Gestion de la confidentialité dans un réseau social

Julien Kerglonou

RAPPORT DE STAGE

2 JUIN 2011

GESTION DE LA CONFIDENTIALITÉ DANS UN RÉSEAU SOCIAL

MASTER 2 RECHERCHE EN INFORMATIQUE

UNIVERSITÉ DE BRETAGNE OCCIDENTALE
TÉLÉCOM BRETAGNE

3 FÉVRIER — 30 JUIN 2011

Encadré par : M. Eric COUSIN

JULIEN KERGLONOU

Résumé

L'avènement des réseaux sociaux en ligne constitue un véritable sujet de société, surtout lorsqu'il s'agit de gestion de la confidentialité des informations partagées. Ce rapport livre une démarche de réflexion concernant l'élaboration d'un système interactif permettant à l'utilisateur des plateformes de réseautage social en ligne de pouvoir y évoluer tout en contrôlant au mieux possible la gestion d'accès à ses données.

Table des matières

TABLE DES FIGURES

Introduction

Ces dernières années, les services de réseautage social en ligne sont en pleine expansion : *Facebook*[1], *Twitter*[2], *YouTube*[3], *Picasa*[4], *MySpace*[5] – pour n'évoquer que les plus populaires – hébergent des centaines de millions de comptes dans le monde. Plus ou moins spécialisés, ils permettent principalement de partager des informations (profil, photos, statut, humeur, vidéos, etc.) avec un réseau "d'amis".

Véritable phénomène de société, ces réseaux sont au cœur de nombreux débats, notamment vis-à-vis des problématiques de confidentialité. La notion de confidentialité a été définie par l'Organisation internationale de normalisation (ou ISO) comme "le fait de s'assurer que l'information n'est seulement accessible qu'à ceux dont l'accès est autorisé". Le principe même des réseaux sociaux en ligne étant le partage d'informations, il est important de savoir comment ce partage est géré, ou encore définir à quel niveau de confidentialité placer de telles informations. A l'heure où une confusion nouvelle entre espace privé et espace public est introduite, il est primordial de garantir aux utilisateurs à la fois une compréhension concernant la visibilité des contenus qu'ils diffusent, et une maîtrise de cette diffusion.

A l'heure actuelle, les outils permettant aux utilisateurs de gérer la confidentialité de leurs données sur les plateformes de réseautage social en ligne (notamment les I.H.M.[6] de gestion de confidentialité) ne rendent pas toujours compte de la complexité des réglages. Il est alors très difficile d'avoir une compréhension globale des modifications que l'on peut apporter, et l'utilisateur peut être très vite perdu. Typiquement, lorsqu'un utilisateur A commente une nouvelle d'un ami B, qui est autorisé à voir ce commentaire ? Seulement A et B ? Les amis de A ? Les amis de B, même s'ils n'ont aucun lien avec A ? On est là confronté aux souhaits de confidentialité de personnes distinctes – voire potentiellement contradictoires – qui ont des cercles d'amis distincts ou appartiennent à des réseaux différents. Cela devient vite très complexe ! Or, les enjeux sont considérables. Les médias se font largement écho de faits divers ([1]), ou autres licenciements ([2], [3]) liés à la consultation d'informations publiées sur *Facebook* par des utilisateurs croyant, à tort, s'exprimer au sein d'un cercle restreint d'"amis". Une fois en possession d'une information personnelle, un utilisateur "indésirable" peut à la guise la copier et la diffuser comme bon lui semble, pouvant alors porter préjudice à son propriétaire. Il est donc essentiel que de tels utilisateurs ne puissent accéder aux informations personnelles d'autrui que si ils y sont autorisés spécifiquement par ceux-ci selon leur volonté. Les risques de se dévoiler se doivent d'être mesurés et contrôlés par l'utilisateur de la manière la plus adaptée possible. Pour cela, le système se doit de l'informer, et l'aider à décider (car c'est un travail difficile).

Dans ce cadre, mon travail est donc de dégager un ensemble de fonctionnalités qui, au sein d'un réseau social en ligne généraliste, aiderait l'utilisateur à garder la maitrise de l'accès aux données personnelles qu'il introduit et diminuerait ses risques d'exposition envers des utilisateurs indésirables, et ce de manière ergonomique, lui permettant de mieux saisir l'impact de la diffusion de ses données.

L'objectif de ce travail est donc de comprendre comment adapter le système d'une plateforme de réseautage social en ligne dans son ensemble afin de répondre à ces problématiques de gestion de confidentialité. *Facebook* étant le plus utilisé et le plus représentatif de ce type de réseaux sociaux en ligne, il nous servira donc de cas d'étude, de repère.

1. `http://www.facebook.com/`
2. `http://twitter.com/`
3. `http://www.youtube.com/`
4. `http://www.picasa.google.com`
5. `http://www.myspace.com/`
6. Interfaces Homme-Machine

Dans un premier temps, nous préciserons le contexte de l'étude afin de bien comprendre son cadre. Puis, dans une deuxième partie, nous nous appuierons sur des études de sociologie, d'analyse des réseaux sociaux (mathématique – théorie des graphes – et du point de vue web sémantique) et de formalisation du contrôle d'accès, pour en faire ressortir un ensemble de fonctionnalités pertinentes répondant aux problématiques de départ. Nous verrons enfin dans une troisième et dernière partie comment utiliser ces fonctionnalités au sein d'une plateforme de réseautage social généraliste.

Chapitre 1

Mise en Contexte

Introduction

Avant tout, il est nécessaire de poser les bases du sujet. Dans cette partie, nous tâcherons donc de bien cerner le contexte de notre travail. Nous définirons ainsi dans un premier temps différentes notions inhérentes aux réseaux sociaux en ligne.

L'utilisateur gérant les droits d'accès à ses données via une interface de gestion de la confidentialité, nous définirons ce qu'est une I.H.M. et nous verrons en quoi la notion d'ergonomie est au cœur de ce concept.

Puis dans un dernier temps, nous nous pencherons plus en détail sur notre cas d'étude, à savoir la plateforme *Facebook*. Nous tenterons d'abord de comprendre son fonctionnement, puis nous émettrons des critiques quant à la gestion d'accès aux données personnelles qu'un utilisateur peut mener sur celle-ci.

1.1 Les réseaux sociaux

1.1.1 Présentation générale

Un réseau social est un ensemble d'entités sociales (telles que des individus ou encore des organisations sociales) reliées entre elles par des liens créés lors des interactions sociales. Il se représente par une structure ou une forme dynamique d'un groupement social [4]. Le terme interaction sociale désigne un ensemble de relations humaines alimentées par des échanges faisant suite à des actions, verbales ou non.

Un réseau social peut donc être représenté par un graphe, appelé "graphe social", dans lequel les sommets désignent les entités sociales, et les arêtes leurs interactions sociales. La figure 1.1 dans [5] représente par exemple, un réseau social dont les entités sont des individus.

La notion de réseau social s'étend aujourd'hui en ligne, par le biais de différentes plateformes telles que *Facebook*, *Twitter*, *Myspace*,... dans lesquels l'utilisateur est amené à renseigner certains champs le concernant (son nom, son âge, ses centres d'intérêts,...). Ces informations définissent alors son identité numérique au cœur d'une page généralement appelé "profil". Par la suite, il se

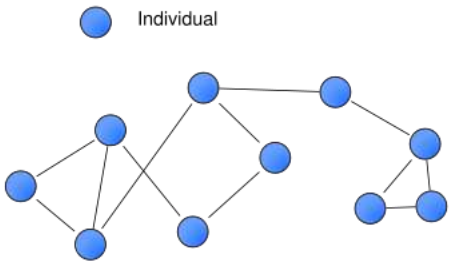


FIGURE 1.1 – Un graphe social

constitue un réseau de connexions par le biais d'interactions avec les autres utilisateurs, partageant alors certaines des informations avec eux.

Les réseaux sociaux en ligne s'inscrivent (notamment depuis l'avènement de la plateforme *Facebook* en 2007) dans une véritable révolution socio-culturelle. Ils représentent en effet de tout nouveaux modes de communication et de sociabilité, permettant à leurs utilisateurs de partager une multitude d'informations avec leurs connexions [1], que ce soit des photos, vidéos, pensées, à titre personnel ou professionnel, le tout en réduisant considérablement les barrières de communications que peuvent constituer l'espace ou le temps. Il permettent également d'entretenir des liens entre les utilisateurs, et même de faire de nouvelles rencontres.

Cependant, ce partage si facilité d'informations entre utilisateurs introduit également de nombreuses problématiques, concernant notamment la sécurité, et plus particulièrement la confidentialité des données.

1.1.2 IDENTITÉ NUMÉRIQUE

L'identité numérique est en soi un concept complexe, pour lequel plusieurs définitions ont été proposées. Globalement, il s'agit de l'ensemble des informations en ligne entré par un internaute, ses contributions, ou les traces qu'il laisse sur les sites web visités. C'est un lien technologique entre une entité réelle et une entité virtuelle [6].

