Métadonnées et Aspects Juridiques: Vie Privée vs Sécurité de l'Information

Manuel Munier, Vincent Lalanne, Pierre-Yves Ardoy, Magali Ricarde

Métadonnées & Aspects Juridiques Vie Privée vs Sécurité de l'Information

Manuel Munier (manuel.munier@univ-pau.fr)*
Vincent Lalanne (vincent.lalanne@univ-pau.fr)*
Pierre-Yves Ardoy (pierre-yves.ardoy@univ-pau.fr)†
Magali Ricarde (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 et de gestion des risques du point de vue de la sécurité de l'information. 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 positionner nos travaux par rapport à ces aspects juridiques.

Mots Clés : métadonnées, vie privée, sécurité de l'information, enjeux socio-économiques

1 Introduction

Quels que soient les domaines d'activité, les nouvelles technologies de l'information (ADSL, ordinateurs portables, smartphones, tablettes...) 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 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). Sans compter que nous emportons nos données sur des clés usb ou dans nos smartphones et les partageons via des communications sans fil (3G, wifi, bluetooth). Dans cette société de l'information, la fiabilité des données manipulées est devenue un enjeu majeur du point de vue de la sécurité.

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 : contextes temporels, (géo)localisation, pré-conditions... Pour activer ou désactiver ces contextes, le système d'information doit collecter et stocker diverses métadonnées : date,

*. LIUPPA, Université de Pau et des Pays de l'Adour, Mont-de-Marsan, France
†. CRAJ, Université de Pau et des Pays de l'Adour, Pau, France
‡. BackPlan, Project Communication Control, Pau, France

adresse IP utilisée, localisation de l'utilisateur, applications... 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. Les métadonnées sont à l'origine du succès des thématiques *data warehousing* et *business intelligence* depuis le milieu des années 90. Les entrepôts de données ont pour objectif de gérer et stocker des masses de données tandis que l'intelligence d'entreprise se focalise sur l'utilisation des données pour l'analyse et l'aide à la décision [Inm95, KRT+08]. Le terme de métadonnée signifie "donnée sur des données". Cependant, cette notion n'est pas (encore) bien connue du droit.

Notre article est organisé de la manière suivante. La section 2 présente les motivations pour l'utilisation de métadonnées, tant du point de vue de la sécurité de l'information que du cœur de métier de certaines sociétés de services. Nous illustrons ces besoins par un exemple de gestion des documents d'entreprise dans le cadre d'un projet Oil & Gas développé avec notre partenaire BackPlan™ [1]. Dans la section 3 nous présentons deux domaines de nos activités de recherche liés à la sécurité de l'information et utilisant le concept de métadonnées. Si les métadonnées sont utiles pour le contrôle d'usage, la traçabilité ou tout simplement la gestion du système d'information, il faut néanmoins se préoccuper des aspects juridiques. La section 4, via des exemples de jurisprudence, met en avant la nécessité de formaliser les métadonnées nécessaires pour mettre en œuvre la politique de sécurité et le cadre dans lequel elles peuvent être utilisées afin de respecter la réglementation en vigueur. La section 5 présente quelques problèmes socio-économiques sous-jacents liés au stockage de données (et métadonnées) dans la société de l'information qui est la nôtre aujourd'hui. Finalement, la section 6 conclut notre article et présente nos perspectives actuelles de travail.

2 Motivations

Comme nous l'avons indiqué au début de notre article, les technologies de l'information nous permettent d'échanger de plus en plus d'informations sous la forme de documents dont la structure devient de plus en plus complexe. Que ce soit dans le cadre d'une "simple" diffusion d'informations (communication unidirectionnelle d'un fournisseur de contenu vers des utilisateurs) ou d'un environnement de travail collaboratif (plusieurs acteurs interagissent pour réaliser des tâches avec un but commun mais éventuellement des objectifs différents), il est donc nécessaire de mettre en œuvre des mécanismes de sécurité qui vont au-delà d'un simple contrôle d'accès : contrôle d'usage (de quelle manière les partenaires peuvent utiliser un document : obligations, workflows, des règles de délégation...), gestion de la cohérence des informations (ex : des documents peuvent en référencer d'autres), traçabilité...

