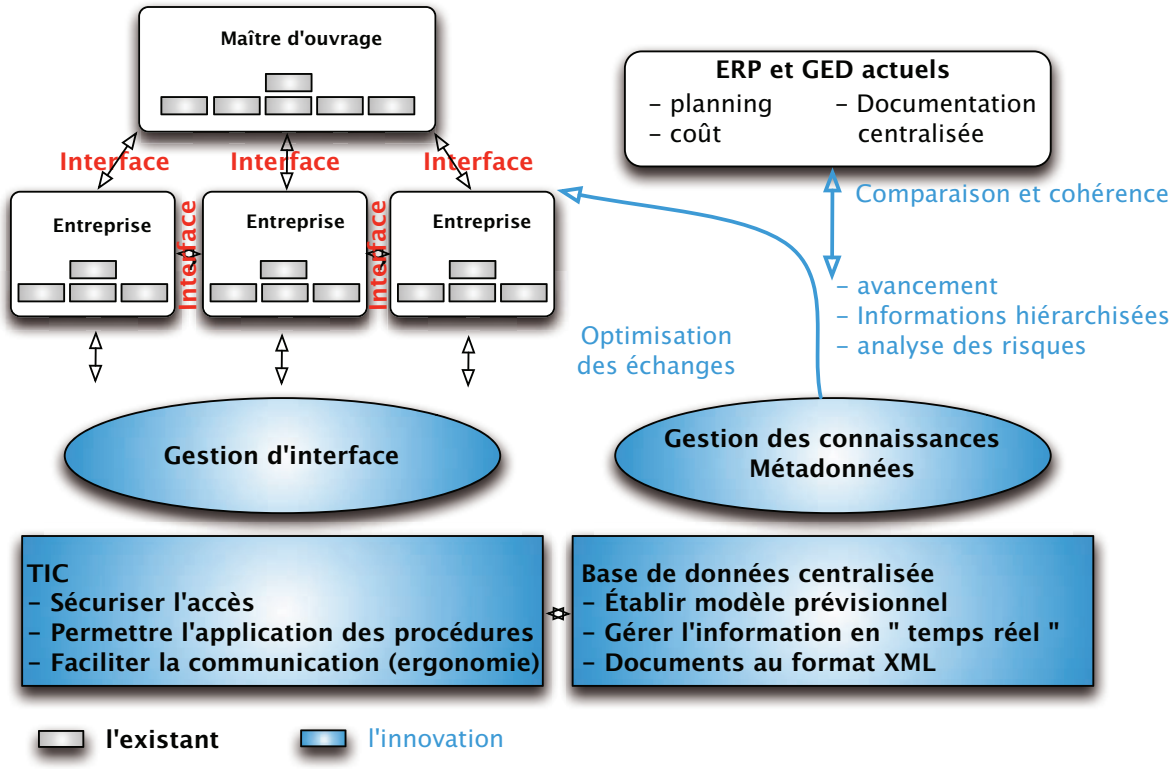


FIGURE 3.3 – La méthodologie selon BackPlan™

la maîtrise des risques.

La société prestataire de service en coordination d'interface :

L'aboutissement de la démarche est de créer une société de prestation de service leader en coordination d'interface et en gestion des risques associés.

3.2 Du projet à la création de la société

3.2.1 La rencontre de personnes autour d'une même idée :

Comme tout projet, il s'agit avant tout d'une rencontre de personnes. La personne à l'origine du projet Backplan est Magali Ricarde, ingénieure de formation, professeure associée à l'école nationale supérieure de génie des technologies industrielles (ENSGTI), école d'ingénieurs interne à l'université de Pau et des Pays de l'Adour. C'est elle, qui après seize années d'expérience dans les domaines du gaz du pétrole et de la chimie, en particulier sur de grands projets industriels, a soulevé le problème de la carence de gestion des interfaces dans ces projets. Elle a été rejointe par Sophie Bernier, diplômée d'une école supérieure de commerce, elle assure la partie prospective et

commerciale. Enfin, je participe sur la partie technique, recherche et developpement à l'évolution des outils et assure le lien avec la partie recherche.

3.2.2 La génèse

- Le projet BackPlan a débuté en 2008. Il a été accompagné par l'Incubateur Régional d'Aquitaine [16] depuis mars 2009. Ce projet a été entouré dés le début par l'université de Pau et des Pays de l'Adour, à travers son laboratoire d'informatique, mais également par l'école nationale supérieure de cognitique de Bordeaux [17] en ce qui concerne l'environnement ergonomique des outils web. Depuis le début du projet, deux étudiants de l'ENSC ont effectué leur stage au sein de l'équipe.
- La marque BackPlan a été déposée le 2 Avril 2009.
- Suite à l'expertise de Mr Guy Le Moal, Directeur de la Plateforme TEPF [18] le 21 Juillet et à la présentation du projet BackPlan au comité scientifique et technique d'Hélioparc le 23 Septembre 2009, l'équipe a intégré la technopole Hélioparc.
- Le lancement du site internet [19] a eu lieu le 1er Décembre 2009 en même temps que la présentation du logo

FIGURE 3.4 – Logo de la société BackPlan™ (2009)

- Année 2010 : publicité du projet et prospection des marchés, test et developpement des outils.
- Création de la société BackPlan en janvier 2011

3.2.3 La société maintenant

Rapidement la société a démarré sur des projets de construction liés au transport du gaz, mettant en œuvre les méthodes et outils développés pendant la phase de gestation.

16. http ://www.incubateur-aquitaine.com
17. http ://www.ensc.fr/http ://www.ensc.fr/
18. Total Exploration Production France
19. http ://www.backplan.fr

Parallèlement c'est toute une activité à la gestion électronique de documents qui se développe : la documentation faisant partie intégrante d'un projet de construction industrielle, cette activité est apparue comme étant un outil important pour d'autres acteurs de l'économie qui sont devenus des clients de BackPlan™ , ils bénéficient ainsi de leur expertise.

Dans un tout autre domaine ce sont les activités de planification qui sont exercées, en particulier sur le démantèlement de l'usine de gaz de Lacq (64 - Aquitaine) de la société TEPF.

Actuellement la société emploie 6 personnes : 2 co-gérantes ainsi que 4 personnes diplômées de l'université de Pau et des Pays de l'Adour pour qui il s'agit du premier emploi (2 ingénieures et 2 Licences professionnelles).

3.3 La gestion du risque dans l'entreprise BackPlan™

Dés la création de la société, nous avons mis en avant la confiance que doivent avoir les clients dans la gestion de leurs informations : traitement et stockage. Aussi, un accent tout particulier a été placé sur la sécurité de l'information : à tous les niveaux il est demandé et rappelé d'être extrêmement vigilant quant à la circulation de ces informations. Ces éléments font bien entendu parti de la politique de la société.

3.3.1 La certification « ISO 27005 Risk Manager »

En France, l'organisme de certification LSTI [20] propose une certification individuelle "ISO 27005 Information Security Risk Manager" et plusieurs entreprises comme HSC [21] ont annoncé la formation correspondante. Enfin l'ISO 27005 est recommandée pour la mise en place d'un SMSI selon la norme ISO 27001, et elle est d'ores et déjà incontournable au niveau international. La norme insiste sur deux points de décisions, soit le travail est satisfaisant, soit il faut réitérer. Cette approche itérative améliore la finesse de l'analyse à chaque itération, fournit une bonne répartition entre le temps et l'effort fourni pour identifier les mesures de sécurité, permet de traiter les risques en fonction des ressources et moyens disponibles, facilite les liens entre les risques et les conséquences sur les métiers, permet d'avancer lorsque les interlocuteurs sont absents, en facilitant la gestion des susceptibilités et des aspects politiques entre interviewés. L'approche itérative de l'ISO 27005 permet à la gestion de risques de tendre progressivement vers une maîtrise des risques de haut niveau et conforme aux besoins de l'organisme.

20. http ://www.lsti-certification.fr/
21. Hervé Schauer Consultants

3.3.2 La formation

La formation certifiante "ISO 27005 Risk Manager" traite de la norme ISO 27005 et de la gestion du risque de sécurité de l'information en général.

Cette formation a permi de mener de bout en bout un processus de gestion du risque et de gérer son cycle de vie.

Les objectifs de la formation sont les suivants :

— Implémenter la norme ISO 27005 et autres méthodes dans toutes circonstances.

— Décrire le processus de gestion des risques et son cycle de vie.

— Pouvoir manager et réaliser une appréciation des risques.

— Communiquer les ressources et les outils disponibles afin de réaliser une appréciation des risques optimale.

— Préparation à l'examen en fin de session.

La durée de la formation est de trois jours : deux jours et demi pendant lesquels exposés, cas pratiques, exercices et études de cas sont alternés et une demi journée réservée à l'examen.

Formateur(s)

La formation est dispensée principalement par Hervé Schauer [22] , Président fondateur de Hervé Schauer Consultants, Cabinet de consultants en sécurité informatique depuis 1989 - Spécialisé sur Unix, Windows, TCP/IP et Internet.

L'examen Lors de la dernière après-midi chaque stagiaire passe un examen conçu et corrigé par LSTI, organisme de certification

3.3.3 Une politique de management de la sécurité de l'information

Ainsi au sein de la société, les nouvelles personnes recrutées, mais également les sociétés clientes, sont sensibilisées à cette problématique.

Actuellement, ce sont trois personnes qui sont certifiées dans le domaine de la sécurité de l'information :

— Certification d'auditeur ISO 27001 : permet d'attester qu'une personne a acquis l'expérience et les capacités à mener un audit selon la norme ISO/CEI 27001

— Certification de mise en œuvre de la norme ISO 27 001 : permet d'attester qu'une personne a acquis l'expérience et les capacités pour conduire un projet de mise en place d'un

22. Certifié ISO27001, Lead Auditor et Lead Implementer, Certifié ISO27005 Risk Manager, Certifié ISO22301 Lead Auditor, Certifié CISSP, ISC2 Certifié GSLC (GIAC Security Leadership Certification), Participant à la normalisation en sécurité à l'AFNOR depuis 1990 et à la CN27 depuis sa création en 1993, Membre du groupe ISO 27001 du Clusif, Animateur du Club 27001

système de management de la sécurité selon la norme ISO/CEI 27001.

— certification ISO/IEC 27005:2011 de management du risque : permet d'attester qu'une personne a acquis les compétences nécessaires pour mener à bien des appréciation de risques dans le domaine de la sécurité de l'information.

3.4 Les collaborations avec le Laboratoire d'informatique de l'université de Pau et des Pays de l'Adour

3.4.1 Pendant la phase d'incubation

Une première convention, sous forme de contrat d'étude, a été signée avec le LIUPPA [23] en mai 2009 pour le développement de l'étude d'un outil informatique qui permettra d'orchestrer la méthode mise en place par la société.

Cette étude repose sur six points. La première étape constitue l'analyse des besoins : il qui les fonctionnalités externes du système sans considérations techniques, en se limitant au niveau du métier et de l'utilisateur. Ensuite il s'agit du découpage fonctionnel et du diagramme des interactions de cette étape il faudra fournir un document précis qui spécifie les fonctionnalités attendues, l'objectif étant de définir les frontières du système, de décrire les fonctionnalités avec les scénarios correspondants, il faudra également définir les interactions entre les modules et les enchaînements entre les écrans. dans une troisième étape faudra identifier les modules applicatifs afin d'y associer les technologies spécifiques ou existantes. Nous possons ensuite à la revue de conception qui va leur permettre de confronter le travail réalisé par le LIUPPA aux besoins du projet BackPlan™ . Puis pour chaque module on précise le niveau d'adéquation au projet, le coût, la robustesse, la pérennité possible dans le temps. C'est dans cet étape que nous allons mettre en œuvre les logiciels et les tester afin de vérifier leur robustesse. Enfin la dernière étape doit permettre d'établir le modèle d'architecture technique et logicielle du système : cette architecture intègre les besoins métiers, la technique actuelle, les priorités et facteurs qualités qui ont été définis (robustesse, efficacité, portabilité etc.). Cette architecture détaille les sous systèmes : application en (interface), domaine (métier) et infrastructures (implémentation).

Après avoir obtenu un cahier des charges rédigé par Magali Ricarde, l'analyse et la conception de cet outil a débuté. L'objectif de cette étude était de créer un prototype à partir de briques informatiques déjà existantes et disponibles (briques à étudier pour pouvoir sélectionner celles dont nous avons besoin). De nombreux outils ont été testés par Manuel Munier (LIUPPA) et

23. Laboratoire d'Informatique de l'UPPA

moi même pour que le prototype s'appuie sur des éléments robustes.

La présentation des éléments constitutifs du prototype a été faite le 12 mars 2010 par Manuel Munier et moi même. A partir de ce prototype, nous avons développé une application industrielle qui servira de démonstration aux clients ainsi qu'un guide utilisateur adapté. Elle s'appuie sur des éléments open sources : moteurs de recherche, moteur BPM (Business Process Management), stockage, le tout interfacé via une technologie Web.

Il s'agissait de démontrer que le modèle et pouvait s'adapter à la méthode proposée.

3.4.2 Un partenariat qui se pérennise

Actuellement, suite au lancement de la société et après un an d'existance, une nouvelle collaboration vient de se mettre en place, et ce, pour trois années.

Il s'agit d'une convention de recherche établie entre la société et le laboratoire d'informatique de l'UPPA qui a pour objet de developper les actions dans le domaine du traitement des métadonnées dans les plateformes documentaires et la publication d'articles scientifiques à portée nationales et internationales : les premiers résultats de ce partenariat sont Trustcom 2012 et Dpm 2013.

3.4.3 Un statut reconnu pour BackPlan^{tm}

Ces travaux et ce partenariat, ainsi que cette thèse ont mené la société BackPlan™ à obtenir le statut reconnu conjointement par le Ministère de l'économie et des finances et le Ministère de l'Enseignement Supérieur et de la Recherche de Jeune Entreprise Universitaire en avril 2013.

C'est actuellement la seule JEU d'Aquitaine.

Chapitre 4

TRUSTCOM 2012
Self-Protecting Documents for Cloud Storage Security

Sommaire

4.1 Présentation

Cet article a été présenté à la conférence IEEE TrustCom 2012 organisée à Liverpool (Grande Bretagne) du 25 au 27 juin 2012. Il s'agit de la 11ème conférence internationale sur la confiance, la sécurité et la vie privée dans les systèmes informatiques et les communications.

C' est également la première collaboration que nous avons faite avec la société backplan ™ : à travers son expérience des métiers du pétrole et du gaz, nous avons pu appliquer notre modèle de documents intelligents sur une problématique bien réelle, mettant en œuvre de nombreuses sociétés autour d'un même projet. Les questions liées à l'usage de tout ou partie du document, la génération de métadonnées et surtout leur exploitation, trouvent là pleinement leur utilisation.

4.2 Introduction

Dans les systèmes d'information la sécurité des données est l'un des aspects les plus importants. Cela concerne tant la confidentialité des informations que leur intégrité ou leur disponibilité. Le problème devient encore plus délicat si un utilisateur souhaite « sortir »un document du système d'information pour, par exemple, y travailler en mode déconnecté ou le diffuser à d'autres personnes extérieures à l'entreprise. En nous inspirant des approches orientées objets, nous avons choisi d'encapsuler dans les documents des composants de sécurité (contrôle d'accès, contrôle d'usage,...) pour aboutir à une architecture orientée documents intelligents pour les DRM [24] d'entreprise (E-DRM).

Les nouvelles technologies (ADSL, ordinateurs portables, smartphones, tablettes,...) nous ont en effet fourni de nouveaux moyens techniques pour communiquer mais nous ont également créé de nouveaux besoins. Par exemple, l'information doit être accessible à tout instant (éventuellement en mode déconnecté) ; le recours à un site centralisé pour les échanges est vu comme une contrainte (architecture client-serveur vs. architecture pair-à-pair) ; les données sont stockées sur des clés usb et sont consultées sur des ordinateurs en libre accès ; nous emportons nos données dans nos smartphones et les partageons via des communications sans fil (3G, wifi, bluetooth). Ces nouvelles pratiques ne sont pas sans soulever quelques problèmes liés à la sécurité de l'information.

Outre les aspects techniques de ces échanges d'information, ce sont également leurs contenus qui ont évolué. Les données sont de plus en plus complexes (notions de document structuré, d'archive, voire même de projet complet). Des données dites publiques cohabitent maintenant avec des données plus confidentielles (notion de restriction d'accès). Le fournisseur de contenu désire parfois contrôler la façon dont l'usager va manipuler ce contenu (notions de contrôle d'usage, de gestion des droits numériques ou DRM). Cela peut concerner tant la consultation

24. DRM : Digital Right Management

des informations (contrôle d'accès, présentation des données) que leur modification (contrôle d'accès, gestion de la cohérence des données, enregistrement et/ou création de métadonnées, gestion des révisions ou traçabilité).

4.3 Les thèmes abordés

Dans le domaine du travail collaboratif, les différents partenaires sont amenés à consulter mais également à modifier les documents. Il est donc nécessaire de mettre en place des mécanismes de sécurité qui vont au-delà d'un simple contrôle d'accès : contrôle d'usage (règles d'utilisation d'un document : obligations, workflow, règles de délégation), maintien de la cohérence des informations (ex : certains documents peuvent en référencer d'autres), traçabilité (suivi des actions, métadonnées attachées aux informations), etc... C'est le rôle d'un Système d'Information qui constitue le cœur de l'entreprise ou du projet.

4.4 Les propositions

Notre idée consiste donc à encapsuler à l'intérieur du document lui-même à la fois les informations qu'il véhicule et des mécanismes de contrôle de la sécurité. Voici un aperçu des principaux composants avant de les détailler dans les sections suivantes :

- la base de données pour stocker le contenu du document ainsi que les métadonnées ; les informations sont structurées et sont représentées par un graphe de nœuds et de relations ; les métadonnées sont des propriétés attachées à ces nœuds et à ces relations
- le noyau de sécurité qui est chargé de faire respecter la politique de sécurité et contrôle donc toutes les actions effectuées sur le document ; il repose pour cela sur différents modules de sécurité dédiés à des tâches spécifiques (ex : contrôle d'accès et contrôle d'usage basé sur le modèle OrBAC, gestion des métadonnées, calcul d'indicateurs,...)
- des services et/ou des applications embarqués pour manipuler le document avec des outils dédiés (ex : mécanismes d'exportation/importation, applications métiers, services de présentation,...)
- la licence de l'utilisateur (stocké indépendamment du document) contenant, entre autres, les permissions, les interdictions et les obligations attribuées à l'utilisateur ; à terme, cette licence contiendra également différentes informations permettant de configurer le noyau de sécurité (et ses modules) telles que la description des métadonnées à collecter ou les règles de gestion des contextes.

Pour illustrer le sujet nous prenons comme exemple un projet du domaine Oil & Gas comme la construction d'un pipeline pour une installation gazière : c'est le type de projet sur lequel

travaille la société BackPlan™ . Le système d'information consiste en de nombreux documents qui ont un rôle central. Leur structure et leur contenu évoluent tout au long du projet : la documentation doit toujours précéder le travail (dessin, procédures de travail), de plus elle est requise à la clôture du projet. La documentation évolue ensuite avec le projet.

Ces documents peuvent être les spécifications, les dessins, les expertises, les procédures...

4.5 Conclusions

L'architecture des documents autoprotégés présentée dans cet article est actuellement au stade de prototype. Nous avons pu représenter différents types d'informations : arbre de fichiers et de répertoires, documents structurés en XML (noeud détaillée par noeud) , tous les fichiers dans un projet avec leurs relations de dépendance. Cela nous a permis de valider notre modèle d'entrepôt de données basé sur une approche multi-vue.

4.6 Texte de l'article

2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications

Self-Protecting Documents for Cloud Storage Security

Manuel Munier, Vincent Lalanne
LIUPPA
Univ Pau & Pays Adour
Mont de Marsan, France
Email: {manuel.munier, vincent.lalanne}@univ-pau.fr

Magali Ricarde
BackPlan Company
Project Communication Control
Pau, France
Email: magali.ricarde@backplan.fr

***Abstract*—Information security is currently one of the most important issues in information systems. This concerns the confidentiality of information but also its integrity and availability. The problem becomes even more difficult when several companies are working together on a project and that the various documents "go out of" their respective information systems. We propose an architecture in which the documents themselves ensure their security and thus can be exchanged over uncontrolled resources such as cloud storage or even USB flash drives. For this we encapsulate within the document itself some security components (e.g. access control, usage control) to achieve an autonomic document architecture for Enterprise DRM (E-DRM). Using such self-protecting documents, a company can ensure security and privacy for its documents when outsourcing storage services (e.g. cloud).**

***Keywords*- self-protecting document; autonomic document; usage control; E-DRM; cloud storage security;**

I. Introduction

Information security is currently one of the most important issues in information systems. Security criteria most commonly used are confidentiality (assurance that information is shared only among authorized persons or organizations), integrity (assurance that the information is authentic and complete), availability (assurance that the systems responsible for delivering, storing and processing information are accessible when needed, by those who need them) and traceability (ability to chronologically interrelate uniquely identifiable entities in a way that is verifiable). The problem becomes even more difficult if a user wants to export a document from the information system to work offline, for example, or to broadcast it to other persons outside the company. Drawing on the object-oriented approaches, we chose to encapsulate in the document itself some security components to achieve an autonomic data management architecture for Enterprise Digital Rights Management (E-DRM).

New technologies (ADSL, laptops, smartphones, tablets,...) have provided us with new technical means to communicate but have also created new needs. For example, information must be accessible at any time (possibly offline); using a centralized site for the exchange is seen as a constraint (client-server architecture vs. peer-to-peer architecture); data are stored on USB flash drives and used on untrusted computers (available on open access); we carry our data in our smartphones and share them via (possibly unsecured) wireless communications like 3G, wifi or bluetooth; cloud storage provides a platform for connected devices (laptops, smartphones, tablets,...) to access data without the need to store it locally on the device. Obviously, these new practices raise some issues related to information security.

Besides the technical aspects of these exchanges of information, their contents have changed. Data are more and more complex (notions of structured documents, whole archives, or even complete projects). Nowadays, public data are sometimes combined with more confidential data (notion of access restriction). The content provider may wish to control how the user is handling the content (concepts of usage control, DRM). This can cover both the consultation of information (access control, data presentation) and its modifications (access control, management of data consistency, logging and/or creation of metadata, update management).

To ensure information security companies typically deploy an information system responsible for providing a network storage and access (and possibly usage) controls. The implementation of such a platform, however, requires certain resources (human and technical). Companies are sometimes tempted to outsource the storage of their data: cloud, data center,... But one important consideration regarding cloud storage is security. This has also been one of the major obstacles to a larger adoption of cloud storage: clients aren't likely to entrust their data to another company without a guarantee that they'll be able to access their information whenever they want and no one else will be able to get at it. To secure data, most systems use techniques like encryption (they use a complex algorithm to encode information; to decode the encrypted files, a user needs the encryption key), authentication processes (the system requires to create a user name and password), authorization practices (the client lists the people who are authorized to access information stored on the cloud system; many corporations have multiple levels of authorization). Even with these protective measures in place, many people worry that data saved on a remote storage system is vulnerable.

978-0-7695-4745-9/12 $26.00 © 2012 IEEE
DOI 10.1109/TrustCom.2012.261

There's always the possibility that a hacker will find an electronic back door and access data. Works exist about security issues in cloud storage [1], [2], [3], [4], or more generally in cloud computing [5], [6], [7].

In this paper we propose an innovative approach that aims to make documents self-secure so they can be stored even on untrusted servers: we encapsulate within the document both the data it carries and the security mechanisms to control the use of such data. Thus security constraints can be (partially) relaxed on the server since all information is encrypted and only the security kernel of the document can access them. Using such self-protecting documents, a company can ensure security and privacy for its documents when outsourcing storage services (eg cloud).

The remainder of this paper is organized as follows: Section II presents our motivations for the design of a new E-DRM architecture based on autonomic documents; this section also presents a concrete case study that we developed with our partner company BackPlan™ [1]; in Section III we describe our autonomic document approach and we give some details about the metadata and the usage control mechanisms; Sections IV and V outline our current work about cloud storage and security issues, and then conclude this paper.

II. Motivations

In the area of collaborative work, the partners are required to consult but also to modify the documents. It is therefore necessary to implement security mechanisms that go beyond a simple access control: usage control (how partners can use a document: obligations, workflows, delegation rules,...), information consistency management (e.g. some documents may reference others), traceability (monitoring of actions, metadata attached to information),... This is the role of the Information System which is the core of the company or project.

A. Architecture

However, centralized architectures for information sharing (Fig.1) have two drawbacks compared to the use of traditional documents:

1) Work on protected documents requires the user to connect to a server responsible for enforcing the security policy. This server hosts all the documents and controls all user actions. Users can not directly exchange information: all the exchanges have to go through this server.
2) These security mechanisms require installation, on the user side, of applications and/or proprietary modules to be able to handle such documents. The user becomes dependent on the security provider for his/her applications.

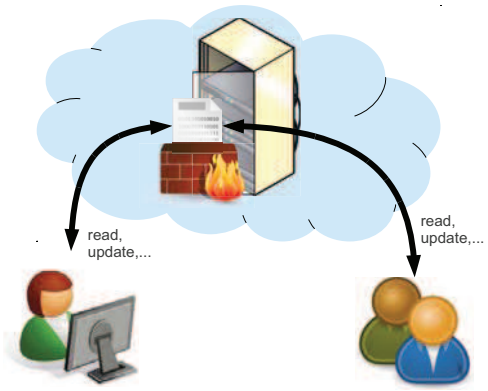


Figure 1. Document security enforced on server side

These are the two findings that led us to propose a new E-DRM architecture inspired by object-oriented concepts where the document becomes an autonomous operating entity able to control by itself how the information it contains can be accessed and used (Fig.2). Such a document is a kind of information system on its own embedding both a data warehouse and various security modules (access control, usage control, metadata,...).