La FING [2] va plus loin et indique qu'à l'heure actuelle, "L'identité numérique n'est plus le simple reflet, plus ou moins fidèle, de l'identité civile. Il faut d'abord parler d'identités numériques, au pluriel : nous en avons tous plusieurs. Et ces identités sont actives : changeantes, mobiles, expressives, négociables, valorisables, elles deviennent les outils et les ressources grâce auxquels l'individu organise son existence numérique, au service de ses valeurs, de ses objectifs et de ses priorités."

Il s'agit donc, dans le cadre de notre travail, de permettre à l'utilisateur de faire évoluer son identité numérique au sein du réseau social comme il le souhaite, tout en ayant connaissance des utilisateurs ayant accès et pouvant interagir sur les données associées à celle-ci.

1.1.3 POLITIQUES DE CONFIDENTIALITÉ

L'introduction de données personnelles est au centre de l'élaboration d'une identité au sein d'un réseau social. Il est donc absolument primordial de pouvoir gérer la diffusion de telles informations

1. Le terme connexions est ici générique ; On peut parler d'amis, de contacts,... L'appellation dépend du réseau social. Le réseau *Facebook* étant notre cadre d'étude, nous parlerons, dans la suite de ce rapport, d'*amis*.
2. Fondation Internet Nouvelle Génération

de manière la plus souple et complète possible, rendant alors dans l'idéal certains types d'information publics, privés, réservés à un certain nombre de ses connexions, etc.

Pour cela, chaque réseau social dispose d'une *politique de confidentialité*, et d'un ensemble d'outils mis à disposition de l'utilisateur pour paramétrer la façon dont ses données sont diffusées sur le réseau. La société responsable du site de réseautage social en question s'engage à respecter cette déclaration, tout comme l'utilisateur.

Une politique de confidentialité est donc un contrat qui décrit comment une société retient, traite, publie et efface les données transmises par ses clients, ici les utilisateurs inscrits sur la plateforme. Par exemple, l'entreprise *Facebook* s'octroie le droit de collecter, à partir de l'appareil grâce auquel l'utilisateur se connecte, des informations sur son type de navigateur, le lieu d'où il l'a fait, son adresse IP ainsi que les pages consultées. Elle s'octroie également le droit de conserver toutes données mise en ligne sur la plateforme, et peut même les partager parfois avec des tiers "afin d'améliorer ou de promouvoir notre service".

Il est donc très important pour un utilisateur de consulter la politique de confidentialité associé à la plateforme sur laquelle il s'inscrit, afin de prendre connaissance de la manière dont la société responsable gère ses informations.

1.1.4 LES DIFFÉRENTS "TYPES" DE RÉSEAUX SOCIAUX EN LIGNE

Le terme "réseau social en ligne", reposant donc sur la notion de communauté d'utilisateurs sur une plateforme, est somme toute assez générique.

En effet, on peut distinguer de nombreux types de plateforme de réseautage social en ligne :

- Les réseaux généralistes (*Facebook*, *LinkedIn*,...)
- Les sites de rencontre (*Meetic*, *e-darling*,...)
- Les sites de partage publics (*Myspace*, *Twitter*, *Youtube*,...)
- Les jeux en ligne (*World of Warcraft*, *Poker Stars*,...)
- etc...

Chaque plateforme a ses propres fonctions et objectifs (voir Figure 1.2), rendant ainsi la politique de confidentialité très différente d'une plateforme à l'autre.

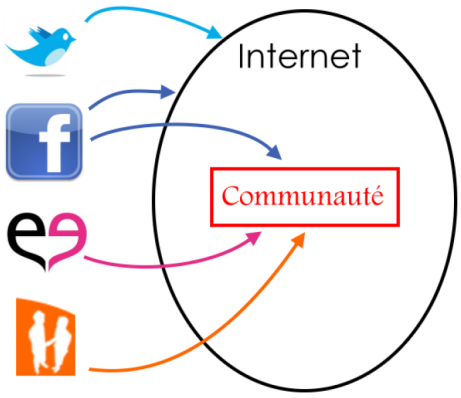


FIGURE 1.2 – Chaque plateforme a ses propres fonctions

En effet, si sur *Facebook* l'utilisateur peut ne rendre disponibles ses informations qu'à un ensemble défini d'internautes – sa famille, par exemple, ou ses amis –, il n'en est pas de même sur la plateforme *Meetic* par exemple où le principe même repose sur le fait de partager avec l'ensemble de la communauté, et uniquement avec elle, des informations personnelles pour faire des rencontres. Ainsi, sur *Meetic*, toute information est consultable par tous les utilisateurs, mais est indisponible pour les autres internautes. La plateforme *Twitter*, quant à elle, a vocation a rendre publiques toutes les informations émises, chacune d'entre elle pouvant faire l'objet de résultat de recherche sur le web : "Ce que vous dites sur *Twitter* peut être consulté instantanément à travers le monde", stipule leur politque de confidentialité. *Youtube* et *Myspace* se ressemblent un peu plus : les utilisateurs peuvent rendre leurs données accessibles à tous les internautes, l'enjeu étant une popularité maximale. Nous approfondirons cette étude plus loin.

Les problématiques de confidentialité ne se posent donc pas de la même manière selon le type de réseau social en ligne que l'on considère.

1.1.5 *Facebook* COMME CAS D'ÉTUDE

Avec plus de 700 millions d'utilisateurs [3], *Facebook* est le service de réseautage social le plus utilisé au monde, et le troisième site web le plus visité (648 millions de visiteurs uniques en novembre 2010 [4]), juste derrière *Google* (970 millions) et *Microsoft* (869 millions).

Facebook étant, de plus, un réseau social généraliste (sur lequel l'utilisateur peut partager tout ce qu'il veut, et avec qui il veut), nous avons décidé d'en faire notre cas d'étude.

1.2 Ergonomie et Interface Homme-machine

1.2.1 INTRODUCTION

Au sein d'une plateforme de réseautage social en ligne, l'utilisateur a la possibilité de paramétrer le niveau d'accessibilité à ses données pour les autres. Ce paramétrage se fait, bien entendu, différemment selon la plateforme considérée.

La "satisfaction" de l'utilisateur dépendra de différents critères, notamment l'ergonomie de l'*I.H.M.* – ou *Interface Homme-Machine*.

1.2.2 L'ERGONOMIE

La notion d'ergonomie a été définie en 1988 par la *SELF* [5] comme étant "la mise en œuvre de connaissances scientifiques relatives à l'homme, et nécessaires pour concevoir des outils, des machines et des dispositifs qui puissent être utilisés avec le maximum de confort, de sécurité et d'efficacité pour le plus grand nombre". L'ergonomie a donc pour but d'aider l'être humain dans la réalisation de tâches par l'introduction d'un système adapté à lui, et aux variabilités des individus constituant son peuple. Prenant ainsi appui sur les Sciences Humaines, l'ergonomie vise à adapter la machine à l'Homme. Pour se faire, elle se base sur trois niveaux :

– Niveau physique : la machine est adaptée à la morphologie de l'Homme.

3. Source : `http://www.socialbakers.com/blog/171-facebook-is-globally-closing-in-to-700-million-users/`
4. Source : Cabinet Comscore (`http://www.comscore.com/`)
5. Société d'Ergonomie de Langue Française

- Niveau social : Compatibilité entre l'environnement de travail et les attentes de l'ensemble des utilisateurs.
- Niveau cognitif : les informations sont traitées et représentées selon les attentes et les capacités de l'ensemble des utilisateurs. Ce dernier niveau est celui qui nous intéresse ici car il définit l'ergonomie des logiciels.

Pour répondre au critère d'ergonomie, un logiciel doit être compatible avec son utilisateur selon :

- l'adaptation des dispositifs d'interaction du système aux caractéristiques physiques de l'utilisateur et la nature de la tâche qu'il doit accomplir.
- la linguistique : l'information doit être représentée avec des symboles compatibles avec l'ensemble des pratiques et usages de la communauté à laquelle l'utilisateur appartient.
- l'activité : la structure du logiciel doit pouvoir s'adapter aux modes de raisonnement de l'utilisateur dans la réalisation de sa tâche.

1.2.3 L'ERGONOMIE AU SERVICE DES INTERFACES HOMME-MACHINES

Une Interface Homme-machine (ou I.H.M.) définit les moyens et outils mis en œuvre, afin qu'un humain puisse contrôler et communiquer avec une machine. Les ingénieurs en ce domaine étudient la façon dont les humains interagissent avec les machines ou entre eux à l'aide de celles-ci, ainsi que la façon de concevoir des systèmes qui soient ergonomiques, efficaces, faciles à utiliser ou plus généralement adaptés à leur contexte d'utilisation.