Dans [MLAR13] nous avons présenté un cas d'étude dans le domaine du génie pétrolier avec un exemple de construction d'un pipeline ou d'une installation pétrolière. Le système d'information d'un tel projet est constitué de nombreux documents. Il a un rôle central, structurant et qui évolue avec l'avancement du projet : la documentation doit toujours précéder l'action (conception, procédure de travail) ; la documentation doit toujours clôturer la construction et ses contrôles (enregistrements, procès-verbaux) ; le document vit

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

au même rythme que le projet ; ces documents sont des spécifications, des schémas, des dossiers d'expertise, des procédures, des enregistrements...

Nous ne donnons pas ici tous les détails de notre étude de cas. Cette section présente l'utilisation des métadonnées pour améliorer la sécurité de l'information (traçabilité, contrôle d'usage) et les questions juridiques qui peuvent se poser.

Aspects Sécurité de l'Information Pour des phases de contrôle telles que la vérification des soudures d'un pipeline ou l'épreuve hydraulique d'un tronçon, il serait intéressant de capturer des (méta)données pour valider des indicateurs d'avancement, améliorer la traçabilité (historique) du processus à des fins de validation et/ou de preuve vis-à-vis de l'administration : photos géotaggées (pour certifier des points de contrôle), métadonnées associées aux plans...

L'utilisation de métadonnées permettrait également de "lier" les revues, certifications, et autres procès verbaux aux documents de conception au sein du SI. Étant donné que plusieurs partenaires collaborent sur un tel projet, chacun pourrait attacher certaines métadonnées aux informations qu'il reçoit en entrée (indice de confiance, criticité d'une modification en termes d'impact sur sa propre tâche...). Sur la base de ces métadonnées il pourrait alors calculer différents indicateurs pour le suivi de sa tâche et ainsi que de nouvelles métadonnées pour les informations qu'il produit. Là encore, l'objectif est d'améliorer la traçabilité, tant au niveau du processus de conception (notion de workflow, de tableaux de bord) qu'au niveau de la gestion des responsabilités en cas de litige (notion de preuve opposable, d'investigation scientifique).

Un autre aspect de la sécurité de l'information consiste à se servir de telles métadonnées pour mettre en place des mécanismes de contrôle d'usage. Il s'agit là de contrôler non seulement les accès aux informations mais également la manière dont un partenaire va pouvoir utiliser les informations auxquelles il aura eu accès. Voici quelques exemples de règles de sécurité que l'on voudrait mettre en place :

- "Il n'est possible de rédiger un livrable du projet que si la confiance dans les différents documents techniques dépasse un certain seuil" (contrôle d'accès évolutif fondé sur des indicateurs de confiance).
- "Accès restreint à certaines parties du document en fonction de données de géolocalisation" (pour éviter, par exemple, que sur un chantier un inconnu ne prenne connaissance d'informations sensibles à l'insu d'un partenaire).
- "Chaque partenaire doit avoir relu chaque étude conceptuelle à laquelle il participe au moins 7 jours avant la date limite de validation" (obligation collective).

Aspects Juridiques Il ne s'agit là que de quelques exemples des possibilités offertes mais il est évident que l'utilisation de métadonnées n'est donc pas anodin 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 (permissions, obligations...) ou calculer certains indicateurs (ex : confiance en un partenaire, qualité d'un document) sont des traitements automatisés et sont donc soumis à un certain nombre de dispositifs réglementaires. Ajoutons à cela les notions de responsabilité et de sanction évoquées ci-dessus et il est évident que les métadonnées deviennent dorénavant des éléments essentiels du système d'information. Elles doivent être considérées comme des données à part entière et être sécurisées au même titre que les informations

"classiques".

La section 4 reviendra plus en détails sur les besoins de formalisation des métadonnées nécessaires à la politique de sécurité et le cadre dans lequel elles peuvent être utilisées. Ceci afin qu'elles puissent être utilisées en tant que preuve en cas de litige (cf. valeur probante) tout en respectant la réglementation concernant les données personnelles.

BackPlan^{tm} est une société française fournissant des services de gestion électronique de documents et de planification dans le cadre de projets industriels d'ampleur, ceci pour améliorer la communication et la transparence au sein du projet, pour gérer les risques liés à la sécurité de l'information et pour garantir la fiabilité des indicateurs ainsi que la conformité réglementaire.