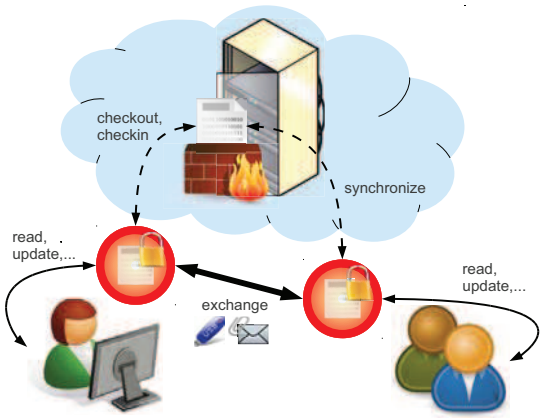


Figure 2. Autonomic document self-manages its security

This allows users to exchange such documents safely, or even portions of documents. Moreover, security constraints are relaxed on the server side since the documents themselves provide their own security. The server will only be used to synchronize versions of documents exchanged by partners.

B. Sample Application

Consider an Oil & Gas project as the construction of a pipeline or an oil installation. The information system

[1] BackPlan™ , Project Communication Control
http://www.backplan.fr

consists of numerous documents, it has a central role, its structure and development evolve along with the progress of the project: the documentation must always precede action (design, work procedures). Documentation is a requirement at the closure of the project along with verifications (records, the minutes,...). The document evolves at the same rate as the project. These documents are specifications, drawings, records of expertise, procedures, records,...

1) Business Aspects: Such a project obviously involves many partners and subcontractors. Here is a representative example of a project timeline:

- 1st level design: this step is performed by the land surveyor who makes a topographical survey of the site where the work will be done.
- 2nd level design: is the design phase of the project itself, and is performed by an engineering company that will plan the work and the construction will be launched from this plan. This level involves various partners: civil engineering, pipefitters, instrumentation engneering, utilities,... These documents are usually schemes that can have many levels or layers, each company or sector can thus share the same document. This stage ends with all engineering documents validated by the stamp "Approved for construction".
- Construction phase: it is based on engineering documents and work procedures. It comes with many documents that are intended to demonstrate compliance of the book in terms of quality and regulatory standards (e.g. multifluid standard, water code, capacity under pressure).
- At the end of construction the land surveyor will come again, and verify the topographic survey: this is a control operation of a project to verify the differences from the planned location and update the data (to know where everything is). The engineering documents then pass status "As built". This operation can also update the geographic information system (GIS) of the place.

As we have stated, the partners will handle many documents. There can also occur unforeseen circumstances during the project. Here are some identified difficulties and problems:

1) Several actors in various fields must work on the same document. Take for example the passage of a pipe under a river: the land surveyor made the survey, but it's not him who knows the techniques of drilling, centering techniques and the radii of curvature of the pipe. These characteristics are a function of pipe diameter, its alloy or its profile,... In 2nd level study partners will therefore have to refine the study of a 1st level.
2) The recordings are very numerous during the construction. For example: inspection of welds. The sections of pipe are welded every 12 to 15m. These welds should be checked: radiography, analysis by a certified individual, hydraulic tests. These controls are spread over time and generate many records that are, once made, a legal value.
3) The code requirements to be applied are numerous and specific whatever the project: standards and regulations, good practice guides, internal standards,... As part of the quality approach, journals are provided to verify that these code requirements are taken into account. It is recommended to involve the regulatory bodies in these reviews. Unfortunately, these documents are currently independent of each other (in terms of information system).
4) Consequences for digital documents: whether for office or mobile use (eg tablet or laptop on a site with no network coverage), the dependence on the document registry can be restrictive (even by a checkout of files for offline viewing). It would be more convenient to group all the information in one document, structured and secure, so that different partners can exchange it directly (and synchronize it occasionally with a server to "publish" their work).

2) Security Aspects: We propose to improve the security of information in two directions: metadata management and usage control.

Metadata: It could be used to "bind" reviews, certifications, and other minutes to design documents within the information system (see item 3 above). The aim is to improve traceability, both in the design process (concept of workflow) in case of litigation (concept of proof of conformity, digital forensics). Take for example a phase control such as checking the welding of a pipe (see item 2 above), it would be interesting to use metadata to improve the traceability of the process for purposes of validation and/or evidence: photos geotagged (to certify checkpoints), metadata associated with the plans,... Since several partners are working on such a project, everyone could also attach some metadata to the information: confidence and trustworthiness indicators, impact risk of a change,...). This metadata would permit to calculate various performance indicators for monitoring the partners' tasks or metadata to the information they produce.

Usage control: Here are some examples of security rules that we would implement to control how partners use documents:

- It is possible to write a deliverable of the project only if confidence in the various technical documents exceeds a certain threshold. It is a dynamic access control based on trust (whether in a document or a partner).
- The security rules may prohibit access to parts of the document based on location data. This prevents, for

example, on a site (or train) an unknown person takes sensitive information over the shoulder of someone. The same argument could be used for putting visual filters in place (blur, masking, wrong information,...).

- A responsive access control may require a partner (or subcontractor), via a mechanism of pre-obligation to accept the terms of a contract (non-disclosure agreement, delegation of responsibility, deadlines,...) before accessing a plan and contribute to the design.
- A user control would be to require a partner to complete parts of design documents (eg. inform the radii of curvature of the pipe, write the study of soil before drilling) before a deadline if he wants to stay a project member (notions of punishment and penalty).
- The usage control can also define collective obligations. For example, each partner must have reread each concept study in which they participate at least 7 days before the deadline for validation.

Currently, the various documents of such a project are managed independently of each other. In the best case, a document registry is set up to try to centralize critical information. The use of a true information system would allow the implementation of such a security policy. However, users now want to handle these documents on their laptops, smartphones, or tablets, and using a centralized site for the exchange is seen as an unacceptable constraint. Hence our idea of developing self secure documents that can be stored anywhere on the cloud or exchanged with USB flash drives.

C. *BackPlan™*

BackPlan™ is a French company providing document management services and collaboration workflow applications to improve project communication, transparency across the project, ability to manage schedule and risks, reliable indicators and regulatory compliance. From the engineering phases to the construction phases, projects involve different companies. All of them will use the collaboration solution BackPlan™ to ensure consistency of information across the project and a complete audit trail of project communication. BackPlan™ document management services are currently provided on a server hosted in a data center: the document registry.

However, as explained in the case study in section II-B, the solution of a centralized information system is not entirely satisfactory. Whether for office or mobile use (e.g. smartphone, tablet or laptop on a site with no network coverage), the dependence on the document registry can be restrictive (even by a checkout of files for offline viewing). It would be more convenient to group all the information in one document, structured and secure, so that different partners can exchange it directly (and synchronize it occasionally with a server to "publish" their work).

III. Autonomic Documents

Our approach is therefore to encapsulate within the document both the data it carries and the security mechanisms to control the use of such data. The architecture we propose for such autonomic documents which can manage their security autonomously is outlined in Figure 3. Here are the most significant components that will then be detailed in the following subsections:

- the database for storing the contents of the document and metadata; structured contents are represented in terms of nodes and relationships; metadata are properties attached to these nodes and relationships
- the security kernel which is responsible for enforcing the security policy and then monitors all actions on the document; it relies on various security modules dedicated to specific tasks (OrBAC for access control and usage control, metadata management, computation of indicators,...)
- embedded applications to operate on the document with dedicated tools; embedded services and/or mechanisms of export and import to handle the document using legacy applications
- the license for the user (stored outside the document) containing the permissions, prohibitions and obligations assigned to this user; later, such a license may also contain other directives for the kernel (and its modules), such as metadata to be collected or context management rules

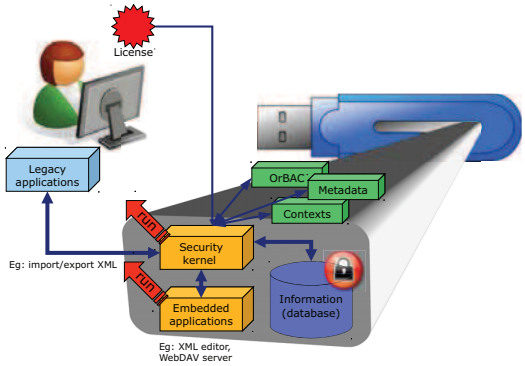


Figure 3. Autonomic document architecture

When the user wants to access the secure document he must first identify himself and provide his license which contains the security policy rules specific to that user. These "rules" may represent permissions, but also prohibitions, obligations, or metadata to be collected during the user's actions.

The document then starts the security kernel of our architecture which will in turn initiate various security modules:

a access control module, a metadata management module, a trustworthiness management module, a module dedicated to collaborative work,... The kernel and the modules are configured with the rules contained in the license provided by the user.

From that moment on, all the accesses to information within the document is made through the security kernel. For any action performed by the user, the kernel requests each security module to validate this action. Some modules will indeed confirm or deny the various actions (e.g. access control module). Others will simply add some metadata to these actions (e.g. who performed this action, from which IP, at what time, with which application, in which context,...) so that other modules can compute their data (e.g. trustworthiness indicator, collaborative work management).

The main difference to "classic" E-DRM schemes is how and where the security mechanisms are implemented. Traditionally, the document (or its subset that is accessible to the user) is encrypted using player keys; decryption takes place in a trusted viewer, which has to run securely on the client side. The trusted viewer is responsible for enforcing a license distributed alongside the document. If the user updates the document, it is generally necessary to publish the amended document on the server (Fig.1) so that another user can access and see the changes (the server must reencrypt the amended document). In the proposed architecture, the document (as an object embedding both data and security modules) is more like a decentralized part of the whole information system. It can manage the updates made by users and store these changes in its embedded encrypted database. The document can be transferred directly to another user without going through a server (Fig.2).

The second difference is that access to the document can be done either through dedicated embedded applications (like "lightweight" trusted viewers), or "heavy" trusted applications (using the API), or by exporting data (e.g. XML format) to some legacy applications.

A. Information Storage

Document data are stored in the embedded database accessible only by the security kernel. This store implements the multi-view model for secure versioned repository as defined in our previous work [8]. In this model, information is organized as a graph of which the nodes are connected by relations of various kinds. We can represent XML documents, trees like files/directories structures, project files when certain items are dependent on other elements. Each version of each node is kept with its own relationships. The view of a user is a subset of this graph computed according to his/her privileges. All user actions (accessing, updating, creating, deleting,... nodes and/or relationships) are performed with respect to his/her view. For instance, if he/she deletes a folder that contains "hidden" files (i.e. not available in his/her view), the execution of this action should result in the removal of this directory in the view of the user. This to ensure the confidentiality property for other information. Refusing the removal because "the directory is not empty" violates this property. Accepting the removal and deleting, as a side effect, an "unauthorized" file also presents a problem of confidentiality. The solution we adopted is to translate each action into elementary operations at the level of the data warehouse in terms of updates of versions of those nodes and their relationships. This is to ensure the integrity of the information. For instance, the "hidden" file that belonged to the deleted directory should not become an orphan. Since the data warehouse stores all the versions, a user who has access to all files in the previous directory could still access the file through the tree, but would also be informed that the parent directory has been "deleted" by another user (along with some files it contained).

B. Security Modules

Obviously the core of our architecture is the security kernel. It is the document interface with the outside world: all the actions performed by the user to handle the document need to be done through the security kernel (like data abstraction in object-oriented programming). From a functional point of view it has two aims: (1) translate the actions performed by the user about his/her view into basic operations on the data warehouse; (2) check that the user's actions are in accordance with the security policy.

The first task is therefore to translate user actions into basic operations at the level of versions of nodes and their relationships. We invite you to consult [8] for more details on this section. For example, to update a node, we create in fact a successor to the version contained in the view of the user. Deleting a node is equivalent to an update with the special value `NULL`. Moving a node results in the creation of a new version (link in succession) but with new relationships.

In this article we focus on the implementation of the security policy. As we stated at the beginning of Section III, the security kernel can be configured to use various security modules. For now we identified three categories of modules as they are responsible for:

- accepting or rejecting the user's actions,
- collecting and attaching metadata to the actions,
- calculating new data as actions go along.

When the user requires the execution of an action, the security kernel performs control in two stages. The first is to validate this action. For this the kernel requests each security module to validate the action. Some modules will indeed accept or reject the action (e.g. access control, usage control). Others will add information to this action (e.g. metadata). If the action is validated by all the security modules it then enters the second stage: the processing of the action. Basic operations implementing this action are then

performed on the data warehouse. Then, the security kernel broadcasts this action a second time to each security module so they can achieve the associated processing: logging (e.g. access control, usage control), computation of additional information (e.g. trustworthiness management, collaborative work management).

The first security module that we developed belongs to the first category and relates to the access and usage control to information contained in the document. From the formal point of view we are using the OrBAC model [9] which can express security policies in terms of permissions, prohibitions and obligations between a subject, an object and an action. These are the rules contained in the license provided to the user. OrBAC model also supports the notion of context [10], [11], [12]. Indeed, security rules in these policies are no longer static but dynamic, depending on the context. They must be also self-adaptive with respect to temporal conditions, the user's location, previous behavior of this user, etc... Many works attempt to extend existing models to deal with access control based on user's location context [13] or temporal context [14]. Some models are specifically dedicated to collaborative environments (in [15] users obtain permissions according to their role and the team in which they are involved) or workflow environment (in [16], [17] subjects gain access to required objects only during the execution of the task). As discussed in [11], authors use contexts in OrBAC to express different types of extra conditions that control activation of rules specified in the security policy:

- the *Temporal context* that depends on the time at which the subject is requesting for an access to the system,
- the *Spatial context* that depends on the subject location,
- the *User-declared context* that depends on the subject objective (or purpose),
- the *Prerequisite context* that depends on characteristics that join the subject, the action and the object.
- the *Provisional context* that depends on previous actions the subject has performed in the system.

Our next step was therefore to develop a second security module to manage contexts in our architecture. By adding context awareness to our intelligent document we can express contextual usage control rules in security policies. To control context activation, the embedded information system must provide the information required to check that conditions associated with the context definition are satisfied or not. To do this, either contexts have direct access to the host system (eg a global clock to check the temporal context) or they use metadata carried by the actions.

A third security module has been developed and relates precisely to metadata collection (which performs the action, at what moment, from which IP,...) and thus ensures the traceability of actions performed on the document by the different users. It is not yet configurable because we first need to define, at a license level, a language to describe the metadata logging policy. However, collecting some predefined metadata, we were able to test and validate context management and, thus, the implementation of security rules to do usage control. One of our current work is to implement a security module using these metadata to compute new information (e.g. trustworthiness evaluation for information updates as in [18], [19], [20]). Such metadata may also be useful for providing collaborative work support as discussed in [21], [22]. To do this we could either write a security module either define a (provisional) context. This remains an open question for now.

C. Opening & Using a Document

As we have previously mentioned, access to information contained in the document can then be done in three ways:

- Export and import mechanisms (XML for example) to manipulate information through existing applications. This requires setting up filters at the security kernel level to format information when exporting (check-out) and to interpret them when importing (check-in). In this case, the level of granularity is the whole file.
- Plugins developed for existing applications to have a more granular level. The plugin can then talk directly with the security kernel to interact at the nodes and relationships level, provided of course that the user has the appropriate privileges.
- Use of services and/or dedicated application embedded in the secure document. During its initialization, after starting the different security components, the document can automatically start running a WebDAV server (for example) in charge of presenting the information in the document as a tree of files/directories. Access to information can then be made from traditional applications via the WebDAV server.

D. License Contents

The last component of our architecture deals with licenses for users. As shown in Figure 4, a license is an XML document containing information of the server that issued the license (the licensor), information relating to the licensee and, of course, various security rules. Having used the OrBAC model to implement our security module in charge of access control and usage control, it is in this example OrBAC rules. Compared with policy language standards like XrML or ODRL, we also need in the future to express how metadata should be collected and what triggers must be deployed to manage contexts. Thus the license does not consist only of security rules but also contains the configuration of the various security modules required for handling the document.

To protect the contents of this license, when created, the rules are encrypted with the user's public key (being the only

one who can decrypt with his/her private key and thus the only person to know, for example, the permissions which are granted to him/her). The license is then signed by the issuer to prevent any further modification.

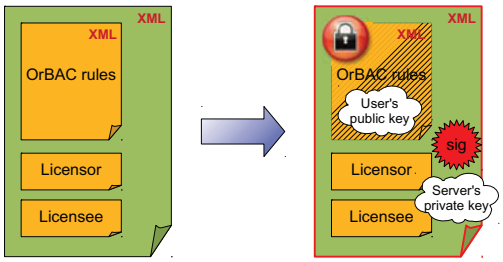


Figure 4. License contents

As we mentioned since the beginning of this article, our approach relies on the concept of metadata for traceability, context activation (dynamic security rules) and, eventually, computation of indicators (collaborative work management). Some of this metadata can however raise privacy issues: user's location, trustworthiness indicator,... One of our perspectives is to define in the license metadata to be collected and for what purposes it will be used. From a legal standpoint, the license will be a user agreement to work on the document.

IV. Cloud Storage

To build our first prototype of autonomic document based E-DRM platform we used standard USB flash drives with an auto-run configuration to launch Java programs. The flash drive represents the document and can be exchanged (physically) between different users. It brings together on the same "support" a database and several executables (security kernel, security modules, embedded services and applications). Such an architecture can run on various operating systems (Microsoft Windows, Linux,...). The only assumptions about the host computer are the availability of a Java virtual machine and enabling the auto-run for removable media. As the user can directly access files on the USB flash drive, and thus in particular to storage files of the database, we encrypted the content of the latter. Thus "raw" information of our document, which might be reached in bypassing the security kernel, is unusable.

This prototype is primarily a proof of concept. The details of "internal" mechanisms such as secret credentials embedded in the security kernel, correct execution of the kernel, ciphering used (embedded database, license contents, communications between the document and external applications,...) are beyond the scope of this paper. We obviously have implemented these mechanisms to experiment our approach, but this part of our work requires further study (e.g. using ISO/IEC 27005:2011 information security risk management approach [23]).

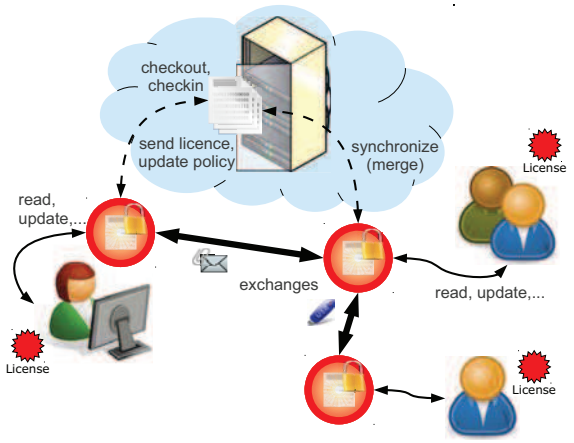


Figure 5. Self-protecting documents in cloud storage

The next step is to develop a storage server on the cloud (Figure 5). This means specifying primitives read (export of a full project document or portion of a document) and write (merging changes made by the partners). This server will also be responsible for distributing licenses and pushing any updates to the security policy. In fact, as with USB flash drives, security constraints can be (partially) relaxed on the server since all information is encrypted and only the security kernel can access them to merge the updates. Using such self-protecting documents, a company can ensure security and privacy for its documents even when outsourcing storage services (e.g. cloud).

V. Conclusions

The architecture of self-protecting documents presented in this paper is currently in the prototype stage. We were able to represent various types of information: tree of files and directories, structured documents in XML (detailed node by node), all files in a project with their dependency relationships. This allowed us to validate, at first, our data warehouse model based on a multi-view approach.

We integrated the OrBAC engine and using metadata and contexts we are currently implementing new security modules specifically dedicated to usage control (e.g. trustworthiness evaluation for information updates, traceability of changes, collaborative work management, schedule and risks management,...). For this we use existing works found in [19], [10], [24], [25], [21].

As explained in section II-B many projects involves several companies, and these companies are increasingly faced with the following dilemma: they want to strengthen controls over the use of their documents, but these documents must also be "dispersed" among users and therefore outside

the information system of the company. We are confident that the self-protecting document architecture presented in this paper is a solution to these new needs. It allows a user to forward the document to another user who handles the document according to his own license. It is no longer necessary to have a centralized server to handle static documents. We only need a drop point (e.g. outsourced cloud storage service) to synchronize the various autonomic documents.

REFERENCES

[1] X. Zhang, H.-t. Du, J.-q. Chen, Y. Lin, and L.-j. Zeng, "Ensure data security in cloud storage," *2011 International Conference on Network Computing and Information Security*, pp. 284–287, 2011.

[2] J.-J. Hwang, H.-K. Chuang, Y.-C. Hsu, and C.-H. Wu, "A business model for cloud computing based on a separate encryption and decryption service," *Information Science and Applications, International Conference on*, vol. 0, pp. 1–7, 2011.

[3] T. Mather, S. Kumaraswamy, and S. Latif, *Cloud Security and Privacy*, M. Loukides, Ed. O'Reilly, 2009, vol. 1, no. 11.

[4] J. Wu, L. Ping, X. Ge, Y. Wang, and J. Fu, "Cloud storage as the infrastructure of cloud computing," *2010 International Conference on Intelligent Computing and Cognitive Informatics*, pp. 380–383, 2010.

[5] B. Grobauer, T. Walloschek, and E. Stocker, "Understanding cloud computing vulnerabilities," *IEEE Security Privacy Magazine*, vol. 9, no. 2, pp. 50–57, 2011.

[6] K. Dahbur, B. Mohammad, and A. B. Tarakji, "A survey of risks , threats and vulnerabilities in cloud computing," *Computing*, pp. 1–6, 2011.

[7] F. Sabahi, "Cloud computing security threats and responses," *Computer Engineering*, pp. 245–249, 2011.

[8] M. Munier, "A multi-view approach for embedded information system security," in *CRiSIS*. IEEE, 2010, pp. 65–72.

[9] A. A. E. Kalam, S. Benferhat, A. Miège, R. E. Baida, F. Cuppens, C. Saurel, P. Balbiani, Y. Deswarte, and G. Trouessin, "Organization based access control," in *POLICY*. IEEE Computer Society, 2003, pp. 120–131.

[10] Y. Elrakaiby, F. Cuppens, and N. Cuppens-Boulahia, "From contextual permission to dynamic pre-obligation: An integrated approach," in *ARES*. IEEE Computer Society, 2010, pp. 70–78.

[11] F. Cuppens and N. Cuppens-Boulahia, "Modeling contextual security policies," *Int. J. Inf. Sec.*, vol. 7, no. 4, pp. 285–305, 2008.

[12] F. Cuppens and A. Miège, "Modelling contexts in the or-bac model," in *ACSAC*. IEEE Computer Society, 2003, pp. 416–427.

[13] M. L. Damiani, E. Bertino, B. Catania, and P. Perlasca, "Geo-rbac: A spatially aware rbac," *ACM Trans. Inf. Syst. Secur.*, vol. 10, no. 1, 2007.

[14] E. Bertino, P. A. Bonatti, and E. Ferrari, "Trbac: A temporal role-based access control model," *ACM Trans. Inf. Syst. Secur.*, vol. 4, no. 3, pp. 191–233, 2001.

[15] C. K. Georgiadis, I. Mavridis, G. Pangalos, and R. K. Thomas, "Flexible team-based access control using contexts," in *SACMAT*, 2001, pp. 21–27.

[16] D. G. Cholewka, R. A. Botha, and J. H. P. Eloff, "A context-sensitive access control model and prototype implementation," in *SEC*, ser. IFIP Conference Proceedings, S. Qing and J. H. P. Eloff, Eds., vol. 175. Kluwer, 2000, pp. 341–350.

[17] E. Bertino, E. Ferrari, and V. Atluri, "The specification and enforcement of authorization constraints in workflow management systems," *ACM Trans. Inf. Syst. Secur.*, vol. 2, no. 1, pp. 65–104, 1999.

[18] E. Bertino and H.-S. Lim, "Assuring data trustworthiness - concepts and research challenges," in *Secure Data Management*, ser. Lecture Notes in Computer Science, W. Jonker and M. Petkovic, Eds., vol. 6358. Springer, 2010, pp. 1–12.

[19] Xingyu Zheng and Patrick Maillé and Cam Tu Phan Le and Stephane Morucci, "Trust mechanisms for efficiency improvement in collaborative working environments," in *MASCOTS*. IEEE, 2010, pp. 465–467.

[20] C. T. P. Le, F. Cuppens, N. Cuppens-Boulahia, and P. Maillé, "Evaluating the trustworthiness of contributors in a collaborative environment," in *CollaborateCom*, ser. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, E. Bertino and J. B. D. Joshi, Eds., vol. 10. Springer, 2008, pp. 451–460.

[21] M. Munier, K. Benali, and C. Godart, "Discoo, a really distributed system for cooperation," *Networking and Information Systems*, vol. 2, no. 5-6, pp. 605–637, 1999.

[22] M. Munier, K. Benali, and C.Godart, "A transactional approach for cross-organizational cooperation," in *GlobeComm99 (Enterprise Applications and Services Symposium)*, december 1999.

[23] ISO/IEC, "ISO/IEC 27005:2011: Information technology — security techniques — information security risk management," International Organization for Standardization, Geneva, Switzerland, Published, 2011.

[24] M. Ben-Ghorbel-Talbi, F. Cuppens, N. Cuppens-Boulahia, and A. Bouhoula, "A delegation model for extended rbac," *Int. J. Inf. Sec.*, vol. 9, no. 3, pp. 209–236, 2010.