Une I.H.M. repose donc sur la notion d'ergonomie et peut revêtir différentes formes en fonction du domaine d'application. Dans notre cadre de travail, nous ne nous intéresserons qu'aux interfaces logicielles.

1.3 Le système de gestion de la confidentialité des données de la plateforme *Facebook*

1.3.1 PRÉSENTATION GÉNÉRALE DE *Facebook*

La plateforme *Facebook* propose à ses utilisateurs de renseigner un ensemble de champs les concernant (nom, photo, sexe, date de naissance, intérêts, opinions politiques ou religieuses, ville de résidence ou d'origine, langues parlées, intérêts,etc.) dans le but de se former un réseau dit d'*amis* (de connexions) et de partager par la suite avec eux, outre ses informations personnelles, un ensemble de données (opinions, documents multimédia, liens, etc.). L'intérêt premier de *Facebook* est donc de rester connecté avec ses amis, quel que soit leur lieu de résidence, et de façon instantanée, mais également de tisser des liens nouveaux avec des utilisateurs partageant les mêmes intérêts par exemple.

La quantité de données que l'on peut partager est donc très grande : photos, vidéos, liens, humeurs,... Au total ce sont plus de 30 milliards de contenus partagés chaque mois par la communauté, et plus de 900 million d'"objets" avec lesquels les utilisateurs peuvent intéragir (pages, groupes, évènements, pages communautaires,...) [6]

L'utilisateur peut ainsi s'élaborer une identité numérique au plus proche de son identité réelle (le système, et le principe même de la plateforme, l'y encourage même fortement).

6. Source : `http://www.facebook.com/press/info.php?statistics`

Lors de sa connexion à *Facebook* (par saisie d'identifiant – son adresse e-mail – et de mot de passe), l'utilisateur accède au *fil d'actualité*. Cette page d'accueil propose une liste des différentes publications, ou changements d'informations liés à ses amis. On y trouve donc statuts, photos, résultats liés à des applications, mais également des indications de commentaires, de pages "aimées",etc. On peut voir également les connexions créés entre un ami et un autre utilisateur de *Facebook*. L'utilisateur peut contrôler les informations affichées sur son fil d'actualité en agissant directement au sein ce dernier, et en paramétrant les publications une à une. Il peut ainsi empêcher l'affichage des informations concernant un de ses ami (que ce soit une publication précise, ou bien toute information à venir concernant cet ami), ou une application.

L'utilisateur peut également consulter son "profil", appelé communément *mur* sur lequel sont récapitulées par ordre déchronologique, toutes les interactions entre lui et ses amis. Si quelqu'un interagit avec lui, il reçoit une notification lui permettant d'accéder directement à la page associée. Il peut également créer des *groupes d'amis*. Un groupe, constitué d'au moins 2 membres, permet aux utilisateurs d'échanger certaines informations de manière plus ciblée, à certaines personnes. Le groupe peut être défini comme étant "ouvert" (la liste des membres et le contenu sont publics), "fermé" (la liste des membres est publique et le contenu privé), ou "secret" (la liste des membres et le contenu sont privés).

Facebook dispose également d'un service de messagerie instantanée, grâce auquel l'utilisateur peut dialoguer avec les amis connectés sur le service au même moment. Il est également possible d'installer des fonctionnalités optionnelles appelées *applications* : quizz, jeux, questions sur ses amis,etc. Une fois installées, l'utilisateur permet à ces applications d'accéder à un certain nombre d'informations personnelles le concernant.

1.3.2 POLITIQUE DE CONFIDENTIALITÉ

NB : La politique de confidentialité de la plateforme Facebook étant en perpétuelle évolution, ce qui suit est valable au 2 juin 2011.

Comme tous les réseaux sociaux, *Facebook* met à disposition des internautes une politique de confidentialité [7] que l'entreprise s'engage à respecter. L'utilisateur peut gérer ses paramètres de confidentialité par le biais d'une interface, dans laquelle il peut contrôler la visibilité de ses données. Je vais, dans ce paragraphe, essayer de rendre compte au mieux possible des informations, inhérentes à notre cadre d'étude, apportées par cette déclaration, mise à jour à la date du 22 décembre 2010 [7].

L'utilisateur peut à tout moment empêcher d'une part les internautes n'ayant pas de compte *Facebook* d'avoir accès à son profil via une recherche Internet (et dans ce cas, même en disposant de l'URL associé à son profil, aucun internaute non connecté à *Facebook* ne peut y accéder), et d'autre part les autres utilisateurs du réseau.

Le nom et la photo du profil sont deux informations qui ne sont pas soumis aux paramètres de confidentialité. Il est cependant possible de ne pas avoir de photo.

Un utilisateur peut contrôler l'accès à chacune des publications (partage de statut, photo(s), vidéo, ou lien) qu'il peut faire. Si il ne précise rien, ce sont les paramètres spécifiés par défaut qui sont appliqués.

A tout moment, un utilisateur peut soit désactiver son compte, soit le supprimer. La désactivation empêche toute personne de le consulter, mais il existe toujours ; il suffit simplement de s'identifier

7. Il est à noter que la politique de confidentialité de *Facebook* n'a cessé d'évoluer depuis sa création en 2005. Ce changement est au cœur de nombreuses polémiques.

pour le réactiver. La suppression, quant à elle est définitive, mais l'entreprise *Facebook* ne garantit pas que des copies d'informations ne subsistent pas, "dans la mesures où elles ont été partagées avec d'autres utilisateurs". De plus, certaines informations partagées avec d'autres utilisateurs ne peuvent être supprimées.

L'utilisateur peut "bloquer" certains utilisateurs, disparaissant alors de *Facebook* pour elles (aucune information, y compris le nom ou la photo de profil ne peut apparaître dans une page *Facebook*, que ce soit pour l'utilisateur bloqué ou pour celui qui a bloqué), mais également des invitations à utiliser une application, ou à participer à des évènements que certains utilisateurs pourraient proposer.

Le partage d'informations sur le profil d'un utilisateur (publication, ou commentaire) est régit par les paramètres de confidentialité de cet utilisateur.

Facebook précise également que si l'utilisateur place la visibilité d'une information via l'interface de gestion de confidentialité sur le paramètre "Tout le monde", cette information est visible par l'ensemble des internautes et "peut aussi être indexée par des moteurs de recherche tiers et être importée, exportée, diffusée et rediffusée par *Facebook* et par des tiers, sans restrictions de confidentialité".

Si un ami de l'utilisateur se connecte à une application, celle-ci peut avoir accès à certaines données concernant cet utilisateur : nom, photo du profil, identifiant, ses connexions, et toutes les informations qu'il aura défini sur le paramètre "Tout le monde". Il y a donc des informations dont le contrôle d'accès n'est pas géré par l'utilisateur.

Lorsqu'un utilisateur décède, *Facebook* peut rendre son compte "commémoratif" : il se retrouve accessible uniquement aux amis, qui peuvent alors écrire sur son mur.

Si l'entreprise *Facebook* se trouve être rachetée, les informations restent soumises à la politique de confidentialité actuelle[8]. sur des serveurs basés aux États-Unis.

L'entreprise *Facebook* indique également qu'elle ne peut pas garantir que les données publiées ne soient vues que par les personnes autorisées, ni qu'elles ne soient un jour rendues publiques, car "rien n'est parfait".

Une information importante, mais qui n'apparaît pas dans [7], est qu'il est impossible d'interdire à un autre utilisateur d'identifier, ou "marquer", une photo (c'est-à-dire indiquer quelle(s) personne(s) apparaît sur celle-ci) que l'utilisateur a publiée, dans la mesure où il y a accès. Il est cependant possible, une fois la photo identifiée, de supprimer ce "marquage" : "Contrôle pour les publications dans lesquelles vous avez été identifié(e) : c'est vous qui contrôlez qui peut voir les photos et vidéos dans lesquelles vous avez été identifié(e) et qui apparaissent sur votre profil. N'oubliez pas cependant que la personne qui publie la photo peut la montrer à d'autres personnes qui ne sont pas dans votre liste d'amis. Si vous ne souhaitez pas que votre nom soit associé à certaines photos, il suffit de le retirer. Cela l'empêchera également d'apparaître sur votre profil."
En effet, il y a une réelle difficulté concernant la gestion des photos sur *Facebook*. Le contrôle est en fait géré de plusieurs façons :

- L'utilisateur publie une photo via la fonctionnalité "Publier" (c'est-à-dire qu'il émet un post, unique, sous la forme d'une photo, et qui s'affiche sur son "mur"). Il peut gérer les droits d'accès à cette photo particulière. Si il "identifie" quelqu'un dessus, cependant, cette personne verra forcément la photo.
- L'utilisateur publie une photo au cœur d'un album (ensemble de photos) : dans ce cas il ne peut gérer les droits d'accès qu'à l'album dans son ensemble, et non pas à la photo uniquement.