La collecte de métadonnées (des données sur les données) au fur et à mesure que les partenaires consultent ou mettent à jour le registre de documents [2] permettrait ainsi d'automatiser certains traitements : calcul d'indicateurs et de tableaux de bord, respect des échéances et anticipation des retards, état d'achèvement des documents... En termes de gestion des risques, ces informations peuvent également être utilisées à posteriori, en cas de litige, pour identifier les responsables d'une erreur ou, au contraire, pour prouver que les normes en vigueur au moment de la phase de construction ont bien été respectées.

3 Métadonnées & Sécurité de l'Information

Nos activités de recherche se déroulent dans le domaine de la sécurité des systèmes d'information. Les questions relatives au cadre juridique pour l'utilisation des métadonnées ont été soulevées lorsque nous avons voulu mettre en œuvre les règles de sécurité contextuelles pour effectuer du contrôle d'usage dans deux projets : le développement d'une architecture de documents autonomes sécurisés, et l'étude de la sécurité dans les architectures orientées services. Ces métadonnées sont bien évidemment utilisées pendant le cycle de vie du document pour le contrôle d'usage. Mais elles peuvent aussi être consultés à posteriori à des fins de traçabilité (des actions réalisées) et peuvent donc également servir de preuve en cas de litige ou d'investigation judiciaire.

Cette section se concentre sur l'utilisation des métadonnées pour améliorer la sécurité de l'information. Dans notre travail, ces "données sur les données" nous permettent de mettre en œuvre des mécanismes de sécurité tels que les politiques de sécurité dynamiques, la gestion du travail collaboratif, le calcul d'indicateurs de confiance et de fiabilité... L'utilisation des métadonnées n'est cependant pas anodin au regard de la loi, et les possibilités technologiques ne doivent pas nous faire oublier les aspects juridiques (cf. Section 4). En outre, ces métadonnées sont même parfois plus importantes que les données auxquelles elles sont associées (cf. Section 5).

3.1 Contrôle d'Usage

Dans le cadre de nos travaux de recherche nous avons développé une architecture E-DRM [3] fondée sur des documents autonomes sécurisés [MLR12]. Nous avons pour cela 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

2. le registre de documents est hébergé chez BackPlan™
3. E-DRM : *Enterprise Digital Right Management*

elle-même la sécurité des informations qu'elle contient et de contrôler la manière dont ces informations sont utilisées. Le modèle de contrôle d'usage utilisé (OrBAC [KBM+03]) 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 [CCB08] : activation ou désactivation de certaines règles, insertion d'une obligation suite à l'exécution d'une action... Le système d'information embarqué doit donc remonter les informations nécessaires pour vérifier si les conditions associées aux définitions des contextes sont satisfaites.

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 (ex : indice de confiance, probabilité de non respect des délais).

3.2 Gestion des Risques

Les architectures orientées services (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" telle que les services web ou les services cloud [LMG13].

Afin d'élaborer un modèle de sécurité pour les communications inter-SI nous utilisons de nouveau une approche orientée contrôle d'usage telle que présentée précédemment. Là aussi, l'utilisation de métadonnées pour la traçabilité des communications (via ces services) nous permettra de remonter des indicateurs qui seront utilisés pour superviser le SI [JMA13]. Notre objectif est que les entreprises puissent garder la maîtrise de leurs informations malgré l'utilisation de technologies cloud.

3.3 Confidentialité des Métadonnées

Des tests préliminaires sur nos documents auto-protégés, effectués avec des métadonnées prédéfinies, nous ont permis de mettre en œuvre des règles de sécurité dynamiques pour le contrôle d'usage. Ces expérimentations nous ont également conduit à nous préoccuper de la confidentialité des métadonnées.