[25] J. Thomas, F. Cuppens, and N. Cuppens, "Environmental constraints management in digital right licences," in *SARSSI 2008 : 3ème Conférence sur la Sécurité des Architectures Réseaux et des Systèmes d'Information, 13-17 octobre, Loctudy, France*, Publibook, Ed. 75015 Paris: RSM - Dépt. Réseaux, Sécurité et Multimédia (Institut TELECOM ; TELECOM Bretagne), 2008, pp. 85–99.

Chapitre 5

PASSAT 2013
Information Security Risk Management in a World of Services

Sommaire

5.1 Introduction

Cet article a été présenté à la conférence internationale sur la vie privée, la sécurité, le risque et la confiance (PASSAT 2013 [25] en septembre 2013 à Washington DC - USA ; Taux d'acceptation de 9,6 %).

Il se place dans un contexte lié aux révélations faites par Edward Snowden en juin 2013 où ont été rendues publiques des informations classées top-secrètes de la NSA [26] sur les pratiques de captations, d'intrusions et d'écoutes mises en place par les États Unis. ***Remarque :*** *ces révélations ont eu lieu seulement deux jours avant la date de soumission de cet article.*

5.2 Présentation

Les architectures orientees services (SOA) offrent de nouvelles possibilites pour l'interconnexion des systemes d'information. L'ouverture de son système d'information pour une entreprise vers l'exterieur n'est toutefois pas anodine du point de vue de la securite. Que ce soit pour utiliser des services proposes par des tiers ou pour offrir les siens, ces technologies introduisent de nouvelles vulnerabilites et, par consequent, de nouveaux risques.

Nos travaux visent a initier une demarche de gestion de ces risques qui s'appuie sur un standard, la norme ISO/IEC 27005 :2011. Nous proposons une evolution de cette norme afin qu'elle puisse prendre en compte pleinement le type « service ». Suite a cette etude nous introduisons egalement un nouveau critere, la maitrisabilite, pour qualifier la securite des systemes d'informations.

5.3 Les thèmes abordés

Dans un contexte actuel d'interconnexion des systèmes, les architectures orientées services (SOA) offrent de nouvelles possibilités massives d'échanges et de traitement des données. Or l'ouverture du système d'information d'une entreprise sur l'extérieur n'est toutefois pas anodine du point de vue de la sécurité. Que ce soit dans l'utilisation des services proposés par les tiers que lorsque l'on offre les siens, ces technologies introduisent de nouvelles vulnérabilités dans le système d'information et, par conséquent, de nouveaux risques.

La première partie est consacrée à un tour d'horizon rapide des technologies qui mettent en œuvre les services Web : ces technologies vont des mécanismes de base (REST, SOAP) jusqu'aux

25. The 5th ASE/IEEE International Conference on Information Privacy, Security, Risk and Trust - http ://www.asesite.org/conferences/passat/2013/index.html
26. National Security Agency

spécifications plus complexes garantissant la sécurisation, la confiance, le contrôle d'usage et la vie privée (WS-security...).

Dans un second temps, ce sont les normes et méthodes employées dans la gestion des risques en sécurité de l'information qui sont succinctement décrites : un détail particulier est porté sur la norme ISO/IEC 27005:2011.

Dans une troisième partie cette norme est présentée à travers la mise en place d'architectures orientées services :

— sous un **aspect technologique** : identification des menaces et vulnérabilités liées à la technologie,

— sous un **aspect sécurité de l'information** : l'identification des conséquences lorsque les données sont confiées à des prestataires extérieurs, celles-ci étant observées d'un point de vue de la propriété de l'information, de la confidentialité, de l'intégrité et de la disponibilité.

5.4 Les propositions

Nos travaux visent à initier une démarche de gestion de ces risques en s'appuyant sur un standard : la norme ISO/IEC 27005:2011. L'utilisation de cette norme nous paraît absolument indispensable, en effet en tant que standard, elle a le mérite d'être reconnue internationaux et peut s'appliquer sur tous les systèmes d'information.

Nous proposons donc une évolution de cette norme afin qu'elle puisse prendre en compte pleinement le type « service »dans la gestion des risques d'un système d'information. Lors de sa création en 2008, les services Web existaient mais bien souvent sous forme propriétaire et/ou privée. Or maintenant, avec la généralisation et la commercialisation de nombreux services de type « CLOUD », il nous semble absolument indispensable de se situer à un niveau de granularité supérieur à celui de la technologie pour se concentrer sur la sécurité même de l'information. En effet, quel que soit les verrous technologiques employés, il ne sera jamais possible d'aller à l'encontre d'une décision de justice imposant à un fournisseur d'accès d'ouvrir l'accès à ses serveurs !

Aussi, afin de tenir compte de ce risque dans l'étude d'un système d'information, nous proposons d'étendre un nouveau type d'actif au sein du système d'information qui serait dénommé « services ».

Suite à cette proposition nous introduisons également un nouveau critère de maîtrisabilité à côté des traditionnels critères de confidentialité, d'intégrité et de disponibilité. Ce critère de maîtrisabilité permettra de qualifier la manière dont est contrôlée l'utilisation des services Web dans le système d'information.

5.5 Conclusions

À travers ces constats et propositions, nous alertons d'abord la communauté sur les risques d'utilisation des services Web (ou cloud) dans un environnement international où les lois et règlements peuvent être différents : protégeant l'information dans un pays, permettant la consultation pour diverses raisons dans un autre.

Grâce à l'ajout d'un critère de maîtrisabilité, nous permettons ainsi à l'architecte du système d'information de qualifier la confiance qu'il a dans le ou les services qu'il utilise.

Nous proposons également l'utilisation de Documents intelligents (Trustcom 2012) qui peuvent, avec un moindre risque, être stockés et transiter via des services dont la confiance est peu maîtrisée.

5.6 Texte de l'article

SocialCom/PASSAT/BigData/EconCom/BioMedCom 2013

Information Security Risk Management in a World of Services

Vincent Lalanne and Manuel Munier
LIUPPA
Université Pau & Pays Adour
Pau, France
Email: {vincent.lalanne,manuel.munier}@univ-pau.fr

Alban Gabillon
GePaSud EA 4238
Université Polynésie Française
France
Email: alban.gabillon@upf.pf

***Abstract*—Service Oriented Architectures (SOA) offer new opportunities for the interconnection of systems. However, for a company, opening its Information System to the "world" is not insignificant in terms of security. Whether to use available services or provide its own services, new technologies have introduced new vulnerabilities and therefore new risks. Our work aims to propose an approach for risk management which is based on the ISO/IEC 27005:2011 standard: we propose a development of this standard (by an extension of Annex D) so that it can fully take into account the type "service" as web services and cloud services. Indeed, a world of services is not limited to link interconnected systems, it is more a relationship between customer and supplier, where notions of trust, accountability, traceability and governance are developed. Following this study we introduce a new security criterion, controllability, to ensure that a company keeps control of its information even if it uses such outsourced services.**

***Keywords*—information security; risk management; SOA; cloud; web services; ISO/IEC 27005; controllability;**

I. Introduction

First information systems (IS) used in companies operate autarky, that is to say closed to the outside world and only supplied by the internal data from the enterprise. The need to connect to other systems quickly emerged, thus increasing the amount of information available, outsourcing some processes and offering new services to both employees (working at home, nomadic users,...) and customers (web portals, information flow,...). The increasing use of mobile devices, smartphones or tablets in the professional world (BYOD[1]) also introduces new risks that companies must face: storing information on the terminal, business applications installed (data, but also configuration parameters like usernames, passwords or server addresses),...

These early infrastructures used private connections, but with the Internet growth, IT designers decided to use this network to connect their information systems. This change has reduced connection costs and provided greater flexibility in the deployment of such infrastructures. From the point of view of the network, security managers have to implement various technologies to control the opening of the corporate network over the Internet: routers, firewalls, VPNs,...

With the emergence of new needs, the interconnection of information systems is the next step. This evolution is a fact in particular through the development of service-oriented architectures (SOA) that have gained great popularity because they allow the creation of new services by composition (orchestration, choreography,...) of existing services over the Internet. They may have very different features: computation, data storage, remote database access (data warehouses, schedules,...). Web Services (WS) is one of the most widely used technologies for such SOA.

Infrastructure design using external services (which we do not control) and/or exposing new services outside the company raises new problems for the information system security (ISS). It concerns not only the classical criteria as confidentiality, integrity and availability, but also new concepts like traceability, trustworthiness and controllability.

The paper is organized as follows. After detailing in Section II various technologies used to secure such web services and a quick reminder of the services offered by the cloud, Section III discusses methods and standards widely used in risk management information systems. Section IV presents our work based on a risk management approach for information system vulnerabilities related to those services. For the purpose of this work, we rely on the ISO/IEC 27005:2011 standard [1] that we extend to "services". We conclude the paper with some related works in Section V and pointers to future research directions in Section VI.

II. Web Services and Security

A. A World of Services

Interconnections between information systems via web services can be carried out either on private infrastructures or through the Internet. For obvious reasons of cost, the use of Internet and its standards is becoming more common. There is another concept very close to web services: the cloud. Indeed, this "technology" involves many services in the field of computing, storage, information processing,... The cloud model takes concepts already known in the world of services, but with aggressive marketing discourse.

The number of available services on the cloud is constantly increasing. Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Software as a Service (SaaS) first appeared, often as variations of existing web services. Now come new concepts such as Monitoring as a Service (MaaS), Communication as a Service (CaaS), Data as a Service (DaaS), inFormation as a Service (FaaS),... The XaaS (Anything as a Service) are born[1]

[1]BYOD: Bring Your Own Device

978-0-7695-5137-1/13 $26.00 © 2013 IEEE
DOI 10.1109/SocialCom.2013.88

and include all services directly accessible from the Internet and that grow on the business model of cloud computing.

Note that we can distinguish two notions: the private cloud and public cloud. Even in a private cloud (internal to the company) it is possible to control the overall infrastructure, it is not the same in public cloud.

B. Web Services Security

Many exchange protocols exist, including REST [2] and SOAP [3] in particular, but we do not detail the XML-RPC[2] [4] protocol, an ancestor of SOAP, unmaintained since 1999. We can note that REST is the native HTTP protocol while XML-RPC and SOAP uses XML over HTTP as a transport protocol over the Internet.

1) REST (REpresentational State Tranfert): REST [2] is a design pattern for implementing connected systems. It is neither a technology nor a standard. It is a type of architecture to publish resources on the web. RESTful architecture meets several principles: Applications are client-server, requests are stateless, clients and servers use a uniform interface; all resources are accessed through well defined methods like HTTP GET, POST, PUT, DELETE, HEAD and OPTIONS. Clients access to named resources; the system understands named resources using URLs such as HTTP URLs (but not only limited to HTTP URLs);

2) SOAP (Simple Object Access Protocol): SOAP [3] establishes a general framework for exchanging complex data in XML. SOAP does not depend on the programming languages (C, Java, PHP, NET, PERL,...) or the operating system on which it is implemented. A SOAP message is a unidirectional transmission between SOAP nodes, from a SOAP sender to a SOAP receiver. SOAP messages are supposed to be combined by applications to implement more complex sequences of interactions: from the basic question and answer model to multiple bidirectional exchanges for "conversational" scenarios.

A SOAP message is an XML document constituted by an envelope containing a Header (optional) and a Body (the message). The envelope is the root of the XML document containing the SOAP message. Header tag allows to pass additional information about this message. This element is optional, but if present it must be the first element in the SOAP envelope of the message. The header can have multiple uses. It may, for example, contain authentication information from the issuer, or the context of a transaction in which the message is only a step. For transport layers (such as FTP) that do not provide return address, one can use the header to identify the sender of the SOAP message. The message body consists of a single Body element containing one or more sub-elements. The message body can carry remote procedure calls, results or error messages. But the practice has become enlarged and the message body is often used to exchange structured data between applications.

Over the SOAP protocol, we can list a number of existing security measures. Additional specifications have been defined over the XML/SOAP stack[5], [6] to strengthen security of infrastructures using web services (Figure 1).

[2]XML-RPC: XML Remote Procedure Call

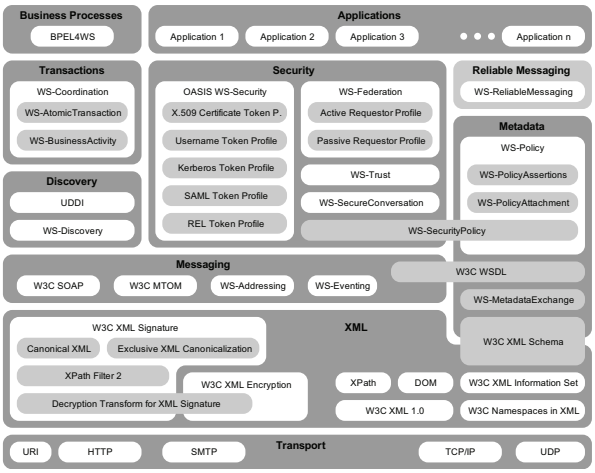


Fig. 1. The Web Services specifications stack (WS-*) [7]

Core technologies around XML are all defined by the World Wide Web Consortium (W3C). This is the case for example for XML Encryption [8] and XML Signature [9] which respectively address the issue of confidentiality (encryption) and data integrity (message authentication and/or signer). Based on these standards, standardization organizations such as OASIS have developed specifications such as Web Service Security [10]. This is a set of SOAP extensions that ensures the message integrity and confidentiality. This specification is flexible and can be adapted to security models as varied as PKI, Kerberos, and SSL. A number of specifications are associated with WS-Security:

- WS-Trust: specification for the generation, renewal and validation of security tokens
- WS-SecureConversation: creation and sharing security contexts (using security context token)
- WS-Federation: mechanisms for allowing disparate security realms to broker information on identities, identity attributes and authentication
- WS-Authorization: expression of authorizations
- WS-Policy: flexible and extensible grammar for expressing the capabilities, requirements and general characteristics of entities (customers or suppliers)
- WS-Privacy: model to indicate how confidentiality requirements and practices related to private data is transmitted between organizations

Other specifications concern more specifically the phases of authentication and authorization. Authentication is the process to validate identity, while authorization is the process of deciding if authenticated user can access such resources or perform such actions. SAML[3] specification [11] defines a framework for exchanging authentication information and authorization between business partners. SAML supports single sign-on (SSO) for affiliated sites. Another specification, XACML[4] [12], [13], defines a language for access control, rule propagation and administration of security policy for information systems.

[3]SAML: Security Assertion Markup Language
[4]XACML: XML Access Control Markup Language

III. Methods and Standards for Information System Risk Management

In this section we will briefly list the main methods used in risk management in information security, then we will, after a short history that led to its creation, detailing the main steps of the ISO/IEC 27005:2011 [1] international standard. We conclude this section with a statement that shows the limits of this standard in consideration of services as we now approach.

A. Why a Standard for Information Security ?

Existing methods for ensuring information security can not be a trust mark for the overall security of the company, because it is often developed internally and difficult to change (long term support ?). To meet the need for overall confidence in the digital economy, work has been initiated to establish international standards for information security.

Since a decade, companies having many data exchanges with other companies (national or international) or with many partners and customers, have experienced the need to agree on standards to secure information and exchange processes. It is precisely this goal that led to the creation of the ISO/IEC 27005 standard. It aims to establish a trust mark for the overall information security within enterprises.

B. Standard or Method ?

A standard is defined as a document based on a consensus covering a broad industrial or economic interest and established by a voluntary process. In contrast, a method is an effective way to achieve a desired and accurate result. But a method does not include the notion of document, neither the concept of consensus. We should not oppose standards and methods, but rather combine them: a method will be the "tool" used to meet a standard.

To effectively implement the ISO/IEC 27005 standard, we can thus rely on a risk management method as CRAMM[5] (United Kingdom), Octave [14] (United States), EBIOS[6] (France).

C. Introduction to the ISO/IEC 27005:2011 Standard

This is the latest standard for information security and it is expected to be widely used in the field of ISMS[7]. Since 1995, several standards for ISMS have been published:

BSI[8] publishes (1995) BS 7799 standard [15], it focuses on ten major chapters that list the actions (one hundred) that can be taken in relation to information security. In 2000, ISO officially adopts it under the reference ISO 17799 (now known as ISO 27002). The creation of ISO 27001 this is BS 7799 plus the requirements that an organization must meet to implement an ISMS, all in one approach closer to ISO 9001. At east ISO 27005 is published in 2008; its purpose is to provide guidelines for information security risk management; it supports the general concepts specified in ISO 27001; it is a consensus between the various methods used in different countries as CRAMM, OCTAVE, EBIOS it will progressively replace. A new version of ISO 27005 standard appeared in 2011.

ISO/IEC 27005:2011 gives recommendations and therefore it quite often uses the conditional. It is not required to follow all the steps of the method: the implementer applies what is most appropriate for his case study. This is a standard that, when applied, allows us to follow a process compliant with ISO/IEC 27001:2005.

The first six chapters of the standard are very short and deal only with generalities. The text begins by clarifying the scope of the standard (Clause 1), recalls some normative references (Clause 2), gives some terms and definitions (Clause 3), shows the general structure of the standard (Clause 4) and points out the advantages of following a risk management approach (Clause 5). Clause 6 gives an overview of the information security risk management process (steps appear in Figure 2). The information security risk management process consists of context establishment (Clause 7), risk assessment (Clause 8), risk treatment (Clause 9), risk acceptance (Clause 10), risk communication and consultation (Clause 11), and risk monitoring and review (Clause 12). Here are the steps that we need to perform to conduct the process.

In the remainder of this section we will briefly present the standard necessary to explain the following main points of our presentation, these items are highlighted in bold.

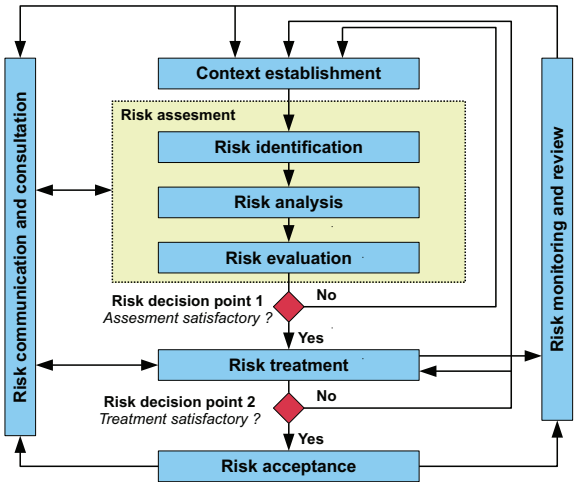


Fig. 2. Information security risk management process

We must first define the scope and **boundaries** of the information security risk management (Clause 7): this step is context establishment. This notion of context is essential, because later in our presentation, it will allow us to set the point of view of risk assessment through the consideration of risks associated with the use of Web Services. The context also sets a number of criteria which are then used as the basis for risk assessment: risk evaluation criteria, impact criteria and risk acceptance criteria.

Clause 8 contains the main steps of the process. It points

[5]CRAMM: **C**CTA (Central Computer and Telecommunications Agency) **R**isk **A**nalysis and **M**anagement **M**ethod

[6]EBIOS: **E**xpression des **B**esoins et **I**dentification des **O**bjectifs de **S**écurité (expression of needs and identification of security objectives)

[7]ISMS: Information Security Management System

[8]BSI: British Standards Institution

in particular that the risk assessment is composed of three sub-tasks: ”risk identification”, ”risk analysis” and ”risk evaluation”. Firstly we need to identify the **assets** that fall within the scope defined above and give them a **value**. The level of granularity of the assets can be refined. This point is very important and will be developed further in the following section about Web Services used by our information system. We then need to identify **threats** that may be of accidental origin or come from inside as well as outside. In the standard, Annex C provides a referential (not exhaustive) of the various possible threats. When an IT architect builds an information system, he implements security mecanisms which he considers essential. ISO/IEC 27005 takes into account this legacy and specifies that an inventory of security measures already deployed (or whose deployment is scheduled) must be done.

The next step is to identify **vulnerabilities** related to the assets that were set previously. For this, Annex D provides a referential (non-exhaustive) of various vulnerabilities that exist for each type of asset with possible threats. We must then identify and assess the **consequences** of incident scenarios that describe a threat from the exploitation of a vulnerability inherent in the asset. These **scenarios** are assessed through the evaluation of their consequences and their estimated **likelihood**. Finally, we must assess and evaluate the **risk** by linking the likelihood and consequences of the scenarios listed above. A list of valued and ranked risks can thus be established.

Clause 9 is devoted to the risk treatment. There are four options to do this: modification, retention, avoidance or risk sharing. Proposals for risk treatment lead to the identification and estimation of **residual risks**, that is to say, the risk remaining after treatment. The next step is to accept or not the risk treatment plan (Clause 10: risk acceptance).

As for any quality process related to risk management (regardless of the domain), such an audit will hardly be effective without the cooperation of the persons concerned and without regular monitoring of the system. Risk communication process (Clause 11) aims to make all stakeholders aware of the concept of risk. Clause 12 of the standard concludes by reminding the concepts of monitoring, review and improvement inherent in an iterative approach. When applying this standard, it is not enough to proceed the steps above once, but instead it is quite clear that we must constantly analyze, monitor and improve the risk assessment. A continuous watch is indeed necessary to take into account new assets, new features, the discovery of new vulnerabilities, the effectiveness of measures taken, the evolution of the security policy of the company,...

D. Observations

The difficulties encountered in risk management in distributed information systems are intrinsically linked to the SOA model. Such a ”logical” architecture relates at the same time hardware aspects (servers of the service provider), software (operating systems, implementation of the SOAP protocol, services themselves) and network (Internet, LAN, routers). Although ISO/IEC 27005 can already take these aspects into account in the study boundaries, it treats them individually: datacenters, network connections, computers, applications,... Moreover, existing technical solutions to improve IS security also address these points individually: backup servers, clusters, redundant links, encryption tools, system administration, application monitoring,...

But now a world of services is not limited to link interconnected systems, it is more a relationship between customer and supplier, where notions of trust, accountability, traceability and governance are developped.

ISO/IEC 27010 provides guidance on information security interworking and communications between industries in the same sectors, in different industry sectors and with governments, either in times of crisis and to protect critical infrastructure or for mutual recognition under normal business circumstances to meet legal, regulatory and contractual obligations. Similarly ISO/IEC 27017 will cover information security aspects of cloud computing, this standard is expected to be a guideline or code of practice recommending relevant information security controls for cloud computing. ISO/IEC 27018 will cover privacy aspects of cloud computing. These standard (not yet published) will intend to normalize communications with a package of good practices but do not reflect the specific risks of using external services.

IV. ISO/IEC 27005 Applied to Risk Management in SOA Based Information Systems

We note that this standard applies to information security has neglected the development of special services in information systems. Indeed the material, human, organizational aspects are planned, but the ability to store, process and use resources via web or generally what is called cloud services. Also, in this section we will proceed in the first part the standard ISO/IEC 27005:2011 considering asset ”service” from a technical point of view but also in terms of impact on the processed information.In the second part we will explain why this standard should evolve and how such a service-oriented approach can improve a risk management process.

A. Conduct of ISO/IEC 27005 Standard

1) Context establishment: Unlike existing work in the literature, we do not deal with the study of securing the web services themselves. We rather propose to study the impact of using a SOA on the information system security from the point of view of the risks related to information security. Indeed, if technologies such as web services bring new features and even generate new needs, they introduce however new vulnerabilities within the IS and, therefore, new risks for IS Security.

Our work concerns the interconnection of information systems (broadly defined) through the use of web services. We do not focus on the ”internal” security of these web services (eg injection of erroneous parameters), but rather on the impact of a ”failure” of a web service on the IS.

2) Identification of assets: In this context, we are led to consider two types of assets:

- web services themselves: input and output data flows, business processes they implement,...
- underlying communications infrastructure: systems (computers and OS), network, software platforms (eg SOAP implementation, servlet container),...

3) Identification of threats: According to the glossary of keywords for information security established by NIST [16], a threat is defined as *"any circumstance or event with the potential to adversely impact organizational operations (including mission, functions, image, or reputation), organizational assets, individuals, other organizations, or the Nation through an information system via unauthorized access, destruction, disclosure, modification of information, and/or denial of service"*.

Within the meaning of ISO standards, a threat is *"a potential cause of an incident, that may result in harm of systems and organization"*. Threats may be of natural or human origin, and could be accidental or deliberate. A risk is *"the potential that a given threat will exploit vulnerabilities of an asset or group of assets and thereby cause harm to the organization; it is measured in terms of a combination of the probability of occurrence of an event and its consequence"*.

With regard to the field of our study, here is a non exhaustive list threats that could affect the operation of web services:

- an incident on the network causes malfunctions (excessive delays, loss of connection)
- a malicious person intercepts messages and forge new posts to harm the IS and/or access certain information
- a software error (accidental or deliberate) on a WS causes the sending of erroneous results

4) Identification of vulnerabilities: Vulnerability can be defined as *"a weakness in the information system, system security procedures, internal controls, or implementation that could be exploited or triggered by a threat source"*. This definition considers not only the vulnerabilities of software components, but also organizational aspects.

When talking about vulnerabilities of service-oriented architectures (SOA), some are now known (OWASP[9], MITRE and the CVE[10] project), but we must also take into account all the components that are part of this technology and are the target of attacks (WS-attacks.org[11]) like web service client, web service server, BPEL engine, intermediaries in the web service architecture (threat as "Man in the Middle"), XML parser, schema validation (XML), signature verification, encryption and decryption processes.