8. On peut s'interroger sur la véritable garantie de cette affirmation.

– Un utilisateur B publie une photo sur laquelle un utilisateur A est "identifié". Dans ce cas, A peut soit retirer l'identification, soit gérer les droits d'accès de l'ensemble des photos dans lesquelles il est identifié, mais il ne peut pas gérer l'accès à cette photo uniquement, étant donné que ce n'est pas lui qui l'a publiée.

Cet exemple concernant la gestion d'accès aux photos illustre la difficulté pour un utilisateur de gérer sa confidentialité au sein du réseau.

Nous savons donc ce que *Facebook* permet ou non à l'utilisateur dans la gestion de sa confidentialité. Intéressons nous à présent à ce qu'il peut concrètement faire.

1.3.3 GESTION D'ACCÈS AUX DONNÉES

Lorsqu'il introduit une nouvelle donnée sur la plateforme, l'utilisateur a le choix de la rendre visible à :

– **Tout le monde** : tous les internautes,
– **Amis seulement**,
– **Amis et leurs amis**,
– **Certains amis** : l'utilisateur spécifie lesquels,
– **Ne pas montrer le contenu à certains amis** : l'utilisateur spécifie lesquels.

Des listes d'amis peuvent également être créés ; et les diverses informations publiées peuvent alors être rendues accessibles seulement à ces listes, ou a contrario, être inaccessible uniquement pour elles.

Une analyse complète du contrôle d'accès que l'utilisateur peut gérer concernant ses données sur la plateforme a été menée dans [8]. L'utilisateur peut contrôler l'accès à l'ensemble des informations qu'il partage (mis à part sa photo de profil et son nom), et ce de manière très complète – il peut même rendre toutes ses informations invisibles pour tout le monde. Il est donc, théoriquement, possible de gérer la quasi-totalité de ses informations (si l'on excepte ce que les autres partagent concernant l'utilisateur et dont il n'a pas connaissance).

Cependant, le problème réside dans la difficulté pour lui de mener à bien les différents réglages. Cela est dû, tout d'abord à leur abstraction : il est difficile de se rendre compte de l'influence que peut avoir le placement d'un paramètre à un certain niveau de visibilité. Concernant le terme "amis et leurs amis" par exemple : un utilisateur a en moyenne 130 amis sur la plateforme[9]. Il a donc, en moyenne, $130 + 130 \times 130$ "amis et leurs amis", soit 17 030 ! Un paramètre placé à "amis et leurs amis" touche donc un nombre véritablement considérable d'utilisateurs, et il est difficile, lors de la configuration des contrôles d'accès, de bien saisir l'impact de diffusion.

La gestion de partage d'informations par "listes d'amis" est également difficile à gérer : besoin de répéter l'opération de partage en fonction du nombre de groupes avec lesquels l'utilisateur souhaite échanger, nécessité parfois de créer d'autres groupes en fonction des informations à partager, etc. Il peut être fastidieux de gérer le partage d'informations par listes.

L'I.H.M. de gestion de confidentialité de *Facebook*, dans laquelle l'utilisateur configure les droits d'accès à ses informations, souffre également de quelques défauts. Bastien et Scapin ont introduit dans [9] huit critères d'évaluation de l'ergonomie d'une I.H.M. Une critique de cette interface est menée à bien dans [8] suivant ces critères. Les points sur lesquels l'I.H.M. de gestion de la confidentialité proposé par *Facebook* posent problème sont les suivants :

9. Source : `http://www.facebook.com/press/info.php?statistics`

- **Concernant les regoupements** : Certains regroupements de paramètres ne paraissent pas justifiés et devraient se situer sur une autre page. De plus, certains paramètres (par exemple les opinions politiques et les opinions religieuses de l'utilisateur) sont groupés et empêchent alors l'utilisateur de modifier leur visibilité un par un. Enfin, la page "Liste de personnes et applications bloquées", dans laquelle l'utilisateur peut définir les membres qu'il désire bloquer, ou dont il désire bloquer les invitations, mériterait d'être plus étoffée.
  Ces quelques regroupements peuvent déboussoler l'utilisateur, qui ne trouve alors pas facilement l'information qu'il recherche.
- **Concernant le retour d'information** : Le retour d'information est vraiment le point qui pose problème. En effet, lorsque l'utilisateur configure la visibilité d'un paramètre, aucune information ne lui est renvoyée. Il peut, par le biais d'un lien lui donnant un aperçu de son profil tel qu'un utilisateur peut voir, et constater alors les changements effectués, mais ce n'est pas clair, et surtout très indirect.
- **Concernant la charge de travail** : L'interface étant divisé en quatre pages, il peut être fastidieux pour un utilisateur de tout paramétrer.

De plus, *Facebook* propose un réglage dit "recommandé", plaçant certains paramètres sur "Tout le monde", notamment les publications. Ces réglages sont ceux définis par défaut lors de la création d'un compte. Pour un utilisateur non expérimenté ou peu intéressé par ces questions de confidentialité, ne pas modifier ces réglages peut s'avérer très dommageable pour lui, car de nombreuses informations le concernant seront alors accessible par l'ensemble des internautes !

La gestion du contrôle d'accès sur *Facebook* est donc difficile, et il est nécessaire d'être un utilisateur aguerri pour en saisir la teneur et l'importance.

1.4 Gestion des risques

Comme on a pu le voir, le partage de données au sein d'un réseau social généraliste peut être très dommageable pour l'utilisateur. Il s'agit alors de s'interroger sur la probabilité qu'un évènement dommageable découlant d'un accès non désiré à certaines informations se produise, les conséquences, ou encore la gravité, le potentiel de nuisance,... Bref, il s'agit de savoir gérer les risques.

[10] est un ouvrage d'avantage centré sur la gestion des risques dans l'industrie, mais peut très bien s'adapter ici à notre travail. Les auteurs distinguent trois manières de gérer le risque, par ordre croissant de coût :

- **La prévention** : Prise de mesures pour limiter l'apparition de l'événement redouté. Cette stratégie est le plus souvent appliquée en premier lieu.
- **L'acceptation** : L'acceptation d'un risque fait suite à une étude de danger. Cette étude permet d'évaluer les dommages pouvant être causés à des personnes exposées si l'événement redouté a lieu.
- **La réduction du risque** : Veille, identification des risques, analyse par la recherche des facteurs de risques et des vulnérabilités, maîtrise des risques par les mesures de prévention et de protection.

Notre travail ici sera donc de réduire au maximum les risques d'exposition non désirée pour un utilisateur sur une plateforme de réseautage social généraliste.

Conclusion

Nous nous sommes donc dans, un premier temps, penchés sur l'ensemble des notions inhérentes au cadre de travail, puis nous avons tâché de comprendre le fonctionnement de la gestion d'accès aux données qu'un utilisateur de la plateforme *Facebook*, notre cas d'étude, peut mener à bien. Nous avons vu qu'il est difficile et fastidieux pour lui de bien gérer la confidentialité des données personnelles qu'il introduit sur le réseau.

Le cadre de notre travail étant situé, entrons à présent dans le vif du sujet, et déterminons les manières de pallier à ces problématiques.

Chapitre 2

Démarche de recherche pour l'élaboration d'une cartographie des domaines inhérents au cadre d'étude

Introduction

Dans le cadre d'un tel travail, il est important de commencer par suivre une démarche de réflexion concernant les besoins inhérents. Nous avons donc commencé par établir une "cartographie" des domaines associés au cadre d'étude afin d'en retirer un ensemble d'outils pouvant aider l'utilisateur dans sa démarche de gestion de la confidentialité de ses données par le biais d'une interface adaptée.

Nous verrons donc ici diverses notions de sociologie (la confiance, l'exposition de soi, la visibilité, et les réseaux sociaux informels), puis quelques apports possibles provenant de la théorie des graphes et du wen sémantique. Enfin, nous étudierons un formalisme de la gestion d'accès aux données d'un réseau social généraliste.

2.1 Aspects sociologiques

De nombreuses notions provenant de la sociologie sont concernées par les réseaux sociaux en ligne. Nous avons donc consulté la littérature sociale afin de saisir différentes notions pouvant être intéressantes à adapter à ce travail : la confiance, l'exposition de soi, la visibilité, et les réseaux sociaux informels.

2.1.1 La confiance

Que ce soit dans le cadre d'un réseau social virtuel ou bien réel, la notion de "confiance" est essentielle et au cœur de nos choix quotidiens. Nous avons donc trouvé intéressant de se pencher sur

la définition sociale de cette notion complexe en vue d'une réflexion sur son implémentation concrète au sein de la gestion de la confidentialité des données pour un réseau social en ligne.

Nous verrons ici comment la définir dans le cadre des réseaux sociaux en ligne, et comment l'utiliser dans le cadre de notre travail.