Selon la norme ISO 8402, la traçabilité concerne l'aptitude à retracer l'historique, l'utilisation ou la localisation d'une entité au moyen d'identifications enregistrées. Dans le cadre de notre architecture, une entité correspond à un document de travail collaboratif autonome. L'historique, l'identification, l'enregistrement et l'utilisation sont des notions pertinentes. Comme nous l'avons indiqué précédemment, l'aspect localisation de l'utilisateur par exemple peut constituer une métadonnée intéressante pour l'expression de règles de sécurité contextuelles. Mais un document pourrait alors révéler la présence de l'utilisateur à un instant donné, ou les différentes révisions qui se sont succédées pour aboutir à la version finale d'un document contractuel ! Il nous faut donc non seulement nous intéresser aux métadonnées qui sont collectées, mais également les protéger de toute utilisation abusive et contrôler les utilisations qui en sont faites (ex : calcul automatique d'indicateurs). Fondamentalement, ce problème de fuite et/ou de détournement d'information est

déjà connu. Toutefois, dans notre approche, la collecte "massive" de métadonnées "de tout type" peut effectivement accentuer le problème.

4 Métadonnées & Aspects Juridiques

Comme nous l'avons expliqué depuis le début de cet article, 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.

La notion de métadonnée n'est pas une notion bien connue du droit. Composé du préfixe grec *méta-* renvoyant à la référence à soi-même, le terme de métadonnée fait référence à des données sur des données, des données permettant de définir, de circonscrire ou de décrire une autre donnée. Ces données sont variées et peuvent intégrer des durées, des dates, des lieux, des éléments d'identification des personnes... Dans un domaine bien précis, le droit donne toutefois une définition des métadonnées. L'article L.127-1, 6° du Code de l'environnement précise en effet qu'il s'agit des *« information(s) décrivant les séries et services de données géographiques et rendant possible leur recherche, leur inventaire et leur utilisation »*. Cette définition, la première posée par un texte légal, ne rend toutefois pas compte de la diversité qui se cache sous le terme de métadonnée.

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 découverts et utilisés.

En second lieu les métadonnées posent la question de leur stockage. Ce stockage suppose une conservation garantissant non seulement l'authenticité des métadonnées (ce qui renvoie à la question de la collecte et de la fiabilité de la source) mais également l'intégrité, la stabilité de celles-ci. Ce qui est au cœur de la question est en fin de compte l'exigence de fiabilité des métadonnées. Il ne suffit pas que celles-ci soient conservées dans de bonnes conditions, il faut encore que soit assurée leur disponibilité, qu'elles soient accessibles, la question étant de savoir par qui pour combien de temps et dans quelles conditions.

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.

Potentiellement, les métadonnées concernent tous les domaines du droit et il paraît difficile voire inutile de les appréhender dans l'abstrait. Il s'agit ici d'identifier des domaines où l'utilisation de métadonnées est susceptible de donner lieu à des questionnements particuliers sans toutefois prétendre à l'exhaustivité. Précisons également qu'il s'agit ici d'un premier questionnement limité au droit interne.

4.1 Droit de la Preuve

C'est essentiellement sur le terrain du droit de la preuve que se placent les rares décisions de cours d'appel se référant aux métadonnées.

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.

C'est le cas tout d'abord des **Ordonnances du Premier Président de la Cour d'Appel de Paris du 25 octobre 2011, Juris-Data n°2011-025553 et du 25 octobre 2011 (inédit, n°09/14462, 09/14501)** [4]. Dans ces espèces, une société avait engagé un recours contre des opérations de visite et de saisie autorisées dans ses locaux par l'Autorité de la concurrence dans le cadre d'une enquête en matière de pratiques anticoncurrentielles. Si le problème juridique ne portait pas directement sur la question des métadonnées, la Cour entérine cependant l'argumentaire de l'Autorité de la concurrence au terme duquel *« la structure particulière d'un fichier de messagerie Outlook et l'obligation de ne modifier ni l'état de l'ordinateur visité, ni les attributs des fichiers (métadonnées contenues dans le fichier lui-même : titre, auteur taille, dates, localisation, signature...) impliquent nécessairement la saisie globale du fichier de messagerie, après avoir vérifié qu'il contient des éléments entrant dans le champ de l'autorisation. »* Dans tous les cas l'accent est mis sur l'exigence selon laquelle il convient de saisir l'intégralité de messageries Outlook afin de ne pas altérer la fiabilité de la saisie par une altération des métadonnées contenues dans les messages.