From these, it is possible to point out some examples of vulnerabilities and associated threats as interception and tampering of a WS request ($\rightsquigarrow$ identity spoofing), tampering WS metadata (WSDL file, WS-security-policy) ($\rightsquigarrow$ abuse of rights), WS-Addressing hacking to change the destination address of the message ($\rightsquigarrow$ data theft), BPEL engine flooding with a huge number of SOAP requests ($\rightsquigarrow$ service no longer available).

Under the ISO/IEC 27005 standard, these vulnerabilities are already referenced in Appendix D within the types "hardware", "software" or "network". Moreover, a number of existing standards and related technologies already allows to strengthen these aspects of WS security (cf. Section II-B2).

[9]OWASP: the Open Web Application Security Community

[10]CVE: Common Vulnerabilities and Exposures

[11]http://clawslab.nds.rub.de/wiki/index.php

Our research activities are at a higher level of granularity: we focus on the security of the information itself (content structure, chain of production and consumption, economic and legal issues related to the information) and not as a "simple" input/output parameter flowing within the IS. If you look at the cycle of information processing (storage, processing, transport) as such, it is certain that many vulnerabilities and threats are emerging with the use of "cloud" services. Thus, we can classify these vulnerabilities according to several categories [17].

a) ***Quality of Service****:* despite the Service Level Agreements (SLA) it is not uncommon to be faced with frequent breakdowns and unavailable services (eg, Amazon EC2, April 2012) which can reach thousands of users around the world. The only possible compensation is financial whereas your image reputation and goodwill may have been heavily impacted. Similarly what about the customer ? Can he migrate to another service provider ? Are there interoperability, portability, transferability when the client wants to take back his information ?

b) ***The location of data and processing****:* these outsourced services can be located anywhere in the world, sometimes without any possibility to choose the country. However, as laws differ from one country to another, some information may be acceptable in one country but prohibited in another.

c) ***Loss of control of information****:* when using outsourced services, companies entrust their data to the provider. That raises various issues about control over critical data of companies (eg research & development, business strategy):

- ***incident*** at the service provider: is it always possible to recover data ? Who is responsible for making backups: clients or the provider ?
- ***reversibility***: it is supposed to allow the client to repossess his data at any time without justification. But the lack of standards for interoperability and migration induces a strong dependency towards the provider.
- ***termination of the contract***: physical erasure of data is rarely complete. Frequently, the provider can not offer any warranty: redundant architectures, backups, archives, indexes, analytics (how many copies ? location ? are there still data somewhere?)

d) ***Information ownership****:* when processing information in the cloud, we entrust the capital of the company to a third party. What happens in case of litigation (eg non-payment, injunction), if the provider goes out of business (eg bankruptcy, acquisition by another company),... ? Can the provider keep your data ? Has he contractually the right to continue to exploit them ?

e) ***The type of information****:* in addition to the traditional data that can be processed in a company, there are those who are particularly sensible: critical data (research and development) and personal data. Regarding personal data, they can be opened by laws in a country whereas they are closed in others (eg USA PATRIOT Act, FISAAA[12]), which may have an impact on privacy.

[12]FISAAA: US Foreign Intelligence Surveillance Act of 1978 Amendments Act of 2008

5) Identification of consequences: From this list of vulnerabilities (not exhaustive) we can now present some basic scenarios illustrating the occurrence of a threat that exploits a vulnerability in a service with its impact on the information system. We distinguish "classical" vulnerabilities (●) that can be addressed using existing technologies and security controls (cf. Section II-B2) and "organizational" vulnerabilities (○) affecting the security of the information itself.

- ● **Identity spoofing:** in a web service context, authentication attacks bypass identification controls by using a number of different techniques, such as session hijacking, session replay, session fixation, identity spoofing, valid credentials theft, authentication module subversion, and brute-forcing. As a result of a successful attack, an illegitimate user will be identified as a legitimate one to access new resources [18].
- ● **Tampering WS metadata**: Maliciously changing the content of the WSDL file [19]. This usually aims at lowering the security requirements of an web service. The information that certain message data is required to be encrypted just gets removed, resulting in an unencrypted communication between web services, enabling the attacker to read the message content.
- ○ **Trust in service**: a failure of one service may affect the operation of another service cascade hosted in this container. This repetition of failure can lead to a domino effect where each service sends its error successively [20].
- ○ **Quality of Service**: The service provider does not fulfill the service by itself; it is a broker with service subcontractors. In case of litigation or digital forensics the issues regarding the quality of the services and/or responsabilities are very difficult to determine.
- ○ **Data stored in a foreign country**: However, some businesses do not like the ability of a country to get access to their data via the court system, for example, a European customer might be concerned about using Cloud Computing system in the United States given the USA PATRIOT Act [21], FISAAA, RIOT[13] and PRISM (since 2007 and revealed in june 2013 by Edward Snowden). In this case the loss of privacy is obvious, whether data is personal or not [22].
- ○ **Prohibited or copyrighted informations**: when sharing a service that is not looking at the origin of the data it stores, the service can be closed overnight eg. Megaupload 2012, resulting in a direct loss of data for the friendly customers of the law.
- ○ **The provider has financial difficulties**: the supplier fails and stops the services it provides. It can also claim money to restore your data (eg 2e2.com 2013).
- ○ **Data recovery**: closing procedures of a service, such as at the end of a contract, should provide the ability to retrieve data (reversibility) but also ensure their final removal on provider's servers (physical erasure of data).

After identifying the incident scenarios, it is necessary to assess the likelihood of each scenario and impact occurring, using qualitative or quantitative analysis techniques. This should take account of how often the threats occur and how easily the vulnerabilities may be exploited.

[13]RIOT: Rapid Information Overlay Technology by Raytheon

B. Evolution of ISO/IEC 27005 Standard Towards Services

As explained in this paper, our work concerns information security as a whole: from contents to economic and legal issues. This approach does not question the concepts established in the ISO/IEC 27005:2011 standard. Instead, it is complementary because our proposal is to add in Annex D of the ISO/IEC 27005:2011 standard a new type named "service" with vulnerabilities and threats related to WS technology and services provided: it is a kind of super-type over existing types "hardware", "software" and "network".

After setting the type "service" in this nomenclature, vulnerabilities, and therefore the associated threats, are viewed in a new light. The main advantage is that the scenarios are closer to reality: we talk about "denial of service" by example rather than "loss of network connection"; "service provider returning erroneous data" instead of "software failure". These vulnerabilities bring simple questions that are at the heart of every service user and allow it to guarantee some "controllability" about his data:

- Can I act freely on my data ?
- How do I know where is my data ?
- How do I know who can access my data ?
- Do I own my data ?
- Can I easily change provider ?
- In what country is my data stored ?
- My data are they legal in the country where they are stored ?
- Does my provider impose respect for the intellectual property to all its customers (to avoid being forced to close) ?

In the Table I, we used the same formalism as in Annex D of the ISO/IEC 27005:2011 standard to present a (non-exhaustive) list of vulnerabilities and threats associated with specific "service" type and taking into account the proposals developed in the previous sections. The new ISO/IEC 27010, 27017 and 27018 standards will not replace our approach. Instead, they merely reaffirm, by their existence, the necessity to take into account the notion of "service" in a process of risk management.

Thus, the concept of "controllability" appears fundamental in the design of information systems using third party services. The IT architect must be absolutely sure that he can organize, manage and thus "control" his services, to avoid that his information system becomes unusable.

From this premise, it is quite possible to define a fourth security criterion specific to services, in addition to the three basic criteria CIA:

1) **Confidentiality**: it prevents the unauthorized disclosure of information (eg illegal read access)
2) **Integrity**: it guarantees the accuracy, completeness, validity and stability of information
3) **Availability**: it ensures continuity of service and system performance
4) **Controllability**: it ensures complete control over services used

Type	Examples of vulnerabilities	Examples of threats
Hardware	...	...
Software	...	...
Network	...	...
Service	Lack of long term support from service provider	Service no longer available
	Unknown life cycle and update policies from service provider	Unexpected change of the service interface
	Unknown country to host services	Spying, data theft
	Failure to comply with the laws in force	Service no longer available
	Provider goes bankrupt	Service no longer available
	Laws on privacy differ in the country of use	Loss of confidentiality
	Laws on information security are less restrictive	Spying, data theft
	Laws on information security are more restrictive	Service no longer available
	Inadequate Service Level Agreement	Breach of maintainability of the IS
	Lack of data recovery procedure	Breach of maintainability of the IS
	Lack of reversibility (migration, interoperability)	Breach of maintainability of the IS
	Lack of data erasure at the end of the contract	Data theft, unauthorized use
	Lack of WS metadata security	Tampering with the web service
	Possible multiple requests to the WS	Identity spoofing
	Lack of traceability of the service provided	Breach of trustworthiness of information
Personnel	...	...
Site	...	...
Organization	...	...

TABLE I. A NEW TYPE "SERVICE" IN ANNEX D OF THE ISO/IEC 27005:2011 STANDARD

V. RELATED WORK

As we previously described, information exchange between systems and particularly in the cloud introduce new risks with regard to information system security. As pointed out in [23] there are two methods to remedy: trust the service provider or implement technical mechanisms to compensate for a trust worthy supplier.

First to compensate the lack of trust in a service it's necessary to put in place mechanisms that enable secure information exchange between systems. In [24] authors discuss security issues for cloud computing and present a layered framework for secured cloud storage. They focus on essential aspects: how to store data in foreign machines, queyring encrypted data and secure this queries.

On the other hand it is also necessary to estimate the security coverage for different type of services. In [25] authors describe a framework which can prescribe the right combination of security tools for different cloud services and according to the level of security assurance required.

Accountability as seen from a holistic point of view, covering legal, socio-economic, regulatory and technical aspects is presented in [26]. That European project named A4Cloud aims at four objects which tackles accountability developing tools that:

1) enable cloud service providers to give their users appropriate control and transparency over how their data is used,
2) enable cloud end users to make choices about how cloud service provider may be use,
3) monitor and check compliance with user's expectation, business policies and regulations,
4) develop recommendations and guidelines for how to achieve accountability for the use of data by cloud services.

The project seems to be a promising approach that tackles preventive, detective and corrective control which allow to achieve accountability.

In [27] authors introduce the notion of autonomous self-controlling objects (SCO), that adapt to the location and other contextual dimensions. The sensitive resources assure their protection by means of adaptive security policies of various granularity, and synchronization protocols.

In the European Network and Information Security Agency (ENISA) a Preparatory Action entitled "Trust and privacy in the Future Internet" covers in one work package identity, accountability and trust in the Future Internet [28]. The objective is to study security models of electronic services and their performance in highly distributed environments, such as today's Internet. Furthermore, ENISA investigates various ways of assuring privacy and accountability on the Internet, reviewing the most prominent methods used, studying their mapping to the underlying architectures and assessing their level of effectiveness and performance. ENISA also works towards the development of recommendations on the use of specific service models in given environments and architectures.

VI. CONCLUSION AND FUTURE WORK

Service Oriented Architecture (SOA) models incuding cloud services are increasingly used because of the benefits they provide. However, from the point of view of information security risk management, these technologies introduce new vulnerabilities with regard to network connections, the use of external services that we can not control, responsibilities related to the provision of service,... Of course, there are

solutions to secure these exchanges and design ”reliable” web services. Our work is complementary to these solutions because we operate at the level of information security within the meaning of the ISO/IEC 27005:2011 standard. Our goal is to refine the process of risk management to the specificities of services in an SOA by taking into account the information in its entirety, including socio-economic and legal aspects. We therefore suggest to extend the ISO/IEC 27005:2011 standard to include consideration of services (web services, cloud) in the risk assessment of information security.

Having identified a number of scenarios that can not be avoided through the existing security mechanisms, the second phase of our work is to propose a security model for communications between IS. To achieve this goal, we consider a usage control oriented approach, as we have already experienced in our previous work on intelligent documents [29], [30], [31] (cf. *Enterprise Digital Right Management*). The use of metadata for traceability of communications (via these services) will also allow us to compute indicators that can, for example, be used to monitor the IS. Thus, these mechanisms allow us to strengthen the security of information we may entrust to a service provider.

Enhancing the importance of providing more transparency and control to processes mediated by the cloud and taking into account that data is dynamic due to the complex responsibility chains, we also propose as a perspective an approach based on preventive, detective and corrective accountability methods [32].

The implementation of this model is obviously one of our prospects. This is what led us to consider Web Services technology since many dedicated security specifications have been developed about the SOAP protocol. Moreover, these specifications (presented in Section II-B2) are ”open”, that is to say that we are free to implement our own models and security policies.

REFERENCES

[1] ISO/IEC, “ISO/IEC 27005:2011: Information security risk management,” International Organization for Standardization (ISO), Geneva, Switzerland, Published, 2011.

[2] R. T. Fielding and R. N. Taylor, “Principled design of the modern web architecture,” in *Proceedings of the 22nd international conference on Software engineering*. ACM, 2000, pp. 407–416.

[3] W3C, “SOAP version 1.2,” W3C Open Source Software, Tech. Rep., june 2003.

[4] UserLand, “XML-RPC,” UserLand Software, Inc., Tech. Rep., 1999.

[5] E. Bertino, L. Martino, F. Paci, and A. Squicciarini, *Security for Web Services and Service-Oriented Architectures*. Springer Publishing Company, Incorporated, 2009.

[6] K. Beznosov, D. J. Flinn, S. Kawamoto, and B. Hartman, “Introduction to web services and their security,” *Information Security Technical Report*, vol. 10, no. 1, pp. 2 – 14, 2005.

[7] C. Geuer-Pollmann and J. Claessens, “Web services and web service security standards,” *Information Security Technical Report*, vol. 10, no. 1, pp. 15 – 24, 2005.

[8] D. Eastlake and J. Reagle, “XML Encryption Syntax and Processing,” W3C, Tech. Rep., 2002.

[9] M. Bartel, J. Boyer, B. Fox, B. LaMacchia, and E. Simon, “XML Signature Syntax and Processing,” W3C Open Source Software, Tech. Rep., 2008.

[10] A. Nadalin, C. Kaler, P. Hallam-Baker, R. Monzillo, and E. Al., “Web Services Security: SOAP Message Security 1.0 (WS-Security 2004),” *OASIS Standard*, vol. 200401, no. February, 2006.

[11] Cantor, Kemp, Philpott, and Maler, “Security Assertion Markup Language (SAML),” OASIS, Tech. Rep., 2005.

[12] T. Moses and S. Godik, “eXtensible Access Control Markup Language (XACML) Version 1.0,” OASIS, Tech. Rep., 2003.

[13] M. Lorch, S. Proctor, R. Lepro, D. Kafura, and S. Shah, “First experiences using xacml for access control in distributed systems,” in *Proceedings of the 2003 ACM workshop on XML security*. ACM, 2003, pp. 25–37.

[14] C. Alberts and A. Dorofee, “An introduction to the OCTAVE method,” *Pittsburgh, PA: Software Engineering Institute, Carnegie Mellon University*, 2001.

[15] “BS 7799:Part 1:1995 information security management code of practice for information security management systems,” BSI British Standards, Tech. Rep., 1995.

[16] “NIST IR 7298 Rev. 1: Glossary of key information security terms,” National Institute of Standards and Technology, Computer Security Ressource Center, Tech. Rep., Fév-2011.

[17] B. Grobauer, T. Walloschek, and E. Stocker, “Understanding cloud computing vulnerabilities,” *Security Privacy, IEEE*, vol. 9, no. 2, pp. 50–57, March 2011.

[18] G. Álvarez and S. Petrović, “A new taxonomy of web attacks suitable for efficient encoding,” *Computers & Security*, vol. 22, no. 5, pp. 435–449, 2003.

[19] M. Jensen, N. Gruschka, and R. Herkenhöner, “A survey of attacks on web services,” *Computer Science-Research and Development*, vol. 24, no. 4, pp. 185–197, 2009.

[20] V. Dialani, S. Miles, L. Moreau, D. De Roure, and M. Luck, “Transparent fault tolerance for web services based architectures,” in *Euro-Par 2002 Parallel Processing*. Springer, 2002, pp. 889–898.

[21] L. T. Lee, “USA PATRIOT Act and telecommunications: Privacy under attack,” *Rutgers Computer & Tech. LJ*, vol. 29, p. 371, 2003.

[22] M. Zhou, R. Zhang, W. Xie, W. Qian, and A. Zhou, “Security and privacy in cloud computing: a survey,” in *Semantics Knowledge and Grid (SKG), 2010 Sixth International Conference on*. IEEE, 2010, pp. 105–112.

[23] W. Pieters, “Security and privacy in the clouds: a bird’s eye view,” in *Computers, privacy and data protection: An element of choice*. Springer, 2011, pp. 445–457.

[24] K. Hamlen, M. Kantarcioglu, L. Khan, and B. Thuraisingham, “Security issues for cloud computing,” *International Journal of Information Security and Privacy (IJISP)*, vol. 4, no. 2, pp. 36–48, 2010.

[25] D. Dasgupta and M. M. Rahman, “Estimating security coverage for cloud services,” in *Privacy, security, risk and trust (PASSAT), 2011 IEEE third international conference on and 2011 IEEE third international conference on social computing (SocialCom)*. IEEE, 2011, pp. 1064–1071.

[26] S. Pearson, V. Tountopoulos, D. Catteddu, M. Sudholt, R. Molva, C. Reich, S. Fischer-Hubner, C. Millard, V. Lotz, M. G. Jaatun *et al.*, “Accountability for cloud and other future internet services,” in *Cloud Computing Technology and Science (CloudCom), 2012 IEEE 4th International Conference on*. IEEE, 2012, pp. 629–632.

[27] A. C. Squicciarini, G. Petracca, and E. Bertino, “Adaptive data protection in distributed systems,” in *Proceedings of the third ACM conference on Data and application security and privacy*. ACM, 2013, pp. 365–376.

[28] “Privacy, accountability and trust – challenges and opportunities,” European Network and Information Security Agency (ENISA), Heraklion, Crete, Greece, A report by the ENISA Ad Hoc Working Group on Privacy and Technology., 2010.

[29] M. Munier, V. Lalanne, and M. Ricarde, “Self-protecting documents for cloud storage security,” in *TrustCom*. IEEE, 2012, pp. 1231–1238.

[30] M. Munier, “A secure autonomous document architecture for enterprise digital right management,” in *SITIS*. IEEE, 2011, pp. 16–23.

[31] M. Munier, “A multi-view approach for embedded information system security,” in *CRiSIS*. IEEE, 2010, pp. 65–72.

[32] E. Jaramillo, M. Munier, and P. Aniorté, “Information security in business intelligence based on cloud: A survey of key issues and the premises of a proposal,” in *WOSIS*, 2013.

Chapitre 6

SARSSI 2013
Gestion des risques dans les systèmes d'information orientés services

Sommaire

6.1 Présentation

Cet article a été présenté sous forme de poster à la 8ème Conférence sur la Sécurité des Architectures Réseaux et des Systèmes d'Information SARSSI 2013 [27] qui s'est tenue du 16 au 18 septembre 2013 à Mont de Marsan à IUT des Pays de l'Adour - Université de Pau et des Pays de l'Adour - France.

6.2 Les thèmes abordés

D'une manière beaucoup plus concise nous abordons ici les thèmes qui ont été développés de manière plus détaillée à la conférence PASSAT 2013. La date de soumission de cet article est antérieure à la conférence de Washington (mais la date de présentation est postérieure) : seules les premières briques sont posées.

Nous faisons d'abord le point sur les services Web et plus particulièrement nous mettons l'accent sur l'aspect sécurité des échanges. Ensuite nous étendons ce concept à la notion de services et plus particulièrement sur les nouveaux aspects commerciaux liés au cloud.

Dans une dernière partie nous appliquons la norme ISO/IEC 27005:2011 à la gestion des risques des systèmes d'information qui met en oeuvre des services Web.

6.3 Les propositions

Nos travaux visent à apporter une modification dans la démarche de gestion des risques en s'appuyant sur un standard : la norme ISO/IEC 27005:2011. L'utilisation de ce standard, à portée internationale, permet de mettre en œuvre un outil qui s'applique à tous les systèmes d'information.

Nous proposons une évolution de cette norme afin qu'elle puisse prendre en compte pleinement le type « service »dans la gestion des risques d'un système d'information. Lors de sa création en 2008, les services Web existaient mais bien souvent sous forme propriétaire et/ou privée. Or maintenant, avec la généralisation et la commercialisation de nombreux services de type « CLOUD ».

Aussi, afin de tenir compte de ce risque dans l'étude d'un système d'information, nous proposons d'étendre un nouveau type d'actif au sein du système d'information qui serait dénommé « services ».

Suite à cette proposition nous introduisons également un nouveau critère de maîtrisabilité à côté des traditionnels critères de confidentialité, d'intégrité et de disponibilité. Ce critère de

27. http ://sarssi2013.univ-pau.fr/

maîtrisabilité permettra de qualifier la manière dont est contrôlée l'utilisation des services Web dans le système d'information.

6.4 Conclusions

À travers ces constats et propositions, nous alertons d'abord la communauté sur les risques d'utilisation des services Web (ou cloud) dans un environnement international où les lois et règlements peuvent être différents : protégeant l'information dans un pays, permettant la consultation pour diverses raisons dans un autre.

Grâce à l'ajout d'un critère de maîtrisabilité, nous permettons ainsi à l'architecte du système d'information de qualifier la confiance qu'il a dans le ou les services qu'il utilise.

Nous proposons afin de minimiser le risque, d'utiliser notre principe de Documents intelligents associés aux métadonnées générées par les échanges entre les services.

6.5 Poster présenté à la conférence SARSSI 2013

SARSSI 2013 - Mont de Marsan

Gestion des Risques

dans les Systèmes d'Information Orientés Services

Vincent Lalanne - Manuel Munier - Alban Gabillon

Résumé : Les architectures orientées services (SOA) offrent de nouvelles possibilités pour l'interconnexion des systèmes d'information. L'ouverture du SI d'une entreprise sur l'extérieur n'est toutefois pas anodine du point de vue de la sécurité. Que ce soit pour utiliser des services proposés par des tiers ou pour offrir les siens, ces technologies introduisent de nouvelles vulnérabilités dans le SI et, par conséquent, de nouveaux risques. Nos travaux visent à initier une démarche de gestion de ces risques qui s'appuie sur un standard, la norme ISO/IEC 27005:2011. Nous proposons une évolution de cette norme afin qu'elle puisse prendre en compte pleinement le type « service ». Suite à cette étude nous introduisons également un nouveau critère, la maîtrisabilité, pour qualifier la sécurité des systèmes d'informations.

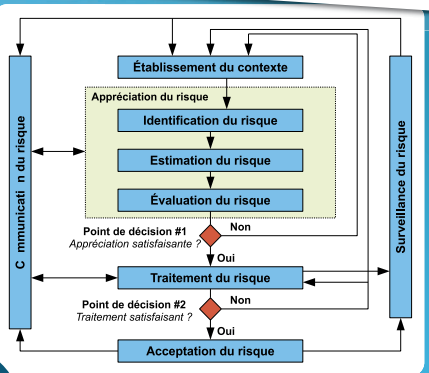
Un monde de services

L'interconnexion des systèmes d'information, avec en particulier le développement des architectures orientées services (SOA), permet la création de nouveaux services par la composition (orchestration, chorégraphie) de services existants sur Internet.

Il n'est plus possible, à l'heure actuelle, de dissocier les services web d'un autre concept qui leur est proche : le « cloud » : ce concept fait appel à de nombreux services dans le domaine du traitement, du stockage, de la transformation de l'information.

La notion de « service » soit être prise en compte dans le processus de gestion des risques au niveau de la sécurité de l'information.

ISO/IEC 27005:2011



Notion de Maîtrisabilité

La qualité de service : Il n'est pas rare d'être confronté à des pannes fréquentes, les services deviennent indisponibles, ce qui peut toucher des milliers d'utilisateurs de par le monde. La seule compensation possible est d'ordre financière alors que votre image a été lourdement dégradée.

La localisation des données et des traitements : ces services externalisés peuvent être situés dans le monde entier, sans qu'il soit possible de pouvoir choisir le pays. Dans la plupart des cas, le prestataire de cloud ignore lui-même où sont les données et où sont exécutés les traitements de l'information .

La perte de contrôle de l'information dont l'origine peut avoir des causes bien différentes. En cas d'incident chez le prestataire de services, est-il toujours possible de récupérer ses données ? Est-ce aux clients ou au fournisseur de services de faire des sauvegardes ? De plus qu'en est-il de la réversabilité qui est censée permettre au client de reprendre possession de ses données, à tout moment, sans justification ?

La propriété de l'information : quand on traite des informations dans le cloud, on confie le capital de l'entreprise à un tiers. Que se passe-t-il en cas de litige (ex : non-paiement, injonction), s'il cesse son activité (ex : faillite, rachat),... ? Le prestataire a-t-il la possibilité de garder vos données? A-t-il contractuellement le droit de les exploiter ?

Dans tous les cas l'architecte doit garder la maîtrise de ses services !

Vulnérabilités liées aux services

Type	Exemples de vulnérabilités	Exemples de menaces
Matériel	...	...
Logiciel	...	...
Réseau	...	...
Service	Pas de support à long terme du fournisseur de service	Le service n'est plus disponible
	Le cycle de vie et les politiques de mise à jour du fournisseur de service sont inconnus	Changement inattendu de l'interface du service
	Hébergement des services dans un pays inconnu	Espionnage, vol de données
	Ne se conforme pas avec les lois en vigueur	Le service n'est plus disponible
	Le fournisseur fait faillite	Le service n'est plus disponible
	Les lois sur la vie privée du pays d'hébergement sont différentes des lois du pays d'utilisation de ces données	Perte de confidentialité
	Permissivité des lois sur la sécurité de l'information	Espionnage, vol de données
	Sévérité des lois sur la sécurité de l'information sont	Le service n'est plus disponible
	Contrat de niveau de service inapproprié	Perte de maîtrisabilité du SI
	Pas de procédure de récupération des données	Perte de maîtrisabilité du SI
	Manque de réversabilité (migration, interopérabilité)	Perte de maîtrisabilité du SI
	Absence de procédure d'effacement des données à la fin du contrat	Vol de données, utilisation non autorisée
	Absence de sécurisation des métadonnées du WS	Falsification du web service
	Possibilité d'appels multiples du WS	Usurpation d'identité
	Absence de traçabilité du service fourni	Absence de confiance dans l'information
Personnel	...	...
Site	...	...
Organisation	...	...