2.1.1.1 Une notion d'engagement

Le sociologue Sztompka définit en 1999 [11] la confiance en une personne comme étant un engagement vis-à-vis d'une action basée sur la croyance que les actions futures de celle-ci mèneront à une issue favorable. Cette croyance se place donc comme étant une fondation dans le but de prendre par la suite un engagement. Il souligne également que la confiance n'est pas une notion binaire : ce n'est pas soit "j'ai complètement confiance" soit "je n'ai pas du tout confiance". La confiance peut donc être assimilée à la probabilité qu'une personne s'engage à effectuer une action, où action et engagement n'ont pas nécessairement à être signifiant.

Deutsch, en 1962 [12], définit le comportement pour une personne A d'accorder sa confiance à une personne B comme survenant lorsque A se trouve dans une situation dont le dénouement (bon ou mauvais) dépend de B. Il a alors deux possibilités :

1. A laisse B le guider et emprunte le chemin que B lui propose ; dans ce cas A s'en remet à B, et pense alors qu'il l'emmènera vers une issue favorable : A fait un choix de confiance.
2. A ne suit pas B : il ne lui fait pas confiance.

Le terme de confiance se justifie alors, à l'issue de ce choix, par le fait que l'impact psychologique provoqué par la mauvaise issue est plus important que celui provoqué par la bonne.

De plus, la confiance n'est pas absolue, mais contextuelle. Dans le cadre des réseaux sociaux, elle s'applique principalement dans le cas du partage de contenu entre utilisateurs ; plus spécifiquement sur le type de contenu, le thème. Si je peux montrer les photos de mes dernières vacances à ma mère (je lui fait "confiance", je pense que partager ce contenu avec elle m'entrainera vers une issue favorable), je peux ne pas vouloir qu'elle voie les photos de ma dernière soirée "arrosée" (je n'ai pas "confiance" en la réaction qu'elle pourra avoir en reçevant ce contenu).

2.1.1.2 La confiance entre utilisateurs dans les réseaux sociaux en ligne

Jennifer Golbeck mène à bien dans [13] une approche concernant l'adaptation de la notion de confiance aux réseaux sociaux. Pour elle, la connaissance par le système de réseautage social du "degré de confiance" constitue une information importante et utile concernant les contextes sociaux entre utilisateurs ; elle propose alors de recommander ceux avec qui l'utilisateur peut partager et interagir. Il y a deux manières de disposer d'une telle information :

(1) Déduire le niveau de confiance à partir des diverses interactions entre utilisateurs sur le réseau (manière implicite).

(2) Demander à l'utilisateur d'exprimer un niveau de confiance envers un autre (manière explicite).

Pour (1), toute la difficulté réside dans la détermination des facteurs impliquant un certain degré de confiance. Selon moi, tout est question d'intention : dans quel but spécifier un degré de confiance ? Pour pouvoir partager des informations de manière sélective ? Pour recevoir des informations provenant d'utilisateurs particuliers ? De plus, il faut savoir repérer et bien utiliser les facteurs entrant en jeu dans l'établissement d'une certaine confiance. Hélas, la plupart de ces facteurs s'établissent

généralement dans les relations sociales "hors ligne" entre personnes, dans leurs vies quotidiennes : vécu, opinions et actions d'une personne, facteurs psychologiques liés à la durée de l'historique et d'évènements partagés, rumeur, influence des autres, gains perçus en étendant sa confiance, etc. Ces données n'étant donc pas disponibles dans le système, il est difficile de se baser sur des analyses sociologiques liées à l'établissement de la confiance entre membres d'un réseau, celles-ci étant menées essentiellement sur des réseaux sociaux hors ligne. En effet, il va de soi qu'il n'est pas possible de recréer l'ensemble des interactions hors-ligne entre utilisateurs d'une plateforme en ligne. Il serait sans doute intéressant de déduire certaines caractéristiques des relations hors ligne entre utilisateurs à partir des données mises en ligne. Bien que cette possibilité ne soit pas directement traitée dans la littérature scientifique, on notera que de nombreux sociologues et psychologues ont montré que la relation de confiance s'établissait principalement par le biais de la **similarité des profils** : plus deux personnes ont des choses en commun, plus elles ont de chances d'établir une relation de confiance mutuelle concernant d'autres contextes. Les informations principales utilisables seraient donc l'étendue et la nature des relations et interactions mutuelles entre utilisateurs, et le nombre d'items qu'ils partagent (intérêts, photos, opinons, etc.). Le système pourrait alors proposer à un utilisateur un degré de confiance concernant un autre, et s'appuyer sur ce degré pour proposer par la suite, par le biais d'algorithmes tels que les *Collaborative Filtering Algorithms* [14], des recommandations par rapport à leurs interactions, ou encore le partage d'informations qu'ils peuvent mener l'un envers l'autre.

Pour étudier (2), l'auteure s'est penchée sur un réseau social cinématographique : *FilmTrust*. Chaque utilisateur de la communauté est amené à définir explicitement un degré de confiance envers chacun de ses amis, et à noter des films. Les relations de confiance et d'amitié (distinctes dans *FilmTrust* : l'utilisateur a des amis, auquels il attribue un degré de confiance) sont asymétriques : A peut être ami avec B sans pour autant que B soit ami avec A, et A et B ne partagent pas forcément le même degré de confiance. Le but de l'auteure est d'étudier l'évolution de la confiance en fonction des notations des films (et donc de la similarité du profil). Elle obtient des critères de recommandations basés sur des statistiques d'évolution du système, et propres à lui. Elle utilise par la suite ces critères dans le but de prédire le niveau de confiance.

Dans le cadre de notre travail, le réseau social *FilmTrust* n'est pas forcément adapté : il ne s'agit que de goûts cinématographiques, ce qui est très particulier et loin d'un réseau plus général comme *Facebook*. Il est cependant très intéressant de constater que la similarité des profils est effectivement l'un des facteurs principaux dans l'établissement d'une relation de confiance en ligne [13], et c'est donc sur cette notion que nous allons par la suite nous appuyer.

2.1.1.3 CRITIQUE ET APPORT POUR NOTRE ÉTUDE

Pour un réseau social généraliste tel que *Facebook*, la quantité d'informations pouvant être échangées entre utilisateurs est grande. Le travail mené par Golbeck sur le réseau social *FilmTrust* est difficilement adaptable à une plateforme aussi générale. Les données partagées sont bien trop variées et nombreuses pour établir un degré de confiance qui se révèlerait alors trop général.

On pourrait en revanche utiliser la propriété de la similarité des profils en imaginant des recommandations thématiques : au fur et à mesure de l'évolution de son identité numérique au sein du réseau, l'utilisateur pourrait définir différents thèmes ("vacances", "tennis", "musique", "religion", etc.), auxquels seraient associées des listes de confiance. Le système suggèrerait quels utilisateurs placer dans ces listes. Une liste de confiance serait donc constituée d'un ensemble de personnes intéressés par le thème associé. Pour effectuer des recommandations, le système pourrait se baser sur l'ensemble des relations et interactions entre deux utilisateurs (sur *Facebook*, par le biais de la fonctionnalité "Liens d'amitié", sorte d'historique d'un couple d'amis). De plus, à chaque nouvelle

demande d'amitié, l'utilisateur pourrait directement placer le demandeur dans une des listes, ou bien attendre d'avoir suffisamment interagi avec lui pour le faire. Il faudrait cependant faire attention et distinguer domaine d'intérêt commun (accointance potentielle, empathie,...), souhait de se dévoiler sur un thème (degré d'exposition de soi sur ce thème - voir plus loin), en général et vis à vis d'une personne en particulier, etc.

Le système pourrait également proposer des suggestions à l'utilisateur concernant la diffusion de ses informations : sur la base de certains critères, le mettre en garde concernant la diffusion d'albums photos plus "sensibles", ou bien encore de statuts mettant en avant des informations plus confidentielles.

En ce qui concerne le partage envers des utilisateurs qui ne sont pas ses amis, on pourrait imaginer que selon certains critères, le système pourrait élaborer des listes de niveau de confiance. Pour les amis d'amis, les critères en question pourraient être le nombre d'amis en commun, ou encore la confiance accordée à ceux-ci.

Je pense que pour bien intégrer cette notion de confiance, il est nécessaire de bien cibler les usages qu'a l'utilisateur de son identité numérique : autant il peut être intéressant pour lui de manipuler ce terme (surtout le mot et son sens pour lui) afin de mieux comprendre l'importance de l'impact que peut avoir la diffusion de ses données à d'autres, autant il peut se retrouver perturbé par l'abstraction de cette notion.