Dans un arrêt de la **Cour d'appel de Versailles, Ch.1, sect. 1 du 30 Septembre 2010 (inédit, n°09/03831)**, un requérant s'appuie sur les métadonnées contenues dans des photographies pour démontrer sa qualité d'auteur des photographies. La Cour retient les métadonnées comme éléments de preuve en relevant que celles-ci comportent l'identité de l'auteur, la date et l'heure des prises de vues, le nom du fabriquant de l'appareil, le modèle de celui-ci, ainsi que le détail des réglages de l'appareil. Si les métadonnées ne peuvent en soi démontrer l'originalité des photographies, elles sont toutefois des éléments permettant d'en établir la paternité et peuvent s'avérer fort utiles dans le cadre d'un procès en contrefaçon comme en l'espèce. De manière proche, la **Chambre prud'homale 7 de la Cour d'appel de Rennes** retient, dans un arrêt du 20 septembre 2011 (inédit, RG n°10/05183), que la production de métadonnées peut servir à déterminer la date de création d'une plaquette publicitaire tout en relevant que la fiabilité de celles-ci n'est pas discutée.

On peut bien évidemment imaginer que les métadonnées peuvent servir d'élément de

4. Voir également les **Ordonnances du 15 novembre 2011 ou du 31 août 2012** du Premier président de la CA Paris rendues également en matière de pratiques anticoncurrentielles.

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.

Au-delà des relations avec les salariés, les métadonnées peuvent intéresser les rapports entre les associés ou les dirigeants d'une société. Elles pourraient par exemple servir à démontrer que tel dirigeant ou tel salarié s'est rendu responsable d'une divulgation de secret de fabrique (article L.621-1 du Code de la propriété intellectuelle et L.1227-1 du Code du travail).

Bien évidemment il ne s'agit là que de quelques utilisations possibles des métadonnées, une liste exhaustive s'avérant sinon impossible, du moins fort délicate. [Dro13] présente une étude plus détaillée des métadonnées en droit de la preuve.

4.2 Vie Privée et Libertés Individuelles

Comme nous pouvons le constater, les métadonnées sont susceptibles de contenir des informations relatives aux personnes, qu'il s'agisse de leur identification ou de leur localisation. Si le droit commence à peine à connaître la notion de métadonnées, il connaît bien en revanche celle de données à caractère personnel ou données personnelles. Il s'agit d'informations qui permettent d'identifier directement ou indirectement des personnes physiques (*Loi informatique et Libertés n° 78-17 du 6 janvier 1978 ; directive 95/46/CE du 23 octobre 1995*). La notion de métadonnée ne correspond pas vraiment à la notion de donnée à caractère personnel dans la mesure où les métadonnées peuvent contenir des informations qui ne sont pas nécessairement des informations permettant d'identifier une personne (informations relatives à la localisation, au temps passé sur un fichier, etc...). Elles peuvent toutefois soulever des questions proches susceptibles d'intéresser la vie privée.

5 Métadonnées & Enjeux Socio-Économiques

Dans la société de l'information actuelle, les métadonnées deviennent 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. C'est un thème d'actualité qui fait peur au public. 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 participants pour assurer la gestion des projets.

L'objectif de cette section 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.

5.1 Les Données, Puissance du Futur

Nous vivons une période de rupture, celle de la numérisation de tout : l'homme, la société, les organisations, le savoir, les interactions... Les données constituent les briques de base de la société de l'information. Leur quantité est en croissance exponentielle. Désormais qualifiées de Big Data dans le monde anglo-saxon, elles représentent déjà des masses considérables pour lesquelles on recourt à de nouvelles mesures. L'infrastructure physique de la société de l'information, les systèmes de télécommunications, les centres de stockage et de traitement des données, les nouveaux services en ligne, constituent des secteurs industriels qui connaissent une croissance inégalée. Les données en elles-mêmes 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.

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. La maîtrise de la donnée permet aussi la maîtrise de certains marchés qui, actuellement, transitent déjà dans certains domaines par les outils de commerce électronique américains. 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.

La captation de données est la priorité absolue de certains pays tels que les États-Unis ou la Chine (qui détiennent respectivement 72% et 16% des 50 premiers sites mondiaux). Dans ces deux pays, les données nationales restent sous contrôle de l'industrie nationale. Et tous deux ambitionnent de récolter la donnée à l'international.

5.2 Localisation des Données

La localisation géographique du fournisseur de cloud computing peut avoir de réelles répercussions sur la protection et la confidentialité des données.