Table : Un nouveau type "service" dans l'annexe D de la norme ISO/IEC 27005:2011

ALBAN GABILLON
alban.gabillon@upf.pf
Université de Polynésie Française
GePaSud EA 4238

VINCENT LALANNE
vincent.lalanne@univ-pau.fr
MANUEL MUNIER
manuel.munier@univ-pau.fr
Université de Pau et des Pays de l'Adour
LIUPPA EA 3000

6.6 Texte de l'article

Gestion des risques dans les systèmes d'information orientés services

Vincent Lalanne (vincent.lalanne@univ-pau.fr)*
Manuel Munier (manuel.munier@univ-pau.fr)*
Alban Gabillon (alban.gabillon@upf.pf)†

Résumé : Les architectures orientées services (SOA) offrent de nouvelles possibilités pour l'interconnexion des systèmes d'information. L'ouverture du SI d'une entreprise sur l'extérieur n'est toutefois pas anodine du point de vue de la sécurité. Que ce soit pour utiliser des services proposés par des tiers ou pour offrir les siens, ces technologies introduisent de nouvelles vulnérabilités dans le SI et, par conséquent, de nouveaux risques. Nos travaux visent à initier une démarche de gestion de ces risques qui s'appuie sur un standard, la norme ISO/IEC 27005:2011. Nous proposons une évolution de cette norme afin qu'elle puisse prendre en compte pleinement le type "service". Suite à cette étude nous introduisons également un nouveau critère, la maîtrisabilité, pour qualifier la sécurité des systèmes d'informations.

Mots Clés : sécurité des systèmes d'information, gestion des risques, maîtrisabilité, ISO/IEC 27005, SOA, cloud, services web.

1 Introduction

Les systèmes d'informations présents dans les entreprises ont d'abord fonctionné en autarcie, c'est-à-dire fermés au monde extérieur et n'étant alimentés que par les données internes à l'entreprise. Très vite la nécessité de se connecter à d'autres systèmes est apparue, permettant ainsi d'accroître la quantité d'informations disponible, d'externaliser certains traitements et d'offrir de nouveaux services, tant au personnel (travail à domicile, itinérance,...) qu'aux clients (portails web, flux d'informations,...). L'usage de plus en plus fréquent des terminaux mobiles, smartphones et autres tablettes tactiles dans le monde professionnel (BYOD [1]) introduit également de nouveaux risques auxquels les entreprises se doivent de faire face : stockage d'informations, applications métier,...

Ces premières connexions ont été réalisées grâce à des liaisons privées mais, avec le déploiement d'Internet, le concepteur s'est orienté vers l'utilisation de ce réseau pour connecter ses différents systèmes d'informations. Cette évolution a permis de diminuer les coûts de connexion et d'apporter une grande souplesse dans le déploiement de telles infrastructures. Du point de vue du réseau proprement dit, les responsables sécurité ont dû mettre

*. LIUPPA, Université de Pau et des Pays de l'Adour, France
†. GePaSud EA 4238, Université de la Polynésie Française, France
1. BYOD : *Bring Your Own Device*

en œuvre différentes technologies pour maîtriser l'ouverture du réseau de l'entreprise sur Internet : routeurs, pare-feux, réseaux privés virtuels,...

L'interconnexion des systèmes d'information est une réalité en particulier avec le développement des architectures orientées services (SOA) car elles permettent la création de nouveaux services par la composition (orchestration, chorégraphie) de services existants sur Internet. Ceux-ci peuvent avoir des fonctionnalités très diverses : le calcul, le stockage des données, la consultation d'informations distantes (catalogues, horaires). Les services web (WS) sont une des technologies actuellement les plus utilisées pour de telles architectures.

La conception d'infrastructures faisant appel à des services extérieurs n'est pas sans soulever des problèmes quant à la sécurité des systèmes d'information (SSI). Cela concerne non seulement les critères classiques de confidentialité, d'intégrité et de disponibilité, mais également des notions telles que la traçabilité ou la confiance.

Après avoir détaillé différentes technologies mises en œuvre dans la sécurisation de ces services web, nous allons proposer une approche innovante mettant en œuvre une gestion des risques liée à l'utilisation de ces services. Cette approche s'appuie sur la norme ISO/IEC 27005:2011 [ISO11] que nous envisageons d'étendre aux services.

2 Services Web et Sécurité

L'interconnexion de différents systèmes d'information via des services web peut être réalisée soit sur des infrastructures propriétaires soit au travers d'Internet. Pour des raisons évidentes de coût, l'utilisation d'Internet et de ses standards est de plus en plus fréquente. Il existe plusieurs protocoles d'échange, dont en particulier XML-RPC [2] [Use99], REST [3] et SOAP [4] [W3C03]. REST est le mode natif du protocole HTTP. La sécurité des informations transmises est assurée par le protocole de transport HTTPS. SOAP utilise quant à lui le format XML avec HTTP comme protocole de transport sur Internet.

2.1 SOAP : Simple Object Access Protocol

Le protocole SOAP [W3C03] définit le cadre général pour l'échange de données complexes structurées en XML. SOAP est indépendant des langages de programmation (C, JAVA, PHP, NET, PERL,...) ou des systèmes d'exploitation sur lesquels il est implémenté. Un message SOAP est une transmission unidirectionnelle entre des nœuds SOAP, d'un émetteur SOAP vers un récepteur SOAP, mais les messages SOAP sont supposés être combinés par les applications pour implémenter les séquences plus complexes d'interactions, de la requête-réponse aux échanges multiples "conversationnels" dans un sens et dans l'autre.

Au-delà du protocole SOAP nous pouvons identifier un certain nombre de mesures de sécurité existantes. Des spécifications supplémentaires ont été définies au dessus de la pile XML/SOAP [BMPS10, BFKH05, OAS06] pour renforcer la sécurité des infrastructures utilisant des services web (cf. pile de spécifications des web services (WS-*) [GPC05]).

Les technologies de base autour d'XML sont toutes définies par le World Wide Web Consortium (W3C). C'est le cas par exemple pour XML Encryption [ER02] et XML

2. XML-RPC : *XML Remote Procedure Call* ; n'est plus maintenu depuis 1999
3. REST : *REpresentational State Transfert*
4. SOAP : *Simple Object Access Protocol*

Signature [BBF+08], permettant de régler respectivement la question de la confidentialité (chiffrement des données) et de l'intégrité des données (authentification du message et/ou du signataire). À partir de ces standards, des organismes de normalisation tels que l'OASIS ont développé des spécifications telles que Web Service Security [NKHB+06]. C'est un lot d'extensions SOAP qui garantit au message son intégrité et sa confidentialité. Cette spécification est flexible et peut être accommodée de modèles de sécurité aussi variés que PKI, Kerberos et SSL. Un certain nombre de spécifications sont associées à WS-Security dont notamment WS-Trust [5], WS-SecureConversation [6], WS-Federation [7], WS-Authorization [8] et WS-Privacy [9].

D'autres spécifications concernent plus particulièrement les phases d'authentification et d'autorisation. L'authentification est le processus pour valider les identités, tandis que l'autorisation est un processus visant à déterminer qu'une partie authentifiée peut accéder à ce type de ressources ou d'effectuer ce type d'actions. La spécification SAML [10] [CKPM05] définit un cadre de travail pour échanger des informations d'authentification et d'autorisation entre les partenaires d'affaires. SAML prend en charge l'authentification unique (SSO) pour les sites affiliés. La spécification XACML [11] [MG03] définit quant à elle un langage pour le contrôle d'accès, la circulation des règles et l'administration de la politique de sécurité des systèmes d'information.

2.2 Un Monde de Services

Il n'est plus possible, à l'heure actuelle, de dissocier les services web d'un autre concept qui leur est proche : le "cloud". En effet ce concept fait appel à de nombreux services dans le domaine du traitement, du stockage, de la transformation de l'information. Le modèle du cloud reprend des concepts déjà connus dans le monde des services mais avec un discours marketing agressif et un nouvel acronyme : les SOA.

Le nombre de services disponibles sur le cloud est en perpétuelle augmentation. Les infrastructures (Infrastructure as a Service, IaaS), les plates-formes (Platform as a Service, PaaS), les logiciels (Software as a Service, SaaS) sont d'abord apparus, n'étant souvent qu'une déclinaison des services web déjà existants, pour laisser place maintenant à la supervision (Monitoring as a Service, MaaS), les communications (Communication as a Service, CaaS), les données (Data as a Service, DaaS), les informations (inFormation as a Service, FaaS), etc... Les XaaS (Anything as a Service) sont nés et regroupent l'ensemble des services informatiques directement accessibles depuis Internet et qui se développent sur le business model du cloud computing. À noter que nous pouvons distinguer deux notions : le cloud privé et le cloud public. Si dans un cloud privé (interne à l'entreprise) il est possible de maîtriser l'infrastructure, ce n'est pas toujours le cas dans un cloud public.

5. WS-Trust : spécification pour la génération, le renouvellement et la validation de *security tokens*
6. WS-SecureConversation : création et partage de contextes de sécurité (via un *security context token*)
7. WS-Federation : définition de mécanismes de fédération d'espaces de confiance hétérogènes
8. WS-Authorization : expression des autorisations
9. WS-Privacy : modèle pour indiquer comment les besoins de confidentialité et les pratiques liées aux données privées sont transmises entre organisations
10. SAML : Security Assertion Markup Language
11. XACML : XML Access Control Markup Language

Vincent Lalanne et Manuel Munier et Alban Gabillon

3 L'ISO/IEC 27005 Appliquée à la Gestion des Risques dans les Systèmes d'Information Orientés Services

Les entreprises ayant de nombreux échanges de données avec d'autres sociétés (nationales ou internationales) ou avec de nombreux partenaires et clients, ont senti depuis une dizaine d'années la nécessité de s'accorder sur des normes pour aider à sécuriser l'information et les processus d'échanges. L'ISO/IEC 27005 est la norme la plus récente en matière de gestion de la sécurité de l'information. Elle est appelée à être très utilisée dans le domaine des SMSI [12]. En 1995 la BSI publie la norme BS 7799 [BSI95] qui s'articule autour de dix grands chapitres qui énumèrent les mesures (une centaine) qui peuvent être prises en matière de sécurité de l'information. En 2000 l'ISO l'adopte officiellement sous la référence ISO 17799 [ISO05a] qui porte désormais le nom de ISO 27002. En 2005 apparait la série des normes ISO 27000 avec la création de l'ISO 27001 [ISO05b] ; il s'agit de la norme BS 7799 à laquelle s'ajoutent les exigences auxquelles doit répondre un organisme pour mettre en place un SMSI, le tout dans une une démarche proche de l'ISO 9001 [ISO08]. En 2008 l'ISO 27005 est mise en place ; c'est une norme de gestion des risques ; elle répond, lorsqu'elle est appliquée, à toutes les exigences de l'ISO 27001. Nous ne détaillerons pas cette norme dans cet article mais nous invitons le lecteur à consulter le site de l'ISO [ISO11] pour en comprendre les mécanismes.

Les difficultés rencontrées pour la gestion des risques dans les systèmes d'information distribués SOA sont fondamentalement liées à la nature même des SOA. Une telle architecture "logique" concerne à la fois les aspects matériel (les serveurs du prestataire de service), logiciel (systèmes d'exploitation, implémentation du protocole SOAP, services eux-mêmes) et réseau (internet, réseau local, routeurs). Certes la norme ISO/IEC 27005:2011 peut déjà prendre en compte ces aspects dans le périmètre d'étude, mais généralement de manière individuelle : datacenter, liaisons, ordinateurs, applicatifs,... Les solutions techniques existantes pour améliorer la sécurité des SI abordent d'ailleurs elles aussi ces points de manière individuelle : redondance de serveurs, liaisons doublées, outils de chiffrement, administration des systèmes, suivi des applicatifs,...

Suite à ce constat nous avons décidé de mener notre étude suivant une démarche de Risk Manager conforme à l'ISO/IEC 27005:2011 avec comme objectif de définir une annexe dédiée à l'interconnexion des SI dans une SOA afin d'enrichir les annexes C et D de la norme avec, respectivement, des menaces et des vulnérabilités spécifiques aux SOA.

3.1 Déroulement de la Norme ISO/IEC 27005:2011

3.1.1 Établissement du contexte

Contrairement aux travaux que nous pouvons trouver dans la littérature, nous ne nous orientons pas vers l'étude de la sécurisation des services web eux-mêmes. Nous étudions l'impact d'une architecture orientée services sur la sécurité des systèmes d'information (SSI) du point de vue des risques liés à la sécurité de l'information. Car si des technologies telles que les services web apportent de nouvelles fonctionnalités, voire sont à même de générer de nouveaux besoins, elles introduisent néanmoins de nouvelles vulnérabilités au sein du SI et, par conséquent, de nouveaux risques pour la Sécurité du SI.

12. SMSI (anglais : ISMS) : Système de Management de la Sécurité de l'information

Notre domaine d'étude concerne l'interconnexion de différents systèmes d'information (au sens large) via l'utilisation de services web. Notre préoccupation n'est pas la sécurité "interne" de ces services web (ex : injection de paramètres erronés) mais l'impact d'une "défaillance" d'un service web sur le SI.

3.1.2 Identification des actifs

Dans ce contexte, les actifs que nous sommes amenés à considérer peuvent être classés en deux catégories : les services web (graphe des flux de données en entrée et en sortie, processus métier qu'ils implémentent,...) et l'infrastructure de communication (systèmes, machines, OS, réseau, plateformes logicielles,...)

3.1.3 Identification des menaces

D'après le glossaire des termes de sécurité de l'information clé du NIST [NIS11], une menace (en anglais *threat*) est définie comme "toute circonstance ou événement susceptible de nuire aux opérations de l'organisation (y compris les missions, les fonctions, l'image, ou à la réputation), aux actifs de l'organisation ou des individus à travers un système d'information via un accès non autorisé, la destruction, la fermeture, la divulgation, la modification de l'information, et/ou déni de service".

Au sens des normes ISO, une menace est un événement qui peut se produire sur un actif ou un ensemble d'actifs. La menace est un événement qui cible un actif en exploitant une ou plusieurs vulnérabilités que possède l'actif. Cet événement peut être de nature délibérée (ex : vandalisme, piratage), accidentelle (ex : rupture d'un câble suite à des travaux de voirie) ou naturelle. Un risque est le potentiel d'une menace à exploiter avec succès une ou plusieurs vulnérabilités particulières et ainsi causer des dommages au système d'information ou à l'organisation. Il se mesure en termes de combinaison des probabilités d'un événement et de ses conséquences. En ce qui concerne notre domaine d'étude, voici une liste (bien évidemment non exhaustive) de menaces qui pourraient altérer le fonctionnement des services web :

- un incident sur le réseau entraîne des dysfonctionnements (délais trop longs, perte de connexion)
- une personne mal intentionnée intercepte des messages et forge de nouveaux messages pour nuire au SI et/ou accéder à certaines informations
- une erreur logicielle (accidentelle ou délibérée) sur un WS provoque l'envoi de réponses erronées

3.1.4 Identification des vulnérabilités

Une vulnérabilité peut être définie comme une "faiblesse dans le système d'information, les procédures de sécurité du système, les contrôles internes, ou la mise en œuvre qui pourraient être exploités ou déclenchés par une source de menace". Cette définition considère non seulement les vulnérabilités des composants logiciels, mais aussi également les aspects organisationnels.

Quand on parle de vulnérabilités des architectures orientées services (SOA) on doit se pencher sur les faiblesses inhérentes à ces technologies et pointer les vulnérabilités maintenant connues (OWASP [OWA], MITRE et le projet CVE [13] [CVE]), mais également

13. CVE : *Common Vulnerabilities and Exposures* ; ce système fournit une méthode de référence pour les vulnérabilités et les expositions (liées à la sécurité de l'information) connues du public.

prendre en compte tous les composants faisant partie de cette technologie et qui sont victimes d'attaques (WS-attacks.org) : le client du web service, le serveur, le moteur BPEL, les intermédiaires, le parseur XML, la validation de schéma (XML), la vérification de la signature, le chiffrement et le déchiffrement.

À partir de ces éléments il est possible de pointer quelques exemples de vulnérabilités et de menaces associées : recopie d'un processus de login vers une autre ressource (usurpation d'identité), usurpation de métadonnées (fichier WSDL, WS-security-policy) (abus de droit), modification du WS-Addressing avec modification du routage (vol de document), le moteur BPEL est sensible aux nombreuses requêtes SOAP (XML flooding).

De manière générale, au sens de la norme ISO/IEC 27005:2011 annexe D, toutes ces vulnérabilités font parties, des types "matériel", "logiciel" ou "réseau" et sont donc déjà implicitement référencées. En outre, un certain nombre de standards et de technologies associées permettent déjà de renforcer la sécurité des WS sur ces aspects (cf. section 2.1).

Notre objectif est de travailler à un niveau de granularité plus élevé et de nous intéresser à la sécurité de l'information en tant que telle (structure du contenu, chaîne de production et de consommation, enjeux économiques et juridiques de l'information) et non pas en tant que "simple" paramètre d'entrée/sortie circulant au sein du SI. Si l'on regarde le service de traitement de l'information (stockage, transformation, transport,...), il est certain que de nombreuses vulnérabilités et menaces se font jour en particulier avec les services de plus en plus présents qui sont communément appelés "**cloud**". Ainsi, il est possible de classer ces vulnérabilités suivant plusieurs catégories :

- ***La qualité de service :*** malgré les conventions de type SLA [14] il n'est pas rare d'être confronté à des pannes fréquentes, des services rendus indisponibles (Amazon EC2, avril 2012), pouvant toucher des milliers d'utilisateurs de par le monde. La seule compensation possible est d'ordre financière alors que votre image a été lourdement dégradée. De même qu'en est-il du client ? Peut-il migrer vers un autre opérateur de services ? Y a-t-il interopérabilité, portabilité, transférabilité lorsque le client veut reprendre ses informations ?
- ***La localisation des données et des traitements :*** ces services externalisés peuvent être situés dans le monde entier, sans qu'il soit possible de pouvoir choisir le pays. Dans la plupart des cas, le prestataire de cloud ignore lui-même où sont les données et où sont exécutés les traitements de l'information ! Un service immédiat et de qualité nécessite une grille de machines sur l'ensemble de la planète, or certaines informations peuvent être acceptables dans un pays mais interdites dans un autre.
- ***La perte de contrôle de l'information*** dont l'origine peut avoir des causes bien différentes. En cas d'***incident*** chez le prestataire de services, est-il toujours possible de récupérer ses données ? Est-ce aux clients ou au fournisseur de services de faire des sauvegardes ? De plus qu'en est-il de la ***réversabilité*** qui est censée permettre au client de reprendre possession de ses données, à tout moment, sans justification. D'autre part en cas de textit**résiliation** du contrat, l'effacement physique des données est rarement complet.
- ***La propriété de l'information :*** quand on traite des informations dans le cloud, on confie le capital de l'entreprise à un tiers. Que se passe-t-il en cas de litige (ex : non-paiement, injonction), s'il cesse son activité (ex : faillite, rachat),... ? Le prestataire a-t-il la possibilité de garder vos données ? A-t-il contractuellement le droit de les

14. *Service Level Agreement*

exploiter ?

– ***La nature de l'information :*** outre les données classiques qui peuvent être traitées dans une entreprise, il en est qui sont particulièrement sensibles : les données stratégiques (recherche-développement) et les données personnelles. En ce qui concerne des données personnelles, elles peuvent être ouvertes par des lois dans un pays alors qu'elles sont fermées dans d'autres (ex : USA Patriot Act, FISAAA [15]), ce qui peut impacter le respect de la vie privée.

3.1.5 Identification des conséquences

À partir de cette liste (non exhaustive) de vulnérabilités nous pouvons maintenant présenter quelques scénarios élémentaires illustrant l'occurrence d'une menace qui exploite une vulnérabilité d'un service avec ses conséquences sur le système d'information. Nous avons distingué les vulnérabilités "classiques" (•) qui peuvent être résolues à l'aide des technologies et mesures de sécurité existantes (cf. section 2.1) des vulnérabilités "organisationnelles" (◦) qui concernent la sécurité de l'information en elle-même.

- Défaut "d'isolation" des services web au niveau du conteneur utilisé ⇝ la défaillance d'un WS peut perturber le fonctionnement des autres WS hébergés dans ce conteneur ⇝ impact potentiel sur la confidentialité, l'intégrité, la disponibilité.
- Usurpation d'identité ⇝ la recopie d'un message contenant les informations d'identification permet de réutiliser ce message une nouvelle fois pour obtenir de nouvelles ressources.
- Falsification des métadonnées du service Web ⇝ une personne malveillante peut falsifier ces métadonnées (WSDL, WS-Security Policy) pour réduire les exigences de sécurité du service web : une obligation de crypter les messages peut être transformée en une transmission des messages en clair permettant ainsi de lire leur contenu.
- ◦ Le fournisseur ne produit pas lui-même le service qu'il propose ; ce n'est qu'un mandataire (avec des sous-traitants) ⇝ problème d'identification des responsabilités en cas de litige.
- ◦ Les données sont stockées sur un territoire étranger ⇝ les lois et réglementations ne sont pas identiques dans tous les pays (ex : SA Patriot Act, FISAAA) ⇝ perte de confidentialité évidente, que les données soient personnelles ou non.
- ◦ Certains clients stockent des données "interdites" ⇝ le service peut être fermé du jour au lendemain (ex : Megaupload 2012, 2e2 2013) ⇝ perte des données même pour les clients respectueux de la loi.
- ◦ Le fournisseur est une entreprise financièrement fragile ⇝ si cette entreprise fait faillite, le service sera interrompu (ex : 2e2 2013).
- ◦ Les procédures de clôture du service (à la fin du contrat) ne sont pas clairement définies ⇝ quid de la réversabilité ? quid de l'effacement physique des données ?

3.2 Proposition d'Évolution de la Norme ISO/IEC 27005

Comme nous venons de le voir, nous traitons la sécurité de l'information dans son ensemble : du contenu jusqu'aux aspects économiques et juridiques. Cette vision de la sécurité de l'information ne remet pas en question les concepts établis dans l'ISO/IEC 27005:2011. Au contraire, elle en est complémentaire car notre proposition consiste à ajouter à l'annexe D de cette norme un nouveau type "service" avec les vulnérabilités et les menaces

15. FISAAA : Foreign Intelligence Surveillance Act Amendments Act

Vincent Lalanne et Manuel Munier et Alban Gabillon

qui lui correspondent et qui tient compte de la technologie des services web mais aussi de l'offre même de service proposé : ce nouveau type "service" peut être vu comme étant un "sur-type" au-dessus des types existants "matériel" , "logiciel" et "réseau".

Ainsi, lorsque l'on fixe un type "service" dans cette nomenclature, les vulnérabilités, et par conséquent les menaces associées, sont vues sous un nouvel angle. Les scénarios sont plus respectueux des faits constatés : ici il est plutôt question de "déni de service" par exemple, plutôt que de "perte de connexion réseau" ; "prestataire de service fournissant des données erronées" au lieu de "défaillance logicielle". De ces vulnérabilités découlent alors des questions simples qui sont au cœur de tout utilisateur de services et qui lui permettent de garantir une certaine "maîtrisabilité" de ses données : Puis-je agir librement sur mes données ? Sais-je où sont mes données ? Sais-je qui a accès à mes données ? Ai-je la propriété de mes données ? Puis-je changer facilement de prestataire ? Dans quel pays sont stockées mes données ? Mes données sont-elles légales dans le pays où elles sont stockées ? Mon prestataire rappelle-t-il le respect de la propriété intellectuelle à tous ses clients et éviter ainsi de se faire fermer ?

Type	Exemples de vulnérabilités	Exemples de menaces
Matériel	...	...
Logiciel	...	...
Réseau	...	...
Service	Pas de support à long terme du fournisseur de service	Le service n'est plus disponible
	Le cycle de vie et les politiques de mise à jour du fournisseur de service sont inconnus	Changement inattendu de l'interface du service
	Hébergement des services dans un pays inconnu	Espionnage, vol de données
	Ne se conforme pas avec les lois en vigueur	Le service n'est plus disponible
	Le fournisseur fait faillite	Le service n'est plus disponible
	Les lois sur la vie privée du pays d'hébergement sont différentes des lois du pays d'utilisation de ces données	Perte de confidentialité
	Permissivité des lois sur la sécurité de l'information	Espionnage, vol de données
	Sévérité des lois sur la sécurité de l'information sont	Le service n'est plus disponible
	ontrat de niveau de service inapproprié	Perte de maitrisabilité du SI
	Pas de procédure de récupération des données	Perte de maitrisabilité du SI
	Manque de réversibilité (migration, interopérabilité)	Perte de maitrisabilité du SI
	Absence de procédure d'effacement des données à la fin du contrat	Vol de données, utilisation non autorisée
	Absence de sécurisation des métadonnées du WS	Falsification du web service
	Possibilité d'appels multiples du WS	Usurpation d'identité
	Absence de traçabilité du service fourni	Absence de confiance dans l'information
Personnel	...	...
Site	...	...
Organisation	...	...