2.1.2 L'EXPOSITION DE SOI

Même si le terme "exposition" paraît être en opposition avec la notion de "confidentialité", il peut être intéressant de comprendre dans quelles mesures les utilisateurs peuvent se montrer aux autres au sein des réseaux sociaux en ligne : cela permet dans un premier temps, de bien saisir de tels usages, puis dans un deuxième temps de mener une réflexion sur l'élaboration d'outils leur permettant de le faire selon leurs envies.

La confidentialité des données au sein d'un réseau social est d'autant plus importante pour les utilisateurs qui s'y exposent beaucoup. Nous verrons en effet que nombre d'entre eux s'exposent par le biais d'interactions diverses vers les membres du réseau auquel ils appartiennent, souvent sans avoir une réelle connaissance de l'impact de cette diffusion ou des risques engendrés. Une étude a été faite en 2008 sur ce phénomène d'exposition en ligne : l'enquête "Sociogeek" [1] dont les résultats ont été publiés dans [15]. Après une présentation de cette étude, nous en verrons les résultats, et l'analyse qui a pu en être fait, justifiant alors l'intérêt de notre travail. Nous réfléchirons alors à la façon d'aider les utilisateurs souhaitant s'exposer à le faire de la manière la plus contrôlée possible.

2.1.2.1 PRÉSENTATION

"Sociogeek" [15] est une enquête collaborative conduite par *faberNovel*[2], *Orange Labs*[3] et le programme "Identités actives" de la *FING*[4] en 2008. L'objectif, via un jeu-enquête en ligne, est de comprendre comment le développement des réseaux sociaux en ligne a modifié la manière par laquelle nous nous exposons et créons des relations en ligne. C'est donc une enquête sociologique sur

1. `http://sociogeek.admin-mag.com/`
2. Société d'expertise de l'innovation.
3. Laboratoire travaillant sur les nouveaux usages des outils de communication.
4. Fondation pour l'Internet Nouvelle Génération ayant pour mission de Repérer, stimuler, et valoriser l'innovation dans les services et usages du numérique et des réseaux.

la mesure de l'évolution de nos comportements à l'heure du web 2.0. Que montre-t-on ? Que cache-t-on ? Avec qui ? Peut-on réellement parler d' "amis" virtuels ?,... sont entre autres les interrogations au cœur desquelles cette enquête s'est développée.

Le jeu se déroule en trois étapes :

1. **Photos :** 20 séries thématiques de 4 photos sur différentes dimensions de l'exposition de soi (sexe, boutons, nudité, grossesse, vacances, maladie, alcool,...) sont proposées. Les photos de chaque série sont classées de l'exposition la plus "modeste" à la plus sévère ; le sujet est amené à cliquer sur les photos qu'il serait prêt à publier, sous la contrainte de devoir publier également les photos précédentes. Une fois les 20 séries effectuées, le sujet a à sa disposition un récapitulatif des photos sélectionnées parmi lesquelles il doit en éliminer 5 et en garder 3.
2. **Questionnaire :** Les sujets sont amenés à répondre à diverses questions telles que leur caractéristiques sociales, ou l'usage qu'ils font des réseaux sociaux.
3. **Choisir des amis :** Le sujet a trois séries de six utilisateurs souhaitant devenir ami avec lui sur un hypothétique réseau social. Pour chaque série, il est amené à dévoiler un certain type d'information les concernant (photo de profil, posts, âge, diplômes, opinions politiques ou religieuses, etc.), puis il élimine un "prétendant". Il répète cette opération jusqu'à ce qu'il n'en reste plus qu'un seul, qu'il accepte alors comme ami. A la fin des trois séries, il classe les élus sur un "podium".

2.1.2.2 Échantillon

7580 internautes ont répondu à cette enquête. Ils ont en moyenne 28 ans, de tous partis politiques et catégories socio-professionnelles (il y a cependant une majorité de cadres et chômeurs, d'avantage de gauche et centre). Parmi eux, 96% utilisent Internet plusieurs fois par jour et sont inscrits sur divers réseaux sociaux tels que *Facebook*, *Copains d'avant*, *Myspace*, *LinkedIn*, *Dailymotion*, *flickr*, *Viadeo*, ou des blogs.

2.1.2.3 Résultats intéressants

Certains résultats de cette enquête justifient l'intérêt de notre thème de travail. Parmi ceux-ci :

- 38% des répondants acceptent presque automatiquement en ami d'autres utilisateurs (23% d'entre eux si ils ont des amis en commun, 6% simplement par la photo, et 7% acceptent toujours).
- 16% des répondants utilisent les réseaux sociaux pour des rencontres.
- Ceux qui s'exposent le plus sont les jeunes (moins de 20 ans).

Il y a donc d'une part, énormément de contacts et d'interactions qui se créent entre des utilisateurs qui ne se connaissent pas ; ils peuvent alors potentiellement révéler l'un à l'autre des informations de nature sensible, intime, ou confidentielle, sans le vouloir (il est facile d'oublier une personne dans une grande liste d'amis et donc lui révéler des choses que l'on ne pense révéler qu'à ses "vrais" amis).

D'autre part, ce sont les jeunes qui s'exposent le plus : s'ils se montrent sans contrôler vraiment le niveau de diffusion de leur exposition, ils pourraient le regretter par la suite (dans leur future vie professionnelle par exemple).

2.1.2.4 Analyse

Les différents sociologues responsables de cette enquête ont alors pu tirer diverses conclusions à partir des résultats obtenus. Ils ont notamment remarqué qu'il n'y avait pas de lien entre la fréquence d'usage et le taux d'exposition de soi : de nombreux utilisateurs affirment passer beaucoup de temps sur certains réseaux sociaux sans pour autant se dévoiler. Ils ont en revanche dégagé un lien fort entre le nombre d'amis et le niveau d'exposition ; il faut donc s'exposer pour être influent [5] au sein du réseau. Ils ont également constaté que l'identité numérique d'un utilisateur était influencé par sa véritable identité : ceux qui ont une forte implication dans leur vie sociale sont plus susceptibles de divulguer des informations en ligne ; on pourrait déplorer ici, le raisonnement fait dans l'absolu, sans prendre en compte le thème, le contexte de l'exposition. Il est cependant à noter qu'un utilisateur peut se révéler être très "actif" sur la plateforme (de nombreuses interactions avec les autres, et un profil toujours en évolution) sans pour autant révéler beaucoup d'informations personnelles sensibles. En ce qui concerne les critères utilisés par la majorité des sujets pour choisir d'accepter un nouvel ami, on remarque que 52% d'entre eux font apparaître en priorité la photo du profil. Viennent par la suite le descriptif bref de la personne ("about me"), les messages échangés avec les amis, et l'âge.

Cette analyse apporte à nouveau une motivation et des pistes : le lien entre exposition et nombre d'amis est éloquent. En effet, si l'on souhaite se faire beaucoup d'amis, alors il est nécessaire de se dévoiler, mais on remarque aussi que ceux qui se dévoilent le plus ont tendance à accepter des amis bien plus facilement (30% d'entre eux acceptent automatiquement !). Pour qu'ils puissent au mieux protéger leur vie privée, et leur garantir alors une exposition contrôlée, il est nécessaire de mener à bien une prévention pertinente, et de leur proposer un moyen simple, pratique et facilement accessible de le faire.

De cette analyse, les sociologues ont conclu qu'il existait cinq formes d'exposition de soi différentes :

- **Modeste (18.9%)** : Publication d'informations peu sensibles, où il est parfois difficile d'identifier l'utilisateur. Concerne d'avantage les femmes, les personnes les plus âgées, ou avec un fort niveau d'éducation. Généralement les personnes incluses dans cette catégorie sont ceux ayant le moins d'amis en ligne (71 en moyenne).
- **L'exposition de soi traditionnelle (24.1%)** : Correspond aux formes habituelles de mise en scène de sa vie privée : la photo de famille, de vacances, de mariages ou de supporter sportif.
- **L'impudeur corporelle (20%)** : Correspond aux formes d'exposition de soi caractérisée par la nudité corporelle, l'intimité sexuelle et la vie amoureuse. Elle rassemble les photos de baisers, de nudité, d'acte sexuel et de grossesse [6].
- **L'exhib' (24.2%)** : Correspond aux formes d'expression de soi dans lesquelles les personnes se montrent dans des poses théâtrales, marquées et très expressives dans un ensemble varié de contextes :en mangeant, décontracté au travail, en colère, ivre, dansant, manifestant ou affichant le désordre de leur lit.
- **Le trash (12.8%)** : Correspond à des formes d'exposition de soi outrancières lorsque les participants exhibent des images "négatives" d'eux-mêmes pleurant, malades ou exhibant des disgrâces corporelles comme des boutons. Elle caractérise la nouvelle culture d'expressivité juvénile, où il s'agit de montrer que l'on est "cool"...

5. Le terme *influence* est ici à prendre au sens de la quantité d'"amis" au sein du réseau : plus l'on a d'amis et plus l'on est influent.