Les obligations légales Les données sensibles peuvent être stockées via une solution de cloud computing. Mais, pour une entreprise française par exemple, il faut alors vérifier que le fournisseur s'engage à conserver ces documents en France. Sinon, le risque est de ne pas pouvoir vérifier le bon traitement des données à caractère personnel et notamment le respect de la législation en vigueur pour l'entreprise (ex en France : durée de conservation des données, possibilité de modification et de suppression des informations...).

De même, il est généralement nécessaire de se conformer à certaines obligations légales en matière fiscales : interdiction de déposer des documents comptables en dehors de l'Union Européenne, déclaration obligatoire à l'administration fiscale pour pouvoir stocker des factures électroniques en dehors du territoire national...

Le "PATRIOT Act" Dans [VHAVE12] des chercheurs en droit néerlandais ont mis en lumière l'intérêt, pour une entreprise européenne, de choisir un fournisseur européen pour externaliser le traitement des données personnelles ou des informations capitales pour l'entreprise. En effet, depuis l'instauration du PATRIOT Act, la législation américaine permet à ses services de sécurité d'accéder à toutes les données à caractère personnel [Lee03] :

— des sociétés américaines, même si les données sont stockées physiquement sur le territoire européen

— de leurs filiales, même si elles sont implantées dans un autre pays du monde
— des serveurs qui sont hébergés aux États-Unis, y compris si la société qui possède ces serveurs est d'une autre nationalité

Le gouvernement américain s'est ainsi doté d'un arsenal juridique qui lui permet contrôler les données privées des citoyens étrangers, dont les Européens, en mettant à contribution ses grandes compagnies telles que Facebook, Google ou Microsoft. Les auteurs du rapport "Combattre le cybercrime et protéger la vie privée sur le Cloud" [EU12a] remis au Parlement Européen fin 2012 dénoncent le "Foreign Intelligence and Surveillance Act" (FISA). Cette législation autorise expressément les agences de renseignement américaines (NSA, CIA...) à mettre sur écoute sans autorisation judiciaire des citoyens américains communiquant avec des étrangers soupçonnés de terrorisme ou d'espionnage. Pour simplifier, un tribunal secret est désormais capable d'émettre un mandat, secret lui aussi (le "secret" des actions peut s'imposer pour une durée indéterminée), obligeant les entreprises américaines à livrer aux agences de renseignement américaines les données privées d'utilisateurs étrangers. Donc vos informations peuvent être dupliquées, conservées et divulguées à des tiers sans que vous en soyez informé... En décembre 2012 cet amendement a été prolongé jusqu'en 2017.

Notre objectif n'est pas de porter un jugement sur le "PATRIOT Act". Nous indiquons simplement que, dans l'état actuel des législations, une entreprise européenne ayant des contraintes fortes sur la confidentialité de ses données doit donc être vigilante quant au choix de son prestataire de services (localisation des données et nationalité du prestataire).

5.3 Vers une CNIL Européenne

Le Parlement Européen a bien avancé sur la réforme de la législation européenne en matière de protection des données proposée par la Commission [EU12b]. Les États-Unis, qui sont en train de réformer leur propre législation, appellent à la convergence réglementaire transatlantique, soulignant qu'ils se montrent tout aussi exigeants que les Européens en la matière. Il faut dire que l'Union a l'ambition de devenir la référence mondiale en matière de protection des données, laissant supposer (d'après les auteurs) que les États-Unis sont plus laxistes.

L'une des questions les plus controversées concerne l'exigence de normes équivalentes pour autoriser le transfert de données européennes vers un pays tiers pour y être traitées. Le problème se pose aussi pour les groupes mondiaux (ex : Google), dont les pratiques de traitement devraient être homologuées par l'Union, alors que les États-Unis aimeraient continuer à utiliser leurs codes de conduite.

En janvier 2013 le Parlement Européen a présenté ses rapports préliminaires sur la future réforme de la directive européenne sur la protection des données personnelles suite aux propositions de la Commission européenne. Souhaitant renforcer la protection des données de ses citoyens, l'Europe s'apprête à entamer la révision des mesures entrées en vigueur en 1995 en assurant vouloir remplacer la directive 95/46/EC sur la protection des données par un règlement européen que l'ensemble des États membres devra appliquer sans discussion. Cette réforme devrait passer par la création d'une autorité administrative indépendante, une "CNIL européenne", qui permettrait d'assurer l'application des règles en matière de protection des données et qui pourrait prendre la forme d'une agence indépendante.