Table 1: Un nouveau type "service" dans l'annexe D de la norme ISO/IEC 27005:2011

Dans le tableau 1 nous avons repris le formalisme utilisé dans l'annexe D de la norme ISO/IEC 27005:2011 pour présenter une série (non exhaustive) de vulnérabilités et de menaces associées spécifiques au type "service". Ces vulnérabilités et menaces tiennent compte des propositions développées dans les sections précédentes.

La notion de "maîtrisabilité" apparaît du coup fondamentale dans la conception des systèmes d'information utilisant des services. L'architecte du SI doit absolument être certain de pouvoir organiser, gérer et donc "maîtriser" ses services, sous peine de rendre son système d'information inexploitable. Il est tout à fait envisageable à partir de ce postulat de définir en plus des trois critères de sécurité fondamentaux CID, un quatrième bien spécifique aux services :

1. **Confidentialité** : prévenir la divulgation non autorisée de l'information (ex : accès en lecture illicite)
2. **Intégrité** : garantir l'exactitude, l'exhaustivité, la validité et la non modification illicite de l'information
3. **Disponibilité** : garantir la continuité de service et de performance du système
4. **Maîtrisablité** : garantir le contrôle total des services utilisés

4 Conclusion

Les modèles d'architectures orientées services (SOA) sont de plus en plus utilisés au regard des bénéfices qu'ils procurent. D'un point de vue gestion des risques du système d'information, ces technologies introduisent néanmoins de nouvelles vulnérabilités de part les accès réseau, l'utilisation de services externes dont nous n'avons pas la maîtrise, les responsabilités liées à la fourniture d'un service, etc... Il existe bien évidemment des solutions pour tenter de sécuriser ces échanges et pour concevoir des services web "fiables". Nos travaux sont complémentaires puisque nous intervenons au niveau de la sécurité de l'information au sens de la norme ISO/IEC 27005:2011. Notre objectif est d'affiner ce processus de gestion des risques sur les spécificités liées aux services dans une SOA en prenant en compte l'information dans sa globalité (jusqu'aux aspects économique et juridique).

Après avoir mis en évidence un certain nombre de scénarios ne pouvant être évités via les mécanismes de sécurité existants, la deuxième phase de nos travaux consistera à proposer un modèle de sécurité pour les communications inter-SI. Nous envisageons pour cela une approche orientée contrôle d'usage telle que nous l'avons déjà expérimentée dans nos travaux [MLR12, Mun11, M. 10] sur les documents intelligents (cf. *Enterprise Digital Right Management*). L'utilisation de métadonnées pour la traçabilité des communications (via ces services) nous permettra également de remonter des indicateurs qui pourront, par exemple, être utilisés pour superviser le système d'information. Ainsi ces mécanismes permettront de protéger l'information lors de son hébergement chez un fournisseur de service. Avec ce modèle nous pourrons ainsi garantir non seulement la Confidentialité, l'Intégrité et la Disponibilité des données, mais également la **Maîtrisabilité** de l'information.

L'implémentation de ce modèle fait bien évidemment partie de nos perspectives. C'est pour cette raison que nous nous sommes orientés vers la technologie Web Services puisqu'un grand nombre de spécifications dédiées à la sécurité ont été développées autour du protocole SOAP. En outre, ces spécifications (présentées à la section 2.1) sont "ouvertes", c'est-à-dire que nous sommes libre de mettre en œuvre nos propres modèles et politiques de sécurité.

Références

[BBF+08] Mark Bartel, John Boyer, Barb Fox, Brian LaMacchia, and Ed Simon. XML Signature Syntax and Processing. Technical report, W3C Open Source Software, 2008.

[BFKH05] Konstantin Beznosov, Donald J. Flinn, Shirley Kawamoto, and Bret Hartman. Introduction to web services and their security. *Information Security Technical Report*, 10(1) :2 – 14, 2005.

[BMPS10] Elisa Bertino, Lorenzo D. Martino, Federica Paci, and Anna C. Squicciarini. *Security for Web Services and Service-Oriented Architectures*. Springer-Verlag Berlin Heidelberg, 2010.

[BSI95] BSI British Standards. BS 7799 :Part 1 :1995 - Code of practice for information security management systems. Technical report, BSI British Standards, 1995.

[CKPM05] Cantor, Kemp, Philpott, and Maler. Security Assertion Markup Language (SAML). Technical report, OASIS, 2005.

[CVE] CVE : Common Vulnerabilities and Exposures. Technical report.

[ER02] Donald Eastlake and Joseph Reagle. XML Encryption Syntax and Processing. Technical report, W3C, 2002.

[GPC05] Christian Geuer-Pollmann and Joris Claessens. Web services and web service security standards. *Information Security Technical Report*, 10(1) :15 – 24, 2005.

[ISO05a] ISO. ISO/IEC 17799:2005 - Code of practice for information security management. Technical report, ISO/IEC, 2005.

[ISO05b] ISO/IEC. ISO/IEC 27001:2005 : Information security management systems. Published, ISO, 2005.

[ISO08] ISO/IEC. ISO/IEC 9001:2008 : Quality management systems. Published, ISO, 2008.

[ISO11] ISO/IEC. ISO/IEC 27005:2011 : Information security risk management. Published, International Organization for Standardization (ISO), Geneva, Switzerland, 2011.

[M. 10] M. Munier. A multi-view approach for embedded information system security. In *CRiSIS*, pages 65–72. IEEE, 2010.

[MG03] Tim Moses and Simon Godik. eXtensible Access Control Markup Language (XACML) Version 1.0. Technical report, OASIS, 2003.

[MLR12] Manuel Munier, Vincent Lalanne, and Magali Ricarde. Self-protecting documents for cloud storage security. In *TrustCom*, pages 1231–1238. IEEE, 2012.

[Mun11] Manuel Munier. A secure autonomous document architecture for enterprise digital right management. In *SITIS*, pages 16–23. IEEE, 2011.

[NIS11] NIST. NIST IR 7298 Rev. 1 : Glossary of Key Information Security Terms. Technical report, National Institute of Standards and Technology (NIST), Fév-2011.

[NKHB+06] A Nadalin, C Kaler, P Hallam-Baker, R Monzillo, and Et Al. Web Services Security : SOAP Message Security 1.0 (WS-Security 2004). *OASIS Standard*, 200401(February), 2006.

[OAS06] OASIS. Web Services Security (WSS). Technical report, OASIS, 2006.

[OWA] Owasp : The open web application security community. Technical report.

[Use99] UserLand. XML-RPC. Technical report, UserLand Software, Inc., 1999.

[W3C03] W3C. SOAP version 1.2. Technical report, W3C Open Source Software, june 2003.

Chapitre 7

DPM 2013
Legal Issues about Metadata Data Privacy vs Information Security

Sommaire

7.1 Introduction

Cet article a été présenté à la 8ème édition de l'atelier DPM [28] qui s'est tenu à Egham (Grande Bretagne) au Royal Holloway, Université de Londres, le 12 et 13 Septembre 2013. Cet atelier se tenait pendant la 18ème édition du Symposium européen sur la recherche en sécurité informatique ESORICS 2013 [29].

7.2 Présentation

Pour les besoins de nos travaux nous utilisons la notion de métadonnées pour mettre en œuvre des mécanismes de E-DRM [30] dans un environnement de documents intelligents.

Ces métadonnées nous permettent d'une part de définir des règles de sécurité contextuelles et d'autre part d'assurer la traçabilité des informations. Leur utilisation peut toutefois avoir des conséquences sur le plan juridique, notamment en ce qui concerne les métadonnées qu'il est possible d'enregistrer (privacy - données personnelles), la manière dont elles doivent être stockées (par exemple la notion de valeur probante en cas de litige) ou les traitements informatiques dans lesquels elles peuvent être impliquées. Nous poursuivons le partenariat avec la société BackPlan™ par l'exploitation des métadonnées qui sont générées lors de la construction de projets Oil & Gas.

L'objectif de cet article est de positionner nos travaux par rapport à ces aspects juridiques. Il marque le premier résultat de collaboration avec le Centre de Recherche et d'Analyse Juridiques [31] de l'université de Pau et des Pays de l'Adour.

7.3 Les thèmes abordés

Quels que soient les domaines d'activité, les nouvelles technologies de l'information nous ont amené à échanger et à stocker des informations en quantité de plus en plus importante. Leur contenu a également évolué. Les données sont plus complexes (notions de document structuré, d'archive, voire même de projet complet). Des données dites publiques cohabitent maintenant avec des données plus confidentielles (notion de restriction d'accès). Sans compter que nous emportons nos données sur des clés usb ou dans nos smartphones et les partageons via des communications sans fil. Dans cette société de l'information, la fiabilité des données manipulées est devenue un enjeu majeur du point de vue de la sécurité.

28. http ://research.icbnet.ntua.gr/DPM2013/
29. http ://homepages.laas.fr/esorics/
30. Enterprise Digital Right Management - DRM pour l'entreprise
31. C.R.A.J. - EA 1929

Concernant la confidentialité, les modèles de contrôle d'usage ont introduit la notion de contexte afin de pouvoir exprimer des règles de sécurité dynamique dans une politique : temporal contexts, spatial contexts, prerequisite contexts, provisional contexts,... Pour activer ou désactiver ces contextes, le système d'information doit collecter et stocker diverses métadonnées : date, adresse IP utilisée, localisation de l'utilisateur,... Ces métadonnées pourront ultérieurement être utilisées à des fins de traçabilité et notamment servir de preuve en cas de litige. Certaines de ces métadonnées peuvent cependant être considérées comme étant des données personnelles de l'utilisateur.

Notre article est organisé de la manière suivante. La section Motivations présente les motivations qui nous ont conduit à proposer une nouvelle architecture E-DRM ainsi qu'un exemple d'utilisation des métadonnées et des mécanismes de contrôle d'usage dans le cadre de documents d'entreprise. Cet exemple a été développé avec notre partenaire BackPlan™ [32]. Dans la section Metadata & Information Security nous développons notre approche orientée documents intelligents en détaillant, en particulier, l'utilisation des métadonnées pour mettre en œuvre les différents mécanismes de sécurité. La section Metadata & Legal Issues, via des exemples de jurisprudence, met en avant la nécessité de formaliser les métadonnées nécessaires pour mettre en vigueur la politique de sécurité et le cadre dans lequel elles peuvent être utilisées afin de respecter la réglementation en vigueur. Finalement, la section Conclusion termine notre article et présente nos perspectives actuelles de travail.

7.4 Les propositions

Notre approche repose fortement sur la notion de métadonnées, tant au niveau de la collecte (traçabilité) que pour l'activation des contextes (règles de sécurité dynamiques) ou, à terme, le calcul de différents indicateurs au fil des utilisations (indice de confiance). A noter que, pour le moment, les éléments d'étude fournis dans cette section ne concernent que la loi française. Pour le passage à l'échelle de nos travaux, les spécificités propres à chaque pays devront bien évidemment être étudiées.

Les métadonnées soulèvent trois types de difficultés : leur collecte, leur stockage et leur utilisation. En premier lieu, elles posent des problèmes s'agissant de leur **collecte**. Bien souvent, cette collecte se fait sinon à l'insu, du moins dans l'ignorance de la personne concernée. Se posent alors la question du droit d'accéder aux informations contenues dans les métadonnées ainsi et surtout que la question du droit de savoir que l'information est collectée. Une autre difficulté surgit aussitôt : celui qui collecte les métadonnées n'a même pas conscience lui-même d'effectuer une telle collecte. Ce n'est que postérieurement à cette collecte que les éléments recueillis seront

32. BackPlan™ , Project Communication Control http ://www.backplan.fr

découverts et utilisés.

Enfin, les métadonnées soulèvent le problème de leur **utilisation**. Cette utilisation peut être de bonne foi, comme par exemple pour la mise en œuvre de mécanismes de sécurité, mais elle peut également être de mauvaise foi, les métadonnées étant détournées de leur utilisation initiale ou faisant l'objet d une falsification.

En droit civil, la preuve d'un fait juridique peut se faire par tout moyen, à certaines conditions toutefois. Il est notamment exigé que la preuve soit rapportée avec loyauté et que le mode de preuve soit fiable. Si les règles sont sensiblement différentes en droit pénal, l'exigence de fiabilité est assurément commune aux deux domaines. Cette exigence de **fiabilité** est bien évidemment au cœur de la recevabilité d'un document électronique à titre de preuve. L'article 1316-1 du Code civil prévoit ainsi que, en matière de preuve des actes juridiques *« l'écrit sous forme électronique est admis en preuve au même titre que l'écrit sur support papier, sous réserve que puisse être dûment identifiée la personne dont il émane et qu'il soit établi et conservé dans des conditions de nature à en garantir l'intégrité. »* C'est donc sans surprise que les rares décisions se référant aux métadonnées exigent la fiabilité de celles-ci.

On peut bien évidemment imaginer que les métadonnées peuvent servir d'élément de preuve dans les domaines les plus variés. Elles peuvent ainsi servir à prouver l'absence sur son lieu de mission d'un salarié. Ainsi en matière de télétravail, les métadonnées ouvrent la possibilité de déterminer si le salarié a bien respecté son temps de travail. Se posent alors les questions de la distinction entre le temps personnel et le temps de travail, du respect de la vie privée du salarié ainsi que la question de la délimitation du pouvoir de direction de l'employeur.

7.5 Conclusions

La confiance dans les données que nous manipulons tous les jours est un des enjeux majeurs de la société de l'information. Il existe de nombreux mécanismes qui nous permettent de collecter, stocker et traiter d'énormes quantités de données et, surtout, de données sur ces données : les métadonnées. Les métadonnées sont un outil indispensable pour la sécurité de l'information : contrôle d'usage pour le partage de documents, investigation judiciaire, preuve en cas de litige,...

Les possibilités technologiques ne doivent cependant pas faire oublier les problèmes juridiques. L'objectif étant de mettre en œuvre une politique de sécurité et d'assurer la traçabilité des informations, il est indispensable de respecter la réglementation existante quant aux métadonnées qu'il est possible d'enregistrer (cf. données personnelles), comment elles doivent être stockées (cf. valeur probante) et les traitements informatiques dans lesquels elles peuvent être impliquées.

Par le biais de cet article nous voulons sensibiliser les gens aux dérives potentielles liées

à l'utilisation de telles métadonnées. Certains travaux ont déjà été effectués pour préserver la vie privée. Un exemple est l'anonymisation des données. Ceux-ci ne sont toutefois pas toujours adaptés à notre problématique de contrôle d'usage où justement certains indicateurs ne doivent pas être anonymes. Dans le cadre des E-DRM il s'agit de projets professionnels entre partenaires. Notre approche consiste donc plutôt à formaliser ce "contrat de collaboration". Pour la communauté informatique, ceci se fera en terme de langage de spécification des métadonnées devant être collectées, par quels moyens, comment elles seront stockées et quelle en sera l'utilisation. Pour la communauté juridique, il faudra d'abord qualifier les métadonnées : doivent-elles être traitées comme des données "traditionnelles" ou bien doivent-elles bénéficier d'un régime juridique spécifique ? Une fois défini le cadre juridique, nous pourrons étudier ensemble sous quelles conditions il est possible d'utiliser des métadonnées et, dans l'autre sens, quelles sont les métadonnées nécessaires pour appliquer certaines lois. Dans l'exemple de la construction d'un projet Oil & Gas comme nous le décrivons à la section Exemple d'Application, il faudra dorénavant inclure dans le contrat de collaboration entre les entreprises l'insertion ou la suppression de ces métadonnées. Doivent elles faire partie intégrante des documents délivrés à la fin du projet ?

7.6 Texte de l'article

Legal Issues about Metadata Data Privacy vs Information Security

Manuel Munier[1], Vincent Lalanne[1], Pierre-Yves Ardoy[2], and Magali Ricarde[3]

[1] Univ Pau & Pays Adour, LIUPPA, Mont de Marsan, France
[2] Univ Pau & Pays Adour, CRAJ, Pau, France
[3] BackPlan Company, Project Communication Control, Pau, France

Abstract. For the purposes of our work we use the concept of metadata to implement enterprise digital right management mechanisms in an intelligent document environment. Such metadata allows us to firstly define contextual security rules and secondly to ensure the information traceability. However, its use may have legal implications, especially with regard to metadata that can be stored (see personal data, privacy), how it should be stored (see probative value in case of litigation, digital forensics) or computer processing in which it may be involved. Another topical issue is the storage and the processing of data using a service provider: the cloud. We must ensure, however, that this solution does not lead to a loss of information controllability for the company. This article aims to position our work with respect to these legal issues.

Keywords: privacy; metadata; information security; socio-economic issues;

1 Introduction

Whatever business areas, new information technologies (ADSL, laptops, smartphones, tablets,...) lead us to exchange and store more and more information. Their content has also evolved. Data is more and more complex (notions of structured documents, whole archives, or even complete projects). Nowadays, public data is sometimes combined with more confidential data (notion of access restriction). Moreover, we carry our data on usb stick or in our smartphones and share it via (possibly unsecured) wireless communications like 3G, wifi or bluetooth. In the information society, the reliability of the data we handle has become a major issue in terms of security.

Security criteria most commonly used are confidentiality (assurance that information is shared only among authorized persons or organizations), integrity (assurance that the information is authentic and complete), availability (assurance that the systems responsible for delivering, storing and processing information are accessible when needed, by those who need them) and traceability (ability to chronologically interrelate uniquely identifiable entities in a way that is verifiable).

Regarding confidentiality, usage control models introduced the notion of context in order to express dynamic security rules in a policy: temporal contexts,

spatial contexts, prerequisite contexts, provisional contexts,... To enable or disable these contexts the information system must collect and store various metadata: date, IP address used, user location,... This metadata can later be used for traceability purposes, including use as evidence in case of litigation. Some of this metadata can, however, be considered as personal data of the user and thus bring privacy concerns. Metadata is one of the important keys to the success of the data warehousing and business intelligence effort since the mid nineties of the last century. Data warehouses are designed to manage and store the data whereas the business intelligence focuses on the usage of the data to facilitate reporting and analysis [1,2]. The term metadata refers to "data about data". However, the concept of metadata is not (yet) a well-known concept of the law.

The remainder of this paper is organized as follows: Section 2 presents our motivations for the use of metadata and usage control mechanisms to enforce information security; this section also presents a concrete case study in the context of business documents; we developed this example with our partner company BackPlan™ [1]; in Section 3 we present two areas of our research activities related to information security using the concept of metadata and how metadata can be useful for usage control, traceability and information system monitoring; Section 4, using examples of jurisprudence, highlights the need to formalize the metadata necessary to enforce the security policy and the framework in which they can be used order to comply the regulations; Section 5 presents some socio-economic issues underlying the storage of data (and metadata) in the today information society; Section 6 concludes the paper and presents some of our perspectives.

2 Motivations

As we stated at the beginning of this article, information technology (IT) allows us to share more and more information in the form of documents whose structure becomes more complex. Whether in the context of a "simple" information dissemination (unidirectional communication from a content provider to the users) or a collaborative work environment (several actors interact to complete tasks with a common goal but possibly different objectives), it is therefore necessary to implement security mechanisms that go beyond a simple access control: usage control (how partners can use a document: obligations, workflows, delegation rules,...), information consistency management (e.g. some documents may reference others), traceability (monitoring of actions, metadata attached to information),...

We do not give here all the details of our case study. This section focuses on the use of metadata to improve information security (traceability, usage control) and legal issues that may arise.

[1] BackPlan™ , Project Communication Control
http://www.backplan.fr

2.1 Sample Application

Consider an Oil & Gas project as the construction of a pipeline or an oil installation. The information system consists of numerous documents, it has a central role, its structure and development evolve along with the progress of the project: the documentation must always precede action (design, work procedures). Documentation is a requirement at the closure of the project along with verifications (records, the minutes,...). The document evolves at the same rate as the project. These documents are specifications, drawings, records of expertise, procedures, records,...

Business Aspects Such a project obviously involves many partners and subcontractors. Here is a representative example of a project timeline (Fig. 1):

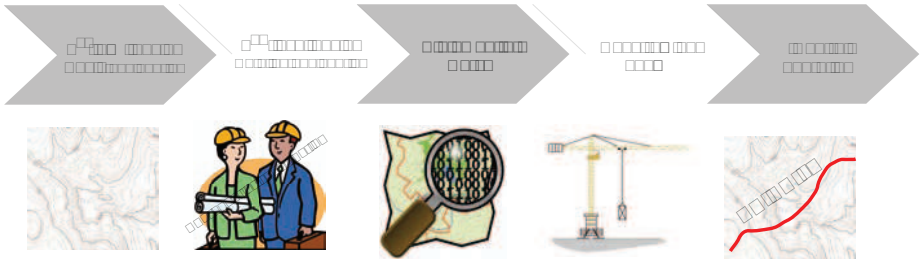


Fig. 1. Oil & Gas project timeline

- 1st level design: Basic Engineering. This step is performed by the land surveyor who makes a topographical survey of the site where the work will be done.
- 2nd level design: Detailed Engineering. This is the design phase of the project itself with the aim, in particular, to minimize the environmental and human constraints; it is performed by an engineering company that will plan the work and the construction will be launched from this plan. This level involves various partners (civil engineering, pipefitters, instrumentation engineering, utilities,...) who share many documents.
- Statement of works: numerous buried and air, public or private networks, go through the territory (water, electricity, gases, dangerous products, telecommunications, irrigation,...). Further to accidents, it is imperative to localize very exactly their position. So, during the realization of a project of construction, gas pipeline for example, companies have the obligation to question a centralized information system common to all the French territory.
- Construction phase: it is based on engineering documents and work procedures. It comes with many documents that are intended to demonstrate

compliance of the book in terms of quality and regulatory standards (e.g. multifluid standard, water code, capacity under pressure). As-builts will have to justify the differences for the administration.

– At the end of construction the land surveyor will come again, and verify the topographic survey: this is a control operation of a project to verify the differences from the planned location and update the data (to know where everything is). The engineering documents then pass status "As built". This operation can also update the geographic information system (GIS) of the place. As-builts must be attached to requests for authorization to operate sent to administrations (also signed by the legal representative).

As we have stated, the partners will handle many documents. Because of the nature of this type of project, a multitude of corporate associations has to work on the same documents. Where from requires it to manage the communications between these interfaces. Besides, it will be necessary to be able to guarantee that each works with documents "up to date" or that the last modifications in date were well taken into account before the "publication" of certain results (cf. usage control and collaborative work management). There can also occur unforeseen circumstances during the project.

Information Security Aspects We propose to improve the security of information in two directions: metadata management and usage control.

Metadata It could be used to "bind" reviews, certifications, good practice guides, standards ,and other minutes to design documents and reports within the information system. The aim is to improve traceability, both in the design process (concept of workflow) in case of litigation (concept of proof of conformity, digital forensics). Take for example a phase control such as checking the welding of a pipe[2]; it would be interesting to use metadata to improve the traceability of the process for purposes of validation and/or evidence: photos geotagged (to certify checkpoints), metadata associated with the plans,... Since several partners are working on such a project, everyone could also attach some metadata to the information: confidence and trustworthiness indicators, impact risk of a change,... This metadata would permit to calculate various performance indicators for monitoring the partners' tasks or metadata to the information they produce.

In case of transfer of work this metadata would allow to improve the follow-up of the project towards the buyer: operations history, context decisions (standards, studies), how and why they have been taken, etc...

Usage Control Here are some examples of security rules that we would implement to control how partners use documents:

[2] The sections of pipe are welded every 12 to 15m. These welds should be checked: radiography, analysis by a certified individual, hydraulic tests. These controls are spread over time and generate many records that are, once made, a legal value.

- It is possible to write a deliverable of the project only if confidence in the various technical documents exceeds a certain threshold. It is a dynamic access control based on trust (whether in a document or a partner).
- The security rules may prohibit access to parts of the document based on location data. This prevents, for example, on a site (or train) an unknown person takes sensitive information over the shoulder of someone.
- A responsive access control may require a partner (or subcontractor), via a mechanism of pre-obligation to accept the terms of a contract (non-disclosure agreement, delegation of responsibility, deadlines,. . .) before accessing a plan and contribute to the design.
- A user control would be to require a partner to complete parts of design documents (e.g. inform the radii of curvature of the pipe, write the study of soil before drilling) before a deadline if he wants to stay a project member (notions of punishment and penalty).
- The usage control can also define collective obligations. For example, each partner must have reread each concept study in which they participate at least 7 days before the deadline for validation.

Legal Issues These are just some examples of opportunities. But the use of metadata is not easy to understand from a legal point of view ! First, regarding the collection and storage, some of this metadata can be in the domain of personal data (including geolocation). But their use to enable/disable contextual security policies (permissions, obligations,. . .) or calculate some indicators (e.g. trust in a partner, document quality) are automatically processes and are therefore subject to a number of regulatory frameworks. Add to that the concepts of accountability and sanctions mentioned above and it is obvious that metadata has now become essential elements of information systems. They should be considered as full data and be secured along with "classic" information.

Section 4 gives more details on requirements to formalize metadata necessary to the security policy and how metadata can be used. One of the objectives is that this metadata can be used as evidence in case of litigation (cf. probative value) while respecting the laws on privacy.

2.2 BackPlan™

BackPlan™ is a French company providing document management services and collaboration workflow applications to improve project communication, transparency across the project, ability to manage schedule and risks, reliable indicators and regulatory compliance. From the engineering phases to the construction phases, projects involve different companies. All of them will use the collaboration solution BackPlan™ to ensure consistency of information across the project and a complete audit trail of project communication. BackPlan™ document management services are currently provided on a server hosted in a data center: the document registry.

