Granular Information Protection Framework

David Doret

28.05.2014

# Point de situation

La rapidité avec laquelle la société de l’information a émergé ces 25 dernières années a débouché sur une situation où le volume d’informations traitées et échangées par ses acteurs a atteint des niveaux historiquement inimaginables et où les solutions mises en œuvre pour sécuriser cette masse d’informations sont incapables de répondre aux problématiques qu’ils engendrent.

Sous l’angle technique, deux techniques clefs utilisées pour sécuriser l’information sont l’encryption et l’authentification. L’encryption permet de réserver l’accès à l’information aux seuls détenteurs de ses clefs et l’authentification permet de garantir son authenticité et son intégrité. La recherche a considérablement progressé dans ce domaine mais il ne s’agit que d’outils. Leur emploi par la multitude hétérogène d’acteurs traitant des informations est chaotique par nature.

Sous l’angle juridique, les parlements élaborent et renforcent continuellement des règlementations visant à protéger les intérêts de leurs citoyens. Ces efforts sont indispensables et bienvenus mais la réglementation reste difficile à appliquer, en particulier pour les acteurs internationaux qui sont confrontés à de nombreuses juridictions. Il n’existe pas de moyen par lequel ces réglementations pourraient être mises en œuvre puis évoluer sans produire des coûts considérables pour l’industrie.

Sous l’angle de la normalisation, divers standards techniques et organisationnels ont émergés. Ces normes permettent aux acteurs d’améliorer l’interopérabilité des systèmes informatiques et d’autre part de démontrer avec un niveau de confiance raisonnable que l’organisation interne d’un acteur est adéquate pour répondre aux exigences du propriétaire des données. Toutes bien conçues qu’elles sont, ces normes ne répondent qu’en partie à cette problématique globale et leur adoption dépend de la bonne volonté des acteurs à en assumer les coûts.

Au final, force est de constater que les personnes physiques et morales perdent le contrôle sur les informations qu’ils détiennent dès lors qu’ils les transmettent à un tiers. Concrètement, lorsqu’un particulier confie une information personnelle par le biais d’un quelconque réseau social sur Internet, lorsqu’une entreprise envoie un document officiel à un partenaire d’affaires, ils abandonnent dans une large mesure leur capacité à vérifier que les contraintes spécifiques liées à cette information seront respectées, à commencer par sa confidentialité. Ils abandonnent dans une large mesure leur capacité à amender ou retirer de la circulation cette information par la suite.

Je propose ici une esquisse de solution à ces défis majeurs: le *Granular Information Protection Framework*, ou *GIPF*. Il s’agit d’une méthodologie supportée par des systèmes informatiques spécialisés. Le GIPF a pour objectifs:

* De protéger les informations en tant que telles plutôt que les canaux et les systèmes qui les traitent,
* De permettre aux propriétaires d’informations de garder un contrôle raisonnable sur leurs informations après transmission à des tiers,
* De permettre aux acteurs tiers de démontrer aux propriétaires d’informations qu’ils sont à même de respecter leurs exigences,
* De réduire les coûts de mises en conformité des acteurs de l’industrie avec d’une part la multitude d’exigences légales émanant des états et d’autre part la multitude d’exigences particulières émanant des propriétaires d’informations.

Ce document a pour objectif de soumettre la solution proposée à la critique pour en éprouver la robustesse et, si cette première étape devait être franchie avec succès, de trouver des soutiens en vue de sa réalisation concrète.

# Une granularité inadéquate

Très sommairement, je montre ici comment le niveau de granularité employé pour protéger les données est généralement inadéquat.

Lorsque le chiffrement est employé pour protéger les informations transmises par un canal, l’ensemble des informations transmises est protégé. Le destinataire des informations à l’extrémité du canal est présumé seul capable de décrypter les informations reçues.

Lorsque le chiffrement est employé pour protéger des données stockées (base de données, système de fichiers, etc.), l’ensemble des informations stockées sont protégées. Le propriétaire du système est présumé seul capable de décrypter les informations stockées.