5.4 Synthèse

Les instances politiques et économiques ont pris conscience de la nécessité de définir un cadre juridique international pour contrôler la collecte, le stockage et l'utilisation des données. Les métadonnées associées à ces données sont également concernées. Elles constituent en effet une valeur ajoutée de la plus haute importance pour les sociétés qui ont pour cœur de métier la gestion de l'information.

6 Conclusion

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 [GZ09, ZPL08]. Ceux-ci ne sont toutefois pas toujours adaptés à notre problématique de contrôle d'usage où justement certains indicateurs ne doivent pas être anonymes. Notre approche consiste plutôt à formaliser le "contrat de collaboration" entre partenaires professionnels. Pour la communauté informatique, ceci se fera en terme de langage de spécification des métadonnées à collecter, 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 (cf. Section 2) 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 ?

Finalement, nous avons abordé à la section 5 certains enjeux socio-économiques sous-jacents au stockage massif de données (et métadonnées) dans la société de l'information qui est la nôtre aujourd'hui. Au-delà de la collaboration entre partenaires sur un projet se pose également la question du recours à des prestataires de services sur le "cloud" (stockage ou traitements). Ces technologies sont devenues incontournables pour les entreprises bien qu'elles introduisent de nouvelles vulnérabilités quant à la sécurité de l'information (perte de la maîtrise de l'information). Ces menaces ne sont pas uniquement techniques (matériel, logiciel, réseau). Elles peuvent également être politiques, ce qui nécessite la définition d'un cadre juridique international pour la protection de l'information.

Références

[CCB08] Frédéric Cuppens and Nora Cuppens-Boulahia. Modeling contextual security policies. *Int. J. Inf. Sec.*, 7(4) :285–305, 2008.

[Dro13] Camille Drouiller. La preuve par les métadonnées. Master's thesis, 2013.

[EU12a] EU. *Fighting cyber crime and protecting privacy in the cloud*. EU Parliament, 2012.

[EU12b] 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.

[GZ09] Paolo Guarda and Nicola Zannone. Towards the development of privacy-aware systems. *Inf. Softw. Technol.*, 51(2) :337–350, February 2009.

[Inm95] William H. Inmon. Tech topic : What is a data warehouse ? *Prism Solutions*, 1, 1995.

[JMA13] Elena Jaramillo, Manuel Munier, and Philippe Aniorté. Information security in business intelligence based on cloud : A survey of key issues and the premises of a proposal. In *WOSIS*, 2013.

[KBM+03] Anas Abou El Kalam, Salem Benferhat, Alexandre Miège, Rania El Baida, Frédéric Cuppens, Claire Saurel, Philippe Balbiani, Yves Deswarte, and Gilles Trouessin. Organization based access control. In *POLICY*, pages 120–131. IEEE Computer Society, 2003.

[KRT+08] Ralph Kimball, Margy Ross, Warren Thornthwaite, Joy Mundy, and Bob Becker. *The Data Warehouse Lifecycle Toolkit*. Wiley Publishing, 2nd edition, 2008.

[Lee03] Laurie Thomas Lee. USA PATRIOT Act and telecommunications : Privacy under attack. *Rutgers Computer & Tech. LJ*, 29 :371, 2003.

[LMG13] Vincent Lalanne, Manuel Munier, and Alban Gabillon. Information security risk management in a world of services. In *PASSAT*, 2013.

[MLAR13] Manuel Munier, Vincent Lalanne, Pierre-Yves Ardoy, and Magali Ricarde. Legal issues about metadata : Data privacy vs information security. In *DPM*, 2013.

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

[VHAVE12] Joris Van Hoboken, Axel Arnbak, and Nico Van Eijk. Cloud computing in higher education and research institutions and the USA PATRIOT Act. *Social Science Research Network Working Paper Series*, November 2012.

[ZPL08] Bin Zhou, Jian Pei, and WoShun Luk. A brief survey on anonymization techniques for privacy preserving publishing of social network data. *SIGKDD Explor. Newsl.*, 10(2) :12–22, December 2008.