6. Cette forme d'exposition reste tout de même limitée sur les plateformes telles que *Facebook* par exemple : il existe en effet une charte qui interdit les photos "inconvenantes", et pouvant entraîner la fermeture du compte. Elle n'est cependant que limitée car les sanctions ne s'appliquant qu'après dénonciation de la part d'un autre utilisateur.

2.1.2.5 CRITIQUE ET APPORT POUR NOTRE ÉTUDE

Cette enquête sociologique est en lien direct avec le sujet du stage : l'analyse des résultats par des sociologues apporte de nombreuses informations concernant les usages qu'ont les utilisateurs de leur exposition au sein des réseaux sociaux. Nombre d'entre eux ont une réelle volonté de s'exposer.

D'une part, c'est un choix qui nécessite un réel contrôle de diffusion. En effet si l'exposition de soi tend à être prononcée, il est important de savoir vers qui on désire le faire : quelques amis proches ? Tous nos amis ? Leurs amis aussi ? Tout le monde ? Le manque de maîtrise de cette exposition peut entraîner de graves conséquences. D'autant plus que les analystes ont révélé que ceux qui s'exhibent le plus sont ceux qui accumulent le plus d'amis. Il est donc important de miser sur la prévention : avant d'accepter quelqu'un en ami, par exemple, insister sur son placement dans une liste d'amis, ou encore si la personne a peu d'amis en commun avec l'utilisateur (en terme de rapport avec le nombre d'amis total bien sûr), lui proposer un avertissement quant à l'implication du partage.

D'autre part, il est important de considérer la nature des informations alors exposées : certaines informations divulguées peuvent en effet se révéler plus nuisibles que d'autres pour le diffuseur. On pourrait introduire la notion de "nuisance potentielle", par exemple, comme critère à utiliser.

Il pourrait également être intéressant de prendre en compte les caractéristiques personnelles que les utilisateurs ont tendance à regarder le plus avant d'accepter une demande d'amitié (photo de profil, "about me", messages aux amis, et âge) et d'insister d'avantage sur le contenu de ces informations et leurs diffusion.

Finalement, ce travail sociologique peut être relié au précédent concernant la notion de confiance ; on pourrait en effet parler de confiance pour l'exhibition. L'exposition de soi pourrait alors être un thème supplémentaire dans les recommandations de diffusion.

Cette étude nous renseigne donc sur la nécessité de permettre aux utilisateurs de contrôler leur exposition au sein des plateformes. Dans le cadre de ce travail, il faudrait donc guider l'utilisateur dans sa démarche d'exposition afin qu'il puisse le faire de façon contrôlée (par le biais de suggestions concernant des listes d'amis adaptés).

2.1.3 LA VISIBILITÉ

Qu'ils décident de montrer un maximum d'informations ou au contraire de très peu se dévoiler, que leur identité numérique soit différente de leur véritable identité, ou au contraire calquée sur elle, les utilisateurs des réseaux sociaux en ligne sont les créateurs d'une multitude d'informations les concernant. Ces informations définissent une personnalité virtuelle qui les caractérise. Les fonctions de cette identité numérique dépend de la plateforme de réseautage social, et de ce que l'utilisateur veut montrer de lui.

Il nous a donc paru intéressant, dans le cadre de notre travail, de pouvoir au mieux caractériser le type d'identité numérique des utilisateurs afin de mieux les guider dans l'élaboration de la confidentialité associée à cette identité. Cette caractérisation est l'objet de cette partie.

Dominique Cardon, sociologue au laboratoire "Sense" d'*Orange Labs*, a proposé dans [16] en 2008 une étude concernant le design de la visibilité dans le web 2.0. Que montre-t-on de soi aux autres ? Comment les liens tissés sont-ils rendus visibles ? Comment retrouver des personnes et en découvrir de nouvelles ? Quelle identité nous fabriquons-nous en ligne ?... Pour tenter de répondre à ce genre d'interrogations, il propose "une typologie des plateformes relationnelles du web 2.0 qui s'organise autour des différentes dimensions de l'identité numérique et du type de visibilité que chaque plateforme confère au profil de ses membres."

Nous allons donc voir ici les formes de visibilité pouvant caractériser les utilisateurs, mais aussi les plateformes. Nous verrons alors que l'enjeu de la visibilité n'est pas la même selon la nature de la plateforme considérée et sa forme.

2.1.3.1 Décomposition de l'identité numérique

L'auteur propose de décliner l'identité numérique autour de deux "tensions" (voir Figure 2.1 de [16]) :

- **Processus de subjectivation** (en abscisse) : Extériorisation de soi. Tension entre les signes se référant à ce que la personne est dans son être (sexe, age,...) et ceux se référant à ce que la personne fait (projet, œuvres, productions,...).
- **Processus de simulation** (en ordonnée) : tension entre les traits se référant à la personne dans sa vie réelle et ceux qui renvoient à une projection, une simulation de soi.

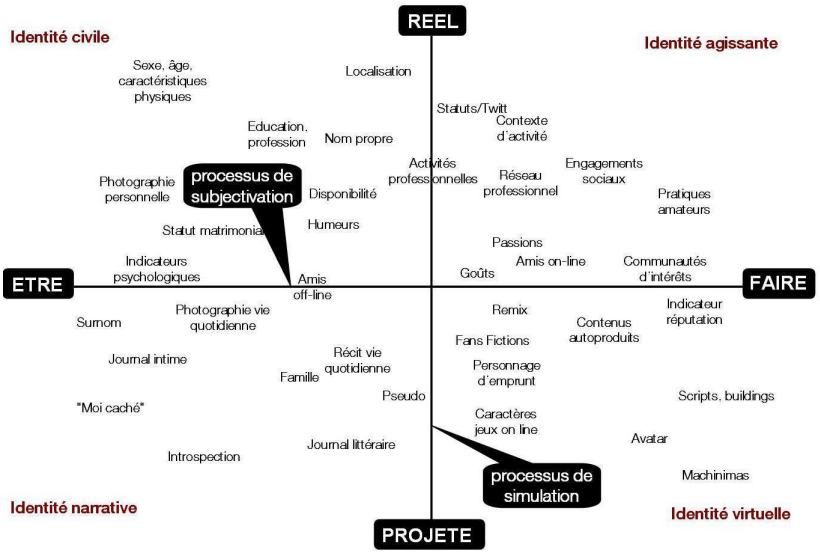


FIGURE 2.1 – Décomposition de l'identité numérique

Ainsi, les utilisateurs situés dans la partie haute du repère (Identité civile - Identité agissante) calquent leur identité numérique sur leur identité réelle. Il montrent aux autres ce qu'ils font (Identité agissante), et ce qu'ils sont (Identité civile), ou du moins ce qu'ils croient être (subjectivation).

Les utilisateurs situés dans la partie basse du repère (Identité narrative - Identité virtuelle) quant à eux, projettent une partie ou une potentialité d'eux-mêmes. Ils expriment ce qu'ils sont vraiment (Identité narrative), ou alors se créent une personnalité virtuelle, au sens premier du terme, devenant alors les concepteurs de leur identité, de l'environnement, ou des actions et des événements auxquels ils prennent part (Identité virtuelle).

2.1.3.2 Formats de visibilité

Cardon propose alors de poser sur ce schéma les cinq formes de visibilité qu'il a pu dénombrer dans les différentes plateformes sociales (voir Figure 2.2 de [16]) :