Using metadata (as data about data) would allow BackPlan™ to enhance existing services and to offer new services to their customers. During the course of the project, metadata would be used to calculate many indicators for project progress, compliance with deadlines, completion status of the various documents,... Once the project is completed, the company makes and delivers to its customers the case-file containing all the business documents for project traceability. In terms of risk management, in case of litigation this information can be used to identify those responsible for error or prove that during the construction phase of the project the standards in force at that time have been complied with.

3 Metadata & Information Security

Our research activities take place in the field of information system security. Issues related to the legal framework for the use of metadata were raised when we wanted to implement contextual security rules to perform usage control within two areas of research: the development of a secure autonomous document architecture, and the study of service oriented architecture security. This metadata will obviously be used during the life cycle of the document to perform usage control. But they can also be accessed later for traceability of actions performed and thus serve as evidence in case of litigation or for digital forensics.

This section focuses on the use of metadata to improve information security. In our work this "data about data" allows us to implement security mechanisms such as dynamic security policies, collaborative work management, calculation of confidence and trustworthiness indicators,... The use of metadata is not, however, trivial in terms of the law. As discussed in Section 4, these technological possibilities must not make us forget the legal issues. In addition, we will see in Section 5 that this metadata can be more important thant the data to which it is associated.

3.1 Intelligent Document Architecture

As part of our research we developed a multi-view model for secure data warehouse [3] and we proposed E-DRM[3] architecture based on secure autonomous documents [4]. While "traditional" information systems centralize all the data on a server which users must connect, we have chosen to define an approach where security mechanisms are relocated closer to the user. However, unlike "conventional" DRM architectures that require the use of a dedicated player (which is responsible for enforcing the security policy), we decided to embed these security mechanisms within the document following object-oriented approach: a document is an autonomous entity capable of ensuring by itself the security of the information it contains and controlling how this information is used. Such a document is a kind of information system on its own embedding both a data

[3] E-DRM: Enterprise Digital Right Management

warehouse and various security modules (access control, usage control, metadata,...). Users can thus exchange the document directly and safely without having to connect to a central site.

The core of our architecture, namely the security kernel, relies on the OrBAC model [5] to express security policies in terms of permissions, prohibitions and obligations between a subject, an object and an action. These security rules are dynamic, that is to say, they can be "adapted" to the context of the actions [6,7,8]: activation or deactivation of rules, the execution of an action triggers the insertion of an obligation,... OrBAC model supports various kinds of contexts: *temporal context* (depends on the time at which the subject is requesting for an access to the system), *spatial context* (depends on the subject (geo)location), *user-declared context* (depends on the subject objective or purpose), *prerequisite context* (depends on characteristics that join the subject, the action and the object) and *provisional context* (depends on previous actions the subject has performed in the system). The embedded information system must therefore provide the information required to check that conditions associated with the context definition are satisfied or not. To do this, either it has direct access to the host system (e.g. a global clock to check the temporal context) or it uses metadata carried by the actions and the nodes contained in the embedded database.

Our approach therefore relies heavily on the concept of metadata, both for collection (traceability) and for context activation (dynamic security rules) or, eventually, computing various indicators throughout the uses (confidence, trustworthiness) as works published in [9,10,11,12].

3.2 Service Oriented Architecture Security

Service Oriented Architectures (SOA) offer new opportunities for the interconnection of systems. Opening its Information System to the "world" is not insignificant in terms of security. Whether to use available services or provide its own services, new technologies have introduced new vulnerabilities and therefore new risks. Our work aims to propose an approach for risk management which is based on the ISO/IEC 27005 standard: we propose a development of this standard so that it can fully take into account the type "service" as web services and cloud services [13].

To develop a security model for communications in inter-organisational information systems we also use a usage control oriented approach as presented above. Again, the use of metadata for traceability of communications (via these services) allow us to compute indicators that will be used to monitor the information system [14]. Our goal is that companies can keep control over their information, despite the use of cloud technologies.

3.3 Data Privacy Concerns

Preliminary tests on our self-protecting document architecture, performed with predefined metadata, allowed us to implement dynamic security rules for usage

control. These experiments have also led us to concern ourselves with the privacy of metadata.

According to the ISO 8402 standard, traceability is the ability to trace the history, use or location of an entity by means of recorded identifications. In our architecture, an entity corresponds to an autonomous collaborative work document. History, identification, registration and use are relevant concepts. As we mentioned above, user localization can be an interesting metadata for expressing contextual security rules. But a document could then reveal the presence of a user at a given time in a certain place, or the various revisions which followed one another leading to the final of a contractual document ! So we must not only focus on the metadata that is collected, but also protect them from unauthorized use and control how it can be used (e.g. automatic computation of indicators). Basically, the problem of leakage and/or misuse of information is already known. However, in our approach, the "massive" collection of metadata "of any kind" can effectively exacerbate the problem.

A concrete example is presented in the article [15] entitled "Metadata: the ghosts haunting e-documents". This story demonstrates the risks of exchanging files with embedded data in negotiating a contract. During negotiations, partners used a common word processing program, Microsoft Word, to edit and propose revisions to the contract, and they utilized the program's track changes feature to allow the other side to see the specific changes proposed. They e-mailed the electronic draft, complete with embedded data, back and forth to each other between rounds of revisions. But without using anything but Microsoft Word's inherent functions, someone can reveal hidden internal comments from adverse party concerning terms of the contract, negotiating positions,... In doing so, a partner can be aware of confidential business information and use it to to pressure his opponent.

This article also raises an interesting question: is that partner (in this case a lawyer) bound by the same obligations that apply when documents in a misaddressed envelope are received or, conversely, is the partner free to use and review the embedded information ?

4 Metadata & Legal Issues

As it has been said from the beginning of this article, our works rely highly on the concept of metadata, both in the collection (traceability) for activation contexts (dynamic safety) and, in the course of time, the computation of various indicators over time (confidence rating or trustworthiness value). Note that, for the moment, the elements of study in this section relate only to the French law. For the scaling of our work, the specificities of each country should obviously be considered.

The concept of metadata is not a well-known concept of the law. Composed of the Greek prefix *meta-* referring to the reference to itself, the term metadata refers to data within data, data which defines limits or describe other data. That data is varied and can include durations, dates, places, elements to identify peo-

ple,... In a very precise domain, however law gives a definition to metadata. Article L.127-1, 6° of the Environmental Code actually specifies that it is the *"information describing sets and spatial data services, making possible their discovery, their inventory and their use"*. However, this definition, the first one put by a legal text, does not report the diversity which hides under the term of metadata.

Metadata raises three types of difficulties: their collection, their storage and their use. First of all, problems raise with regards to their collection. Very often, its collection is done unbeknown to the authors, at the least in the ignorance of the concerned entities. Concerning the law, this raises the question of the right of access to information contained in metadata and above all the question of the right to know the information is collected. Another difficulty arises straigh away: the entity who collects the metadata is not even aware he is making such a collection, it is only, subsequent to the collection, that the meditative elements are discovered and used.

Secondly, metadata raises the question of its storage. This storage implies guaranteeing conservation, not only of the authenticity of the metadata (which refers to the question of the collection and the reliability of the source), but also of its integrity, its stability. What is at the heart of the matter is ultimately to require the reliability of the metadata. It is not only important that it is stored in good conditions, it is also necessary as to ensure its availability, its accessibility, what is needed to be know, by whom, for how long and under which conditions.

Finally, metadata raises the problem of its usage. Such a usage may be done in good faith, as for example in the implementation of security mechanisms, but it can also in bad faith, the metadata being diverted from its original usage or be subject to falsification.

Potentially metadata affects all fields of the law and it is difficult or even useless to apprehend it in the abstract. It consists in identifying areas where the use of metadata is likely to raise specific questions without claiming completeness. Also note that this is a primary questioning limited to the French domestic law.

4.1 Evidence of Law

It is mainly in the field of the law of evidence that the rare decisions of the court of appeal which refers to metadata are found.

In civil law, proof of a legal act can be give by any means, nevertheless within certain conditions. It is in particular required that the evidence be reported fairly and that the type of evidence is reliable. Even if the rules are significantly different from criminal law, the reliability requirement is definetely common to both domains. This reliability requirement is obviously at the heart of the admissibility of an electronic document, proofwise. Article 1316-1 of the Civil Code thus provides that, for proof of legal acts *"the writing in electronic form is admissible in evidence equal to a written document on paper, provided that the person who issued the document and its establiment and storage are executed under the conditions so that its integrity can be duly identified.* It is thus not a surprise that the rare decisions referring to metadata require its reliability.

This is foremost the case of the Ruling of the First President of the Paris Court of Appeal on 25 October 2011, Juris-Data N°2011-025553 and October 25, 2011 (unpublished, N°09/14462, 09/14501)[4]. In this specific case, a company filed a legal complaint against the search and seising in its premises authorized by local competent authorities with regards to the investigation of anticompetitive practices. Even if the legal problem did not directly address the issue of metadata, the Court nevertheless adopted the argument of the Competent Authority at the end of which *"the structure of a particular Outlook mail file and the obligation not to change the state of the computer visited nor the characteristics of a file (metadata in the file itself: title, author size, dates, location, and signature,...) necessarily imply the complete seising of the mail file after verifying that it contains elements falling within the scope of the authorization."* In all cases the emphasis is on the requirement that it should seize the entire Outlook messaging so as not to affect the reliability of the input by an alteration of metadata contained in the messages.

In a judgment of the Court of Appeal of Versailles, Ch.1, sect. 1 of 30 September 2010 (unpublished, N°09/03831), an applicant pressed on the metadata contained in photographs to prove his position as author of the photographs. The Court retained the metadata elements as evidence noting that these included the identity of the author, the date and time of shooting, the name of the manufacturer of the device, the model of the latter and the description of the camera settings. If metadata may in itself not demonstrate the source of the photographs, it is however the elements which can establish the paternity and can therefore be very useful in the case of a counterfait lawsuit. Within the same context, the Labour Chamber of the Court of Appeal of Rennes retained, in a judgment of 20 September 2011 (unpublished, RG N°10/05183), that the production of metadata may be used to determine the creation date of an advertising brochure while noting that the reliability of the latter is not discussed.

It is of course conceivable that metadata can serve as evidence in a variety of areas. It can thus be used to prove the absence of an employee. In telework, metadata opens up the possibility to determine whether the employee has met with his working hours. Thus, arise questions of the distinction between personal time and working time, respect of privacy and where the bounderies of the employer's power are situated.

Beyond relations between employer and employees, metadata may be of interest in relationships between business partners or directors of a company. It could for example be used to demonstrate that this officer or that employee was responsible for the disclosure of a trade secret (Article L.621-1 of the Intellectual Property Code and L.1227-1 of the Labour Code).

Moreover, in the context of the mission of prevention and repression measures of illegal downloading which lie within the Hadopi law, metadata can be useful to show which person is actually responsible for downloading illegal copyright protected work. In July 2013, French legislators struck down the heavy-handed

[4] See also Orders of November 15, 2011 or August 31, 2012 from the First President of the Paris Court given in matters relating to anticompetitive practices.

Hadopi online copyright law. Under the law's "three strikes" rule, users who violated copyright restrictions three or more times could be punished by having their Internet connections cut. But Hadopi suffered great controversy when France's highest court, the Constitutional Council, declared access to the internet a basic human right. French legislators are now seeking policy reforms that will shift the focus of law enforcement towards commercial piracy issues.

These are obviously only a few possible usages of metadata, an exhaustive list does not seem possible or at least very delicate.

4.2 Privacy and Individual Liberties

Indeed, metadata may contain information about people, whether it concerns identification or location. For instance, a picture of a group of friends can be published on the Internet (we recall that The Internet is a public space) it may contain the names of the people in the picture, the place where it was taken, the date and time of the snapshot (tags on Facebook photos).

So these are issues that concern not only image rights (assuming the image was published without the agreement of different people), but also the respect of someone's privacy set by Article 9 of the Civil Code. For a number of reasons, in fact, a person may want a certain amount of information not to be disclosed. (such date, such time,...).

If the concept of metadata is just beginning to be understood by the law, it has on the other hand understood those of personal data. It concerns information that can directly or indirectly identify individuals (*Information Act N°78-17 of 6 January 1978, Directive 95/46/EC of October 23, 1995*). The concept of metadata does necessarely correspond to the concept of personal data insofar metadata may contain information which is not necessarily information which can be used to identify a person (information about location, time spent on a file,...). They may, however, raise questions which may be of interest to privacy.

As such, metadata collection is likely to be considered as a collection of personal data. In this sense it seems possible to read Deliberation N°2011-423 of 15 December CNIL 2011 authorizing the company GEOLSEMANTICS to implement, on a trial basis, as part of a research project, the treatment of personal data, necessary for the development of a tool, called SAIMSI (eng: follow adaptive inter-lingual and multi-source information). The metadata items concerned are those attached to the collected documents, that is to say, those *"corresponding of how the information was collected (if applicable: document URL source, date of registration, date, time and place of the issue body text and the source)"*.

4.3 Digital Protection and Intellectual Creations

The French DADVSI Act of 1 August 2006 introduced the ability to protect intellectual works by systems which limit or prohibit any copying. In French law, these devices are called Technical Protection Measures (MTP), better known by the acronym DRM (Digital Right Management). The Intellectual Property Act has developed a complete,relatively complex system, designed to ensure that

such DRM can not be used by producers or publishers aimed at anticompetitive purposes unrelated to the protection of copyrights. The presence of DRM should thus not prevent interoperability, that is to say the ability for the works to be read by the most diverse materials.

5 Metadata & Socio-Economic Issues

In the information society today, metadata becomes sometimes more important than the data which it is associated. Whether in the field of privacy (personal data) or professional (business data of a company), many companies have developed their business on it. Lately, the media focus on large multinational companies such as Google, Facebook or Microsoft regarding the protection of privacy and personal data. This is a hot topic that scares the public. In the professional world, same issues arise about critical data of companies (e.g. research & development, business strategy). This is the case for example of the BackPlan™ company whose business uses metadata on information exchanged between participants for project communication control.

The objective of this section is not to criticize the practices of a particular country or to denigrate the work of a particular company. We just want to highlight the socio-economic issues about information in today's society and the need to harmonize the laws of different countries to define an international legal framework.

5.1 Data are Future's Power

We live in a transitional period, the digitization of everything: people, society, organizations, knowledge, interactions,... Data is the basic building block of the information society. Its quantity is growing exponentially: we are talking about Big Data. The physical infrastructure of the information society, telecommunication systems, storage facilities and data processing, new online services, are industries experiencing unprecedented growth. Data per se offers tremendous potential that we begin to use to generate new knowledge.

Personal data, both that produced by the users (texts, photos, videos,...) and that generated by the systems we use often unknowingly, is the heart of the economy of the information society, and therefore the heart of the economy. Control of the data also allows control of certain markets, which currently are already using U.S. electronic commerce tools in some areas. Control of the information society gives power still difficult to evaluate and far beyond the areas of the market economy.

Data capture is the top priority in some countries such as the United States or China (which hold respectively 72% and 16% of the top 50 sites worldwide). In both countries, national data remain under control of the domestic industry. And both aspire to collect the data at the international level.

5.2 Data Location

Geographical location of the cloud provider can have a real impact on the protection and confidentiality of data.

Legal obligations Sensitive data can be stored using a cloud computing solution. But for a French company, for example, it is necessary to check that the provider undertakes to keep these documents in France. Otherwise, the company may be unable to ensure that the processing of personal data complies with the legislation in force for it (e.g. in France: duration of data retention, ability to modify and delete information,...).

Similarly, it is generally necessary to comply with certain legal tax obligations: prohibition to store account books outside the European Union, mandatory reporting to the tax authorities in order to store electronic invoices outside the national territory,...

The "USA PATRIOT Act" Dutch legal researchers have published a study [16] that highlights the importance for a European company to choose a European provider to outsource the processing of personal data or information vital to the company. Indeed, since the establishment of the USA PATRIOT Act, U.S. law allows security services to access all personal data [17]:

- data from U.S. companies, even if the data is physically stored on the European territory
- data from their subsidiaries, even if they are located in another country in the world
- data stored on servers that are hosted in the United States, even if the company that owns the servers is of another nationality

The U.S. government has now established a legal arsenal which allows personal data control of foreign citizens, including Europeans, by leveraging its major companies such as Facebook, Google or Microsoft. At the end of 2012, the European Parliament's Committee on Civil Liberties, Justice and Home Affairs ("LIBE") released a study titled "Fighting cyber crime and protecting privacy in the cloud" [18]. Authors denounce the "Foreign Intelligence and Surveillance Act" (FISA). This amendment expressly authorizes U.S. intelligence agencies (NSA, CIA,...) to wiretap (without judicial authorization) U.S. citizens communicating with foreigners suspected of terrorism or spying. Shortly, a secret tribunal is now able to issue a warrant, secret too (the "secret" for actions may be required for an indefinite period), forcing American companies to deliver to U.S. intelligence agencies the private data of foreign users. Therefore your information may be duplicated, stored and disclosed to third parties without notifying you... In December 2012 the amendment was extended until 2017.

Our aim is not to pass judgment on the "USA PATRIOT Act" and other secret projects such as Riot or PRISM (since 2007 and revealed in june 2013 by Edward Snowden). We simply point out that in the current legislation, a European company with strong constraints on the information confidentiality

must therefore be vigilant when choosing a service provider (data location and nationality of the provider).

5.3 Towards a European CNIL

The European Parliament has made good progress on the reform of the EU legislation on data protection proposed almost a year ago by the Commission [19]. The United States, which are reforming their own legislation, call for transatlantic regulatory convergence, noting that they are just as demanding as the Europeans in this area. The EU has the ambition to become the global standard for data protection, suggesting (according to the authors) that the United States are more lax.

One of the most controversial issues is the requirement of equivalent standards to allow European data transfer to a third country for processing. The problem also arises for global corporations (e.g. Google), whose processing practices should be approved by the Union, while the United States would continue to use their codes of conduct.

In January 2013 the European Parliament presented its preliminary report on the future reform of the EU Directive on the protection of personal data in response to proposals from the European Commission. Wishing to strengthen the protection of data of its citizens, Europe is about to start revising the measures that came into force in 1995 ensuring wanting to replace Directive 95/46/EC on the protection of data by a European regulation that all Member States should apply without discussion. This reform will require the creation of an independent administrative authority, ie a European CNIL, which will enforce the rules on data protection, which could take the form of an independent agency.

However, for the French Data Protection Authority ("CNIL"), the text proposed by the European Justice Commissioner Viviane Reding "presents considerable progress" but also "elements of concern". The President of the CNIL, Isabelle Falque-Pierrotin acknowledges that it has the "major advantage" to submit to the European law all data processing on a European resident by a company not established in Europe: in other words it is the European law that would apply to a French victim of abuse by an American internet company, for example. But, says the President of the CNIL, the European text raises the problem of the concept of "principal place of business", according to which the competent regulatory authority in the event of a dispute with a European citizen is that of the place where the company and not the complainant.

5.4 Synthesis

The political and economic authorities have become aware of the need to establish an international legal framework to control the collection, storage and use of data. Metadata associated with data is also included. It is indeed value of the highest importance for companies whose business is the management of information.

6 Conclusion

Confidence in the data that we handle every day is one of the major challenges of the information society. There are many mechanisms that allow us to collect, store and process huge amounts of data and especially data on this data: metadata. Metadata is an essential tool for information security: usage control for document sharing and cloud security, digital forensics, evidence in case of litigation,...

Technological possibilities must not however make us forget the legal issues. The objective being to implement a security policy and to ensure information traceability, it is essential to respect existing regulations regarding the metadata that can be stored (see personal data, privacy), how it should be stored (see probative value) and computer processing in which it may be involved.

Through this article we want to raise awareness of potential abuses related to the use of such metadata. Some work has already been done to preserve privacy. An example is the anonymization of data [20,21]. These are not always suitable for our problem of usage control where precisely some indicators should not be anonymous. In the context of E-DRM we talk about business projects between partners. Thus our approach is rather to formalize this "collaboration agreement". For the IT community, this will be in terms of language as specification for metadata to be collected, by what means, how it is stored and what will be the use. For the legal community, it must first qualify the metadata: should it be treated as "traditional" data or should it receive a specific legal regime ? Once defined the legal framework, we can study together under what conditions it is possible to use metadata and, in the other way, what are the metadata necessary to apply certain laws. For instance, in the Oil & Gas case study described in Section 2.1, it will now be necessary to include in the contract between companies ("collaboration agreements") the insertion or the deletion of this metadata. For instance, should metadata appear within documents delivered at the end of the project ?

Finally, in Section 5 we discussed some socio-economic issues underlying the mass storage of data (and metadata) in the today information society. Beyond collaboration between partners on a project, we must also study the use of service providers on the "cloud" (storage or processes). These technologies have become unavoidable for companies although they introduce new vulnerabilities for the information security (loss of information controllability). These threats are not just technical (hardware, software, network). They can also be political, which requires the definition of an international legal framework for data protection.

References

1. Inmon, W.H.: Tech topic: What is a data warehouse ? Prism Solutions **1** (1995)
2. Kimball, R., Ross, M., Thornthwaite, W., Mundy, J., Becker, B.: The Data Warehouse Lifecycle Toolkit. 2nd edn. Wiley Publishing (2008)
3. Munier, M.: A multi-view approach for embedded information system security. In: CRiSIS, IEEE (2010) 65–72

4. Munier, M., Lalanne, V., Ricarde, M.: Self-protecting documents for cloud storage security. In: TrustCom, IEEE (2012) 1231–1238
5. Kalam, A.A.E., Benferhat, S., Miège, A., Baida, R.E., Cuppens, F., Saurel, C., Balbiani, P., Deswarte, Y., Trouessin, G.: Organization based access control. In: POLICY, IEEE Computer Society (2003) 120–131
6. Elrakaiby, Y., Cuppens, F., Cuppens-Boulahia, N.: From contextual permission to dynamic pre-obligation: An integrated approach. In: ARES, IEEE Computer Society (2010) 70–78
7. Cuppens, F., Cuppens-Boulahia, N.: Modeling contextual security policies. Int. J. Inf. Sec. **7**(4) (2008) 285–305
8. Cuppens, F., Miège, A.: Modelling contexts in the or-bac model. In: ACSAC, IEEE Computer Society (2003) 416–427
9. Bertino, E., Lim, H.S.: Assuring data trustworthiness - concepts and research challenges. In Jonker, W., Petkovic, M., eds.: Secure Data Management. Volume 6358 of Lecture Notes in Computer Science., Springer (2010) 1–12
10. Zheng, X., Maillé, P., Le, C.T.P., Morucci, S.: Improving the efficiency of collaborative work with trust management. In Agoulmine, N., Bartolini, C., Pfeifer, T., O'Sullivan, D., eds.: Integrated Network Management, IEEE (2011) 1172–1179
11. Xingyu Zheng and Patrick Maillé and Cam Tu Phan Le and Stephane Morucci: Trust mechanisms for efficiency improvement in collaborative working environments. In: MASCOTS, IEEE (2010) 465–467
12. Le, C.T.P., Cuppens, F., Cuppens-Boulahia, N., Maillé, P.: Evaluating the trustworthiness of contributors in a collaborative environment. In Bertino, E., Joshi, J.B.D., eds.: CollaborateCom. Volume 10 of Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering., Springer (2008) 451–460
13. Lalanne, V., Munier, M., Gabillon, A.: Information security risk management in a world of services. In: PASSAT. (2013)
14. Jaramillo, E., Munier, M., Aniorté, P.: Information security in business intelligence based on cloud: A survey of key issues and the premises of a proposal. In: WOSIS. (2013)
15. Hricik, D., Scott, C.E.: Metadata: The ghosts haunting e-documents. In: FindLaw. (March, 2008)
16. Van Hoboken, J., Arnbak, A., Van Eijk, N.: Cloud computing in higher education and research institutions and the USA PATRIOT Act. Social Science Research Network Working Paper Series (November 2012)
17. Lee, L.T.: USA PATRIOT Act and telecommunications: Privacy under attack. Rutgers Computer & Tech. LJ **29** (2003) 371
18. EU: Fighting cyber crime and protecting privacy in the cloud. EU Parliament (2012)
19. EU: Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protecting Regulation). Comm. European Communities, Bruxelles (2012)
20. Guarda, P., Zannone, N.: Towards the development of privacy-aware systems. Inf. Softw. Technol. **51**(2) (February 2009) 337–350
21. Zhou, B., Pei, J., Luk, W.: A brief survey on anonymization techniques for privacy preserving publishing of social network data. SIGKDD Explor. Newsl. **10**(2) (December 2008) 12–22

Chapitre 8

CNRIUT 2013
Securite de l'Information
Metadonnees & Aspects Juridiques

Sommaire

8.1 Présentation

Cet article a été présenté au 2ème Congrés National de la Recherche en IUT CNR IUT 2013 [33] (19ème édition) qui s'est tenue du 12 au 14 juin 2013 à l'IUT de Corté - Université de de Corse - France.

8.2 Le contexte de l'article

Cet article s'inscrit dans une politique de présentation des activités de recherche au sein des IUT. Nous montrons ainsi que dans des sites délocalisés il est possible de faire de la recherche, souvent basée sur des réseaux forts. Nous avons l'occasion d'afficher des axes de recherche trans-disciplinaires comme nous le faisons avec le CRAJ.