Lorsque le chiffremetn est employé pour protéger un document, l’ensemble du document est chiffré et seuls les détenteurs des clefs de décryption seront capables d’en déchiffrer le contenu.

Mécaniquement, ces cas d’utilisation ne permettent pas de répondre aux exigences particulières et variées des informations individuelles transmises, traitées ou stockées. Les informations sont traitées sans discrimination.

Ce manque de granularité est la pierre d’achoppement qui rend impraticable la protection des informations.

# Qu’est-ce qu’une information ?

Avant de poursuivre, définissons rapidement ce que nous entendons ici par *information*.

Informatique : Elément de connaissance susceptible d’être représenté à l’aide de conventions pour être conservé, traité ou communiqué. (Larousse 2012)

Renseignements consignés sur un support quelconque dans un but de transmission des connaissances. (Fiche terminologique de l’Office québécois de la langue française, <http://gdt.oqlf.gouv.qc.ca/ficheOqlf.aspx?Id_Fiche=8363906#eng> ).

Un élément de connaissance, ce peut-être :

* Je propose à la société X la vente du produit Y pour la somme Z,
* Ma date de naissance est le \*\*/\*\*/\*\*\*\*,
* Mon numéro de carte de crédit est le \*\*\*\*-\*\*\*\*\*-\*\*\*\*-\*\*\*\*,
* Ma petite amie m’a quitté, je suis furieux et je déclare que c’est une trainée.

# Protéger l’information avant tout

L’idée fondamentale du GIPF est d’adopter une perspective différente et de protéger, avant tout, les informations elles-mêmes plutôt que leur support. En général, c’est en effet en lien avec des informations spécifiques que leurs propriétaires ont des intérêts à défendre et des exigences particulières.

## Le lien de confiance

Certaines évidences méritent parfois d’être rappelées : lorsqu’un acteur souhaite confier une information à un tiers, il doit lui faire confiance. La valeur du lien de confiance doit être au moins équivalente à la valeur de l’information transmise.

Pour établir ou renforcer le lien de confiance, un acteur peut mettre en œuvre divers moyens dont:

* Prendre acte et accepter les conditions contractuelles proposées,
* Exiger des conditions contractuelles particulières,
* Se renseigner sur la réputation du tiers,
* Se renseigner sur les mesures techniques et organisationnelles mises en œuvre par le tiers pour protéger les informations,
* Exiger qu’un auditeur indépendant atteste de l’efficacité des mesures techniques et organisationnelles mises en œuvre.

Mais en pratique, ces moyens sont difficiles à mettre en œuvre, coûteux, difficiles à vérifier et nécessitent des compétences techniques, juridiques et organisationnelles pointues.

Le GIPF ne change évidemment rien à la nature des hommes ou de leurs organisations. La confiance restera toujours un élément primordial lorsque des informations seront confiées à un tiers. Cela dit, le GIPF permet la mise en place de contrôles techniques tels que cette confiance puisse être établie sur la base de facteurs objectifs et vérifiables. Mieux, le GIPF permet au propriétaire de l’information, sous certaines conditions, de garder le contrôle sur ses informations après les avoir transmises à un tiers.

# Exigences spécifiques du système GIPF

En synthèse, j’inventorie ici les exigences nécessaires auxquelles le système GIPF devrait satisfaire pour adresser cette problématique :