8.3 Les thèmes abordés

Nous présentons la notion de métadonnées afin de mettre en œuvre les mécanismes de contrôle d'usage dans un environnement de documents intelligents (E-DRM) ou dans la supervision d'architectures orientées services (SOA). Ces métadonnées nous permettent d'une part de définir des règles de sécurité contextuelles et d'autre part d'assurer la tracabilité des informations. Leur utilisation peut toutefois avoir des conséquences sur le plan juridique, notamment en ce qui concerne les métadonnées qu'il est possible d'enregistrer (cf. données personnelles), la manière dont elles doivent etre stockées (cf. valeur probante en cas de litige) ou les traitements informatiques dans lesquels elles peuvent etre impliquées. Un autre problème d'actualité est le stockage et le traitement des données via un prestataire de service : le cloud. Il faut neanmoins veiller à ce que cette solution ne conduise pas a une perte de maîtrisabilite de l'information pour l'entreprise. L'objectif de cet article est de présenter notre collaboration entre informaticiens et juristes sur ce sujet.

33. http ://cnr-iut.univ-corse.fr/Le-CNR-IUT_a21.html

8.4 Texte de l'article

Sécurité de l'Information
Métadonnées & Aspects Juridiques

Manuel Munier, Vincent Lalanne
LIUPPA (informatique)
IUT des Pays l'Adour, dpt R&T
Mont de Marsan, France
Email : manuel.munier@univ-pau.fr
vincent.lalanne@univ-pau.fr

Pierre-Yves Ardoy
CRAJ (droit)
Univ Pau & Pays Adour
Pau, France
Email : pierre-yves.ardoy@univ-pau.fr

Magali Ricarde
BackPlan Company
Project Communication Control
Pau, France
Email : magali.ricarde@backplan.fr

Résumé

Pour les besoins de nos travaux nous utilisons la notion de métadonnées pour mettre en œuvre des mécanismes de contrôle d'usage dans un environnement de documents intelligents (E-DRM) ou dans la supervision d'architectures orientées services (SOA). Ces métadonnées nous permettent d'une part de définir des règles de sécurité contextuelles et d'autre part d'assurer la traçabilité des informations. Leur utilisation peut toutefois avoir des conséquences sur le plan juridique, notamment en ce qui concerne les métadonnées qu'il est possible d'enregistrer (cf. données personnelles), la manière dont elles doivent être stockées (cf. valeur probante en cas de litige) ou les traitements informatiques dans lesquels elles peuvent être impliquées. Un autre problème d'actualité est le stockage et le traitement des données via un prestataire de service : le cloud. Il faut néanmoins veiller à ce que cette solution ne conduise pas à une perte de maîtrisabilité de l'information pour l'entreprise. L'objectif de cet article est de présenter notre collaboration entre informaticiens et juristes sur ce sujet.

Mots Clés

sécurité de l'information ; métadonnées ; droit ; vie privée ; cloud ;

⋆ ⋆ ⋆

– **Correspondant pour cette communication :** Manuel Munier

IUT des Pays de l'Adour - département R&T
371 rue du Ruisseau, BP 201
40004 Mont de Marsan Cedex
Email : manuel.munier@univ-pau.fr

– **Thème :** pluridisciplinarité
– **Type de session :** présentation

⋆ ⋆ ⋆

http://iut-adour.univ-pau.fr/ http://www.backplan.fr/

I. Problématique

Quels que soient les domaines d'activité, les nouvelles technologies de l'information (ADSL, ordinateurs portables, smartphones, tablettes,...) nous amènent à échanger et à stocker des informations en quantité de plus en plus importante. Leur contenu a également évolué. Les données sont de plus en plus complexes (documents dits structurés, archives représentant un projet complet avec sous-dossiers, documents de travail, références normatives,...). Des données dites publiques cohabitent maintenant avec des données plus confidentielles (notion de restriction d'accès). Les modalités d'échange ont elles aussi évolué. Nous emportons nos données sur des clés usb ou dans nos smartphones (dénoté en anglais par l'acronyme BYOD [1]), nous les partageons via des communications sans fil (3G, wifi, bluetooth) qui ne sont pas toujours forcément sécurisées, nous les hébergeons sur le cloud (*anywhere, anytime, any device*) sans avoir réellement conscience des risques induits. Dans cette société de l'information, la maîtrise des données est un enjeu majeur de la sécurité.

A. Les IUT Réseaux & Télécoms

De l'Internet aux réseaux sociaux, des téléphones aux jeux vidéos, des ordinateurs aux réseaux domotiques, des véhicules à la T.N.T, les Réseaux et Télécommunications sont présents chaque jour un peu plus dans notre quotidien. La spécialité R&T propose une formation scientifique qui aborde deux domaines en pleine convergence : les réseaux de télécommunications qui sont longtemps restés spécialisés dans le transport des informations audio/vidéo et les réseaux informatiques dédiés au transport de données. Aujourd'hui ils sont devenus indissociables : un même réseau unifié constitue l'infrastructure sur laquelle transitent toutes les données numériques autour de la planète (Internet, www, mail, téléphonie numérique, visioconférence, télévision numérique,...).

Les IUT R&T proposent également la licence professionnelle ASUR [2] dédiée à l'administration et à la sécurité des systèmes et des réseaux avec, entre autres compétences, effectuer une analyse des risques, concevoir un règlement de sécurité, installer les mesures de protection adéquates.

B. La Sécurité de l'Information

La sécurité informatique ne se limite toutefois pas à mettre en œuvre un intranet au sein de l'entreprise et à s'assurer du bon fonctionnement des serveurs et du réseau. En effet, nous disposons actuellement d'un arsenal technologique conséquent : serveurs de fichiers, annuaires centralisés, redondance de serveurs, VLANs, pare-feux, VPNs, architectures cloud et SOA [3],... Néanmois, il nous est de plus en plus souvent nécessaire de mettre en place des mécanismes de sécurité qui vont au-delà d'un traditionnel contrôle d'accès à une ressource : contrôle d'usage (comment un utilisateur peut agir sur un document : actions obligatoires, respect du flux de travaux, délégation de responsabilité,...), cohérence des informations (ex : certains documents peuvent faire référence à d'autres documents), traçabilité (supervision des actions réalisées, notion de preuve opposable en cas de litige, investigation scientifique),... Ces mécanismes collectent pour cela les métadonnées que l'on peut attacher aux information : qui exécute l'action, depuis quelle adresse IP, à quel instant, avec quelle application, depuis quel endroit (géolocalisation),... Ils peuvent également calculer certains indicateurs (nouvelles métadonnées) : confiance dans un partenaire, qualité d'un document, convergence ou divergence d'une tâche collaborative,...

Il ne s'agit là que de quelques exemples des possibilités offertes mais il apparaît clairement que l'utilisation de métadonnées n'est pas anodine d'un point de vue juridique ! Non seulement, du point du vue de la collecte et du stockage, certaines de ces métadonnées peuvent relever du domaine des données personnelles (géolocalisation notamment), mais leur utilisation pour activer/désactiver des règles de sécurité contextuelles ou calculer certains indicateurs sont des traitements automatisés et sont donc soumis à un certain nombre de dispositifs réglementaires. Ajoutons à cela les notions de responsabilité et, éventuellement, de sanction en cas de litige, et il est évident que les métadonnées deviennent dorénavant des éléments essentiels du système d'information.

Dans la société de l'information actuelle, les métadonnées deviennent ainsi parfois plus importantes que les données auxquelles elles sont associées. Que ce soit dans le domaine de la vie privée (données personnelles) ou professionnelle (données métier d'une entreprise), de nombreuses sociétés en ont fait leur business. Depuis quelques temps, les médias s'agitent autour de grandes multinationales telles que Google, Facebook ou Microsoft quant à la protection de la vie privée et des données personnelles. Dans le monde professionnel, il s'agit cette fois des données stratégiques de l'entreprise (ex : recherche & développement, stratégie commerciale). C'est le cas par exemple de la société BackPlan™ dont le business utilise les métadonnées associées aux informations échangées entre les partenaires pour assurer la gestion documentaire sur de gros projets. L'objectif de nos travaux ne consiste pas à critiquer les pratiques de tel ou tel pays ni à dénigrer les activités de telle ou telle entreprise. Nous voulons simplement mettre en avant les enjeux socio-économiques de l'information dans la société actuelle et la nécessité d'harmoniser les législations des différents pays afin de définir un cadre juridique international.

II. Nos Travaux de Recherche

Notre thématique de recherche sur la sécurité de l'information, initialement très "technico-informatique", nous a donc naturellement amené à collaborer avec nos collègues juristes pour traiter ces aspects juridiques.

A. Informatique

1) Documents autonomes sécurisés: Dans le cadre de nos travaux de recherche nous avons développé une architecture E-DRM [4] fondée sur des documents autonomes sécurisés. Alors que les systèmes d'information "traditionnels" centralisent toutes les données sur un serveur auquel les utilisateurs doivent se connecter, nous avons choisi de définir une approche où les mécanismes de sécurité sont délocalisés au plus proche de l'utilisateur. Toutefois, à l'inverse des architectures DRM "classiques" nécessitant l'utilisation d'un player dédié (lequel

1. BYOD : *Bring Your Own Device*
2. ASUR : Administration et SécUrité des Réseaux
3. SOA : *Service Oriented Architecture*
4. E-DRM : *Enterprise Digital Right Management*

est chargé de faire respecter la politique de sécurité), nous avons décidé d'intégrer ces mécanismes de sécurité au sein même du document en nous inspirant des concepts orientés objets : un document est une entité autonome capable d'assurer elle-même la sécurité des informations qu'elle contient et de contrôler la manière dont ces informations sont utilisées. Un tel document peut être considéré comme un système d'information en soi encapsulant à la fois un entrepôt de données et différents modules de sécurité (contrôle d'accès, contrôle d'usage, gestion des métadonnées,...). Les utilisateurs peuvent ainsi se l'échanger directement et en toute sécurité sans avoir à passer par un site central.

Le modèle de contrôle d'usage utilisé permet d'exprimer des politiques de sécurité dynamiques, c'est-à-dire dans lesquelles les règles (permissions, interdictions ou d'obligation) peuvent être "adaptées" en fonction du contexte d'exécution des actions : activation ou désactivation de certaines règles, insertion d'une obligation suite à l'exécution d'une action,... Pour cela, le système d'information embarqué doit remonter les informations nécessaires pour vérifier si les conditions associées aux définitions des contextes sont satisfaites. Pour cela, il peut soit accéder au système hôte (ex : obtention d'une horloge globale pour vérifier les contextes temporels), soit utiliser les métadonnées attachées aux actions et aux nœuds de la base de données embarquée.

Notre approche repose donc fortement sur la notion de métadonnées, tant au niveau de la collecte (traçabilité) que pour l'activation des contextes (règles de sécurité dynamiques) ou, à terme, le calcul de différents indicateurs au fil des utilisations (indice de confiance).

2) Sécurité des architectures orientées services: Les SOA offrent de nouvelles possibilités pour l'interconnexion des systèmes d'information (SI). L'ouverture du SI d'une entreprise sur l'extérieur a toutefois des conséquences sur la sécurité. Que ce soit pour utiliser des services proposés par des tiers ou pour offrir les siens, ces technologies introduisent de nouvelles vulnérabilités dans le SI et, par conséquent, de nouveaux risques. Nos travaux visent à initier une démarche de gestion de ces risques qui s'appuie sur un standard, la norme ISO/IEC 27005. Nous proposons une évolution de cette norme afin de prendre en compte pleinement cette notion de service.

Afin d'élaborer un modèle de sécurité pour les communications inter-SI nous envisageons une approche orientée contrôle d'usage telle que présentée précédemment. L'utilisation de métadonnées pour la traçabilité des communications (via ces services) nous permettra également de remonter des indicateurs qui pourront, par exemple, être utilisés pour superviser le SI. Notre objectif est que les entreprises puissent garder la maîtrise de leurs informations malgré l'utilisation du cloud.

B. Droit

Les données personnelles, tant celles produites par les usagers (textes, photos, vidéos,...) que celles générées par les systèmes que nous utilisons souvent à notre insu, sont au cœur de l'économie de la société de l'information, et donc de l'économie. L'émergence du cloud n'a fait qu'accentuer le phénomène : que ce soit à titre privé ou à titre professionnel (voire les deux), nous sommes de plus en plus nombreux à confier nos données à de grandes multinationales telles que Google, Facebook, Amazon ou Microsoft qui ont fait de l'information leur cœur de métier. La maîtrise de la donnée, et donc la maîtrise de la société de l'information, donne une puissance qu'on soupçonne encore peu et qui dépasse de loin les secteurs de l'économie marchande. Les tensions actuelles entre l'Union Européenne et les États Unis quant à la sécurité de l'information (dont les données personnelles) sont la preuve des enjeux stratégiques sous-jacents.

A ce titre les métadonnées jouent un rôle particulier. Or non seulement le rôle des métadonnées est très largement ignoré du public, mais plus encore, l'existence des métadonnées elles-mêmes est souvent inconnue du public. Ces données sur les données sont même la plupart du temps collectées à l'insu des utilisateurs d'ordinateurs, de smartphones,... et donc, par hypothèse, sans leur consentement. Les métadonnées posent donc, avant tout, un problème de sécurité. De ce point de vue, la collaboration entre informaticiens et juristes semble être une voie importante pour tenter de cerner ce que sont les métadonnées tant du point de vue technique que du point de vue juridique, identifier les difficultés qu'elles posent dans les domaines respectifs de l'informatique et du droit, et proposer des solutions fondées technologiquement et juridiquement.

En premier lieu, il est indéniable que les métadonnées sont des objets de droit. S'agissant par exemple de la confidentialité (entre autres critères de sécurité), les modèles de contrôle d'usage ont introduit la notion de contexte afin de pouvoir exprimer des règles de sécurité dynamiques dans une politique (temporel, géolocalisation, prérequis,...). Pour activer ou désactiver ces contextes, le système d'information doit collecter et stocker diverses métadonnées : date, adresse IP utilisée, localisation de l'utilisateur,... Certaines permissions, par exemple, pourront n'être octroyées que si l'utilisateur se trouve à l'intérieur d'une enceinte sécurisée, à certaines plages horaires ou en fonction d'actions préalablement réalisées. Cette "contextualisation" des données ne peut cependant pas se faire du strict point de vue informatique et le droit doit permettre de déterminer les données susceptibles d'être collectées dans telle ou telle situation, ou les contraintes que le système peut imposer à un utilisateur (contraintes qui, le cas échéant, peuvent avoir comme conséquence d'empêcher l'utilisateur de réaliser certaines tâches).

En second lieu, les métadonnées peuvent efficacement servir le droit dès lors qu'il est question de sécuriser un contenu qui a vocation à influencer une prise de décision dans le cadre d'un procès par exemple. Ainsi, les métadonnées peuvent être utilisées à des fins de traçabilité et notamment servir de preuve en cas de litige. Elles peuvent en effet aider à déterminer des responsabilités, à déterminer l'étendu d'un droit par exemple en matière de bornage), ou d'une manière plus générale elles sont susceptibles d'être admises comme éléments de preuve.

III. Synthèse

Les données constituent les briques de base de la société de l'information. Elles offrent un potentiel extraordinaire que l'on commence à exploiter. Elles permettent de générer des connaissances, qui étaient soit hors d'atteinte, soit inexistantes, parce que hors du domaine du pensable. Cette évolution sociétale ne doit toutefois pas se faire au détriment des utilisateurs et nécessite donc d'étudier et/ou adapter le cadre juridique.

Chapitre 9

Conclusions et Perspectives

Conclusions

L'objet de cette thèse est la gestion des risques appliquée aux systèmes d'information distribués.

Nous nous sommes tout d'abord intéressés à la sécurisation des architectures orientées services, en particulier les systèmes de gestion de droits de contenus numériques ou DRM. Nous avons traité les problèmes d'interopérabilité au sein de ces structures afin de rendre ce service plus efficace et mieux sécurisé. Nous avons exposé l'intérêt de l'utilisation de langage d'expression de droits pour l'édition des licences. Nous avons mis en avant les standards Web services Security pour l'échange de messages entre les acteurs de l'architecture DRM et nous avons proposé pour chaque composant une interface avec un service Web, les messages se faisant par le protocole SOAP.

Notre deuxième contribution, sous forme d'article, a concerné la distribution du contenu sécurisé ou auto protégé au sein des structures de type cloud. Dans cet article nous sommes partis du postulat que la sécurité des informations (ou des données) dans une architecture publique du type cloud n'est pas garantie. C'est-à-dire que nous avons pris comme point de départ une crainte de pertes de confidentialité ou d'intégrité. Nous avons développé ensuite notre modèle de DRM d'entreprise : il s'agit d'une architecture orientée documents intelligents à destination des entreprises. Un document intelligent est un document qui contient son propre système d'information, nous avons développé ce concept à travers un modèle de DRM. Nous avons abordé également avec ce concept la possibilité des documents collaboratifs.

À travers ces années de thèse j'ai eu l'occasion de participer à une aventure qui est celle de la création d'une entreprise innovante, la société BackPlan™ . Cette société a pour objet de développer un service dédié aux entreprises qui se retrouvent sur de grands projets industriels. Le principe d'un grand projet est de mettre ensemble pour une durée finie un grand nombre de sociétés qui n'ont jamais collaboré. La transmission de l'information est primordiale dans ce genre de situation : c'est sur ce créneau que se situe la société BackPlan™ . Elle définit ce qu'elle appelle une interface et se propose de gérer ces interfaces entre les sociétés. Le concept est innovant car il n'existe pas, bien souvent tout repose sur le chef de projet qui a du mal à coordonner la circulation des bonnes informations. Associé à cette méthode la société a eu besoin d'un outil informatique de diffusion et partage de l'information, c'est sur ce point-là que nous sommes intervenuss. Un accent particulièrement fort a été mis sur la gestion de la sécurité de l'information, c'est à cette occasion que nous avons passé la certification ISO/IEC 27005:2011, standard de gestion des risques en sécurité de l'information. Nous avons développé un prototype d'architecture basée sur les services qui puissent subvenir aux besoins de BackPlan™ .

Avec cette société innovante nous avons l'occasion de nous appuyer sur des cas concrets dans le domaine qui est le leur, l'Oil & Gas. Nous avons également bâti un partenariat durable

sous forme d'un contrat de recherche qui nous permettra de développer des axes communs en particulier dans le domaine des métadonnées. En termes de résultats que pouvons nous féliciter du classement officiel de cette entreprise en JEU [34].

Cette certification ISO/IEC 27005:2011 nous a permis de développer une partie service dans une norme dédiée aux systèmes d'information. En effet cette norme de gestion des risques ne tient pas compte de la notion même de service tel qu'on la comprend actuellement : cloud, stockage en ligne, traitement des données etc. Nous avons donc proposé une extension de cette norme pour tenir compte de ce monde de service. Nous avons également introduit un critère de maitrisabilité des services : en d'autres termes l'architecte d'un système d'information doit pouvoir maîtriser les services qu'il contractualise avec ses prestataires.

Dans les deux derniers articles nous avons développé notre concept de DRM d'Entreprise qui utilise les métadonnées pour deployer des contextes dans les modèles de contrôle d'usage. Nous avons également développé la nécessité de formaliser les métadonnées nécessaires à la mise en vigueur de la politique de sécurité tout en garantissant le respect de la réglementation : cette partie a été réalisée en partenariat avec les juristes de notre université. Les exemples liés au contexte ont été tirés études de cas fournis par la société BackPlan™ .

Quand on fait le point sur les thématiques abordées et les contributions apportées, nous avons eu comme fil rouge l'utilisation des services et leur sécurisation : c'est une constante. Mais si l'on regarde du côté de la sécurité, nous avons évolué d'une sécurité de type informatique, tout à fait classique, vers une sécurité des données et de l'information. Cette évolution n'est pas anodine car elle correspond à une abstraction de la notion de sécurité. Nous passons de la micro sécurité, c'est-à-dire la sécurité technique, celle des outils, à la macro sécurité, c'est-à-dire celle des systèmes et des organisations. Les dernières révélations de captation et de stockage des données personnelles par la NSA [35], en juin 2013, ne penvent que nous convaincre de cette conséquence : quelque soit le verrou placé sur la porte d'une pièce, si le mur de l'enceinte est percé, il ne servira à rien !

Perspectives

Nos travaux de thèse ouvrent de nouvelles pistes de recherche que nous tenons à explorer dans ce qui suit.

En premier il faut se réjouir de la bonne santé de la société BackPlan™ qui maintenant après trois années d'existence se « porte bien ». Nous avons des perspectives de contractualiser à nouveau sur deux axes de recherche. Le premier axe concerne l'utilisation des métadonnées tant au niveau de l'infrastructure que du côté du pilotage des échanges (tableaux de bord, business

34. Jeune Entreprise Universitaire
35. National Security Agency

intelligence, Big Data), le second s'articule autour de notre concept de documents intelligents dans un monde nomade : les chantiers de construction se déroulent en extérieur et rarement dans les bureaux ! Ceci implique que les utilisateurs doivent pouvoir travailler tout en étant déconnectés du réseau de leur entreprise.

Ensuite, nos collaborations avec le Centre de Ressource et d'Analyse Juridiques commence à porter ses fruits : deux publications en commun ont été acceptées et un projet PEPS CNRS a été soumis. L'objectif est de poursuivre sur la thématique « métadonnées & sécurité de l'information »avec également deux axes qui sont les métadonnées pour le contrôle d'usage et l'investigation numérique (digital forensics) qui menent dans le domaine du droit aux notions de preuve probante et de pré-constitution de preuves.

Au sein de notre équipe de recherche, les voies qui ont été ouvertes autour de la gestion des risques et de l'exploitation des métadonnées ne se referment pas. Nous continuons l'intégration de la notion de service dans les normes ISO 27k dédiées à la sécurité de l'information. La thèse d'Elena Jaramillo démarrée en décembre 2012 a pour objectif de prendre en compte, dès la phase d'analyse et de conception, les aspects « sécurité de l'information »dans le domaine des architectures orientées services (ex : Web Services, technologies cloud,...). Outre la modélisation des politiques de sécurité et la formalisation des interactions entre services, elle adopte une démarche IDM afin de pouvoir intégrer ces travaux d'un point de vue Génie Logiciel. Ces travaux s'inscrivent dans une problématique plus générale communément appelée « privacy by design ». Cette thèse est financée par le Conseil Général des Landes. Ces travaux on déjà permis de publier deux articles en un an : WOSIS 2013 et SARSSI 2013.

D'autre part notre équipe a vu l'arrivée d'un nouveau Maitre de Conférences qui a été recruté sur ce profil.

Nous pouvons constater que les axes qui ont été les nôtres pendant ce travail de recherche, ont permis d'ouvrir de nouveaux champs d'investigation qui ne vont pas manquer d'être parcourus.

Bibliographie

[AD01] Christopher Alberts and Audrey Dorofee. An introduction to the octave method. *Pittsburgh, PA : Software Engineering Institute, Carnegie Mellon University. http ://www.cert.org/octave/methodintro.html*, 2001.

[BBF+08] Mark Bartel, John Boyer, Barb Fox, Brian LaMacchia, and Ed Simon. XML Signature Syntax and Processing. Technical report, W3C Open Source Software, 2008.

[BS795] Bs 7799 :part 1 :1995 information security management code of practice for information security management systems. Technical report, BSI British Standards, 1995.

[CKPM05] Cantor, Kemp, Philpott, and Maler. Security Assertion Markup Language (SAML). Technical report, OASIS, 2005.

[CML09] James Caverlee, Prasenjit Mitra, and Mary Laarsgard. Dublin core. In Ling Liu and M. Tamer Özsu, editors, *Encyclopedia of Database Systems*, page 949. Springer US, 2009.

[CRA] Uk government's risk assessment methodology : Cramm.

[CVE] Cve : Common vulnerabilities and exposures. Technical report.

[Dem50] JW. Edwards Deming. Elementary principles of the statistical control of quality. *Union of Japanese Scientists and Engineers*, 7(1), 1950.

[EBI10] Ebios 2010 expression des besoins et identification des objectifs de sécurité. Technical report, 2010.

[ER02] Donald Eastlake and Joseph Reagle. XML Encryption Syntax and Processing. Technical report, W3C, 2002.

[FT00] Roy T Fielding and Richard N Taylor. Principled design of the modern web architecture. In *Proceedings of the 22nd international conference on Software engineering*, pages 407–416. ACM, 2000.

[GPC05] Christian Geuer-Pollmann and Joris Claessens. Web services and web service security standards. *Information Security Technical Report*, 10(1) :15 – 24, 2005.

[ISO05a] ISO. ISO/IEC 17799 :2005 - Code of practice for information security management. Technical report, ISO, 2005.

[ISO05b] ISO/IEC. ISO/IEC 27001:2005 : Information security management systems. Published, ISO, 2005.

[ISO11] ISO/IEC. ISO/IEC 27005:2011 : Information technology — security techniques — information security risk management. Published, International Organization for Standardization, Geneva, Switzerland, 2011.

[Lee03] Laurie Thomas Lee. Usa patriot act and telecommunications : Privacy under attack, the. *Rutgers Computer & Tech. LJ*, 29 :371, 2003.

[LPL+03] Markus Lorch, Seth Proctor, Rebekah Lepro, Dennis Kafura, and Sumit Shah. First experiences using xacml for access control in distributed systems. In *Proceedings of the 2003 ACM workshop on XML security*, pages 25–37. ACM, 2003.

[MG03] Tim Moses and Simon Godik. eXtensible Access Control Markup Language (XACML) Version 1.0. Technical report, OASIS, 2003.

[OAS06] OASIS. Web Services Security (WSS). Technical report, OASIS, 2006.

[OCT] Octave® information security risk evaluation. Technical report.

[OWA] Owasp : The open web application security community. Technical report.

[PSS00] Jaehong Park, Ravi Sandhu, and James Schifalacqua. Security architectures for controlled digital information dissemination. In *Computer Security Applications, 2000. ACSAC'00. 16th Annual Conference*, pages 224–233. IEEE, 2000.

[Use99] UserLand. XML-RPC. Technical report, UserLand Software, Inc., 1999.