* R-1 : Le système doit permettre aux propriétaires d’information de transmettre des informations chiffrées, déchiffrables par les seules personnes autorisées,
* R-2 : Le système doit permettre aux propriétaires d’information de manipuler les informations protégées pour composer des documents dans le format de son choix et à partir de multiples informations,
* R-3 : Le système ne doit pas se substituer aux applications informatiques utilisées par les acteurs mais doit permettre à ces applications de devenir compatible avec lui,
* R-4 : Le système doit apporter un bénéfice à ses utilisateurs sous forme des gains de productivité de sorte à en faciliter l’adoption,
* R-5 : Le système ne doit pas dépendre d’une plateforme ou d’une technologie spécifique,
* R-6 : Le système doit être évolutif de sorte à s’adapter aux exigences futures que pourront avoir les acteurs en lien avec le stockage et le traitement des informations,
* R-7 : Le propriétaire d’information doit pouvoir marquer les informations qu’il transmet de sorte à ce qu’il puisse être vérifié par la suite que c’est bien lui qui en est bien le propriétaire,
* R-8 : Mais dans certaines circonstances, l’anonymat du propriétaire d’information doit pouvoir être garanti après transmission,
* R-9 : Les informations doivent être identifiables,
* R-10 : Les informations transmises à un tiers doivent être spécifiquement identifiables,
* R-11 : Les informations doivent pouvoir être structurées de sorte à en faciliter l’exploitation automatique,
* R-12 : L’authenticité et l’intégrité des informations doit pouvoir être établie par les acteurs traitant ces informations,
* R-13 : Le propriétaire d’information doit pouvoir adosser des exigences particulières aux informations qu’il transmet. Par exemple, le propriétaire d’information peut imposer que ses informations ne soit pas transmises à des tiers, ne soient pas transmises à l’étranger, puissent être utilisées de façon agrégées à des fins statistiques ou d’autres exigences diverses,
* R-14 : Le propriétaire d’information doit pouvoir confier ses informations à des tiers en ayant un niveau de confiance raisonnable que ses exigences particulières seront respectées,
* R-15 : Dans certaines circonstances, le propriétaire d’information doit être en mesure d’amender les informations après transmission,
* R-16 : Dans certaines circonstances, le propriétaire d’information doit pouvoir effacer les informations après transmission.

# Le modèle GIPF

Je décris ici, dans les grandes lignes, les différents éléments et processus caractérisant le GIPF.

En premier lieu, les informations originales sont gérées séparément des documents ou des systèmes qui les contiennent généralement. Elles sont individuellement contenues dans des enveloppes protectrices.

Leur stockage se fera typiquement dans des bases de données spécifiques.



|  |  |
| --- | --- |
| <Information>  <UnlockedContent>Je suis né le \*\*.\*\*.\*\*\*\*\*.</UnlockedContent>  </Information> | L’enveloppe  Le contenu de l’enveloppe |

Les enveloppes protectrices permettent de décorer l’information avec des informations contextuelles riches. Ces informations contextuelles peuvent par exemple: identifier l’information elle-même, identifier son propriétaire, informer des exigences particulières exigées par leur propriétaire, etc.



Les enveloppes protectrices permettent d’appliquer des mécanismes de protection sur l’information qu’elles contiennent et de décrire les mécanismes utilisés. Il peut s’agir de chiffrement :

|  |  |
| --- | --- |
| <Information>  <Lock scheme="System.EAX.1">  <PublicKey>...</PublicKey>  <AuthenticationTag>...</AuthenticationTag>  </Lock>  <LockedContent>...</LockedContent>  </Information> | L’algorithme d’encryption  Une référence à une clef |

Ou d’une signature électronique permettant d’authentifier le document :

|  |  |
| --- | --- |
| <Information>  <Lock scheme="System.CMAC.1">  <PublicKey>...</PublicKey>  </Lock>  <LockedContent>Je suis né le \*\*.\*\*.\*\*\*\*\*.</LockedContent>  </Information> | L’algorithme d’encryption  Une référence à une clef |

Les enveloppes protectrices peuvent contenir des enveloppers protectrices et constituer ainsi une hierarchie d’informations auxquelles différentes protections sont appliqués :

|  |  |
| --- | --- |
| <Information>  <UnlockedContent>  <Information>  <UnlockedContent>...</UnlockedContent>  </Information>  <Information>  <UnlockedContent>...</UnlockedContent>  </Information>  </UnlockedContent>  </Information> | L’algorithme d’encryption  Une référence à une clef |

Les inforemations peuvent être manipulées, transmises et strockées.

le propriétaire d’information, qu’il s ’agisse d’une personne phyisque ou morale,

O. Propriétaire d’information

A. Acteur traitant de l’information.

EXEMPLES

Pour expliquer ce qu’est le GIPF et démontrer par quels mécanismes il entend répondre à ces différentes problématiques, je propose cas d’utilisation pratiques.

CAS 1 : Informations commerciales transmises à une entreprise

La société X envoie à la société Y une offre pour diverses prestation. Il est dans son intérêt de garantir la confidentialité de ces informations car des sociétés concurrentes pourraient utiliser ces informations pour proposer des offres concurrentielles et lui voler des parts de marché.

CAS 2 : Informations personnelles postées sur un réseau social

Jean poste des infromations personnelles sur un réseau social. Ces informations sont d’ordre privé et il souhaite que seul un cercle restreint d’amis y aient accès. Par la suite, sa situation personnelle ayant évolué, Jean souhaite modifier ces informations. Plus tard encore, Jean souhaite les effacer complètement.

CAS 3 : Carte de crédit

Jean achète en ligne un produit quelconque. Pour ce faire, il fournit diverses informations sensibles parmi lesquelles son numéro de carte de crédit.

CAS 4 : Le manager X confie à l’employé Y la rédaction de divers éléments dans un document. Ce document comprend un jeu d’informations spécifique que l’employé Y n’est pas habilité à voir, par exemple des projections financières.

Pour démontrer comment le GIPF propose de protéger ces informations, nous illustre

Le Granular Information Protection Framework a pour ambition de repondre prefisemment a cette demande-la et de redonner aux proprietaires des informations le controle des informations qu ils diffusent.

Il permet a un acteur, soit une personne privee ou morale, de proteger des informations specifiques.

Exemple: un post sur un reseau social

Exemple: un message prive a son conjoint

Anne et Louis sont amis depuis quelques mois

Anne envoie un message coquin a Louis

Elle souhaite que ce message reste en toute circonstance prive

L email se retrouvera par exemple stocke sur l ordinateur personnel de Louis

Anne est proprietaire de son message meme apres la rupture avec Jean

Exemple: un citoyen fourni une informationa une administration publique

Exemple: carte de credit

Jean Dupont achete une trotinette sur internet. Il paye avec sa carte de credit. Le magasin en ligne lui demande des informations personelles. Cette demande peut etre consideree comme legitime ou non. Mais jean, s'il veut acheter sa trotinette, devra se plier a cette contrainte. A partir du moment ou il aura fourni ses informations personellles au magasin en ligne, il n exercera plus aucun controle sur les donnees.

Jean donc saisi une information protegee. L'information contient toujours sses informations personelles mais ces dernieres sont signees digitalement. Mieux, divers mecanismes sont mis en oeuvre pour . Jean saura que cette information la il l a fournie a ce magasin precisemment.

Commercialisation

Un modele ouvert et open source pour attirer un maximum d acteurs

Commercialement , lobjectif sera

de constituer l'institut qui assurera la gouvernance du systeme

De commercialiser des produits (add ons , apis, etc) permettant aux entreprisees et aux particuliers de consommer plus facilement le framework

Criteres pour une solution

Les informations sont identifiables

Les mecanismes d encrPtion, authentication, etc. Sont ouverts et evolutifs pour eviter l obsolescence rapide technologique

Les mecanismes de protection peuvent etre chaines, appliques les uns sur les autrs pour crer des schemas de protection complexe

Les informations peuvent etre identifiees, classees, retrouvees, etc.

Un prototype

# Le modèle agency

Ce modèle enrichi le GIPF original en appliquant une organisation particulière au modèle.

Plutôt que de confier ses informations à d’innombrables tiers pour lesquels il lui est difficile de vérifier le serieur, il confie ses informations à une agence. Les tiers n’auront d’autre choix que d’utiliser l’API, etc.