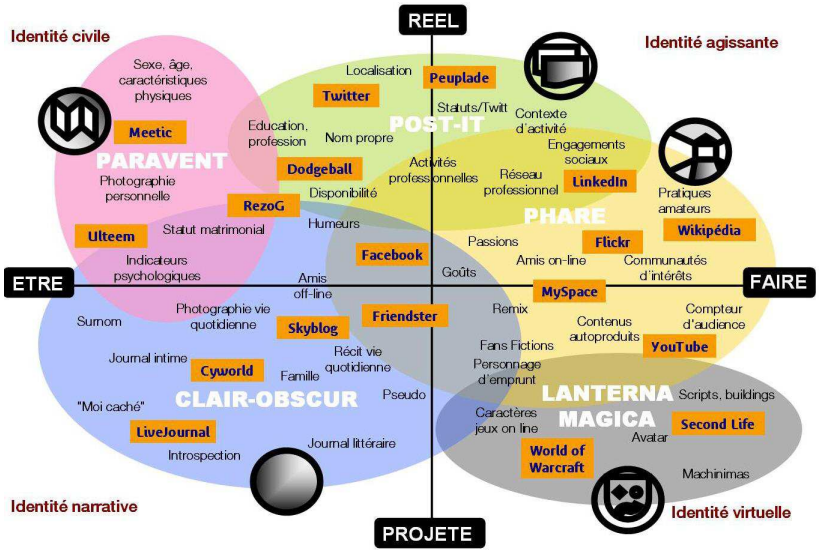


FIGURE 2.2 – Formats de visibilité

- **Le Paravent** : Dans le "monde" du *Paravent*, les utilisateurs ne sont visibles qu'à travers un moteur de recherche spécifique. Ils sont cachés derrière des catégories, et ne se dévoilent que lors d'une interaction décidée. Le réseau social *Meetic* en est l'un des exemples.
- **Le Clair-obscur** : Les utilisateurs des réseaux sociaux du *Clair-obscur* rendent généralement visible intimité, quotidien, et vie sociale, mais à un réseau de proches ; l'accès est très limité aux autres utilisateurs. Ils peuvent cependant s'ils le souhaitent s'ouvrir à la "nébuleuse" d'amis des amis ou à des réseaux proches. *Skyblog*, *Friendster*, *Facebook* en sont des exemples.
- **Le Phare** : Le principe des plateformes du *Phare* est de rendre visible beaucoup de choses concernant son identité réelle, ou "un volet particulier" de celle-ci, pour être reconnu dans un domaine spécifique (exemple : des photos sous-marines, des activités musicales, des productions vidéos,...). Le profil est facilement accessible, afin de favoriser les échanges et les contacts avec des inconnus. Dans le monde du Phare, les utilisateurs créent des collectifs sur des contenus partagés et sont à la recherche d'audience, de connectivité, ou encore de réputation. Parmi ces plateformes, on trouve *Flickr*, *Myspace*, ou *YouTube*.
- **Le Post-it** : On rend visible sa disponibilité et sa présence par de nombreux indices contextuels et on réserve l'accès à un cercle relationnel restreint. Le couplage territoire– temps y est très fort (géolocalisation, planifier des rencontres,...). Twitter, par exemple, appartient au monde du *Post-It*.
- **La Lanterna magica** : Les utilisateurs sont représentés par un avatar et découplent leur identité. Les principaux réseaux sociaux de la *Lanterna Magica* sont les jeux en ligne (*World of Warcraft*, *Second Life*,...).

2.1.3.3 Enjeu de la visibilité

L'avènement des plateformes sociales en ligne offre aux utilisateurs de toutes nouvelles formes de visibilité, peu compatibles entre elles (Figure 2.3 de [16]).

Sur certaines plateformes, l'enjeu est de ne pas beaucoup se dévoiler pour se découvrir par la suite lors d'une rencontre réelle (**Se cacher, se voir**). Sur certaines autres, on peut s'élaborer une toute nouvelle identité, virtuelle et potentiellement très différente de son identité réelle ; sur ces plateformes, aucune rencontre réelle : c'est cette nouvelle identité qui définit entièrement l'utilisateur (**se voir caché**). L'utilisateur peut également s'exposer de manière très contrôlable et ne montrer que certains aspects de lui-même, floutant parfois le pire pour ne montrer que le meilleur, ou encore se rendre peu reconnaissable, peu retrouvable (**montrer caché**). A contrario, on peut ne rien cacher et s'exposer entièrement afin de s'assurer le plus de notoriété possible (**tout montrer, tout voir**).

Chaque plateforme offre donc à l'utilisateur une visibilité propre. Cette diversité lui permet donc de s'élaborer l'identité numérique qu'il souhaite, et d'avoir ainsi un contrôle adapté de sa distance à soi.

On peut cependant s'interroger sur "l'agrégation" des plateformes dans *123People*[7] ou *Google*[8], qui font en quelques sorte "tomber les barrières" entre les réseau sociaux. Bien sûr, certaines plateformes, comme *Facebook*, permettent au utilisateurs de disparaitre de ce genre de moteurs de recherche, mais pas toutes.

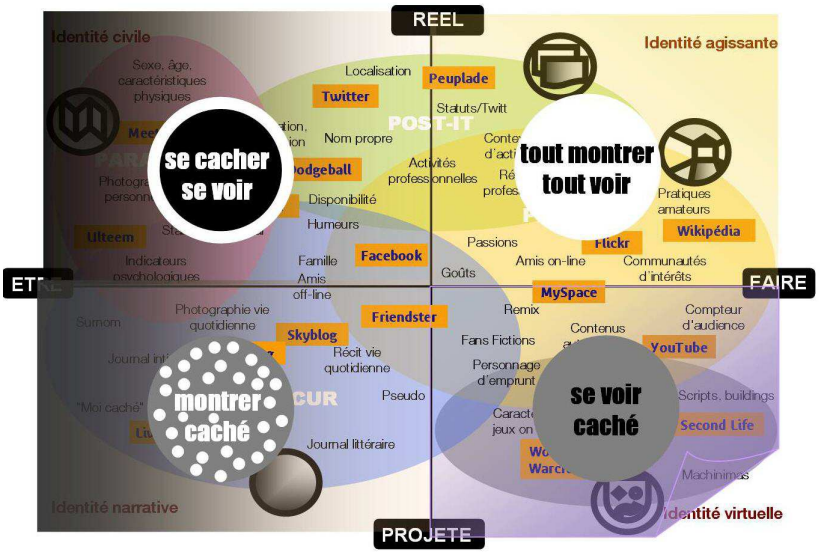


FIGURE 2.3 – Quatre formes de visibilité

De par leurs différences de nature, les communautés des réseaux ne se mélangent donc que très peu. Cependant, de plus en plus de réseaux sociaux réels se confondent dans un seul et même réseau social réel ; c'est le cas de *Facebook* par exemple, où l'utilisateur mélange sous son véritable nom, l'ensemble de ses connaissances. Cette pratique introduit alors un tout nouveau risque identitaire : ce

7. http://www.123people.fr/
8. http://www.google.com/

mélange nouveau de tous les milieux de son entourage derrière une visibilité contrôlée nécessite des compétences sociales et relationnelles spécifiques et inégalement distribuées.

2.1.3.4 FORME DES RÉSEAUX SOCIAUX

Selon la plateforme sociale, il est important de savoir différencier leur taille et leur forme (voir Figure 2.4 de [16]), toutes deux dépendantes à la fois de la nature de la plateforme et des intentions de l'utilisateur. Ainsi , dans le monde du Paravent, le réseau d'amis n'est jamais montré. Dans celui du Phare, le réseau est très étendu mais généralement non fortement connecté, contrairement au Clair-obscur.

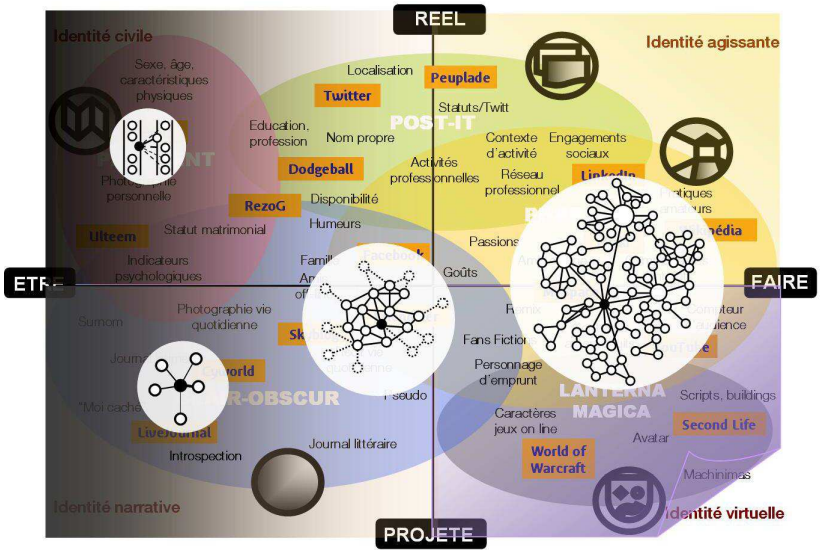


FIGURE 2.4 – La taille et la forme des réseaux sociaux dépendent de leurs natures

Bien sûr, sur de nombreuses plateformes, la constitution d'un réseau social dépend également du niveau de visibilité du profil de l'utilisateur, et de l'utilisation qu'il en fait. Mais ce contrôle de visibilité n'est pas forcément bien utilisé par tous, et on trouve de plus en plus souvent des réseaux constitués d'un entremêlement d'amis réels et d'amis "utiles", mettant en avant une logique opportuniste et calculatrice en décalage avec les attentes initiales des utilisateurs. La nécessité d'organiser les zones de visibilité et de bien trier les contacts est donc essentielle.

2.1.3.5 CRITIQUE ET APPORT POUR NOTRE ÉTUDE

Ce travail propose une classification des réseaux sociaux et de leurs utilisateurs sur un repère, basé sur des notions sociologiques et psychologiques solides. Cela constitue donc un très bon appui concernant notre travail. En effet, on pourrait réfléchir aux outils permettant de placer les différents utilisateurs sur ce repère, et adapter alors l'interface de paramétrage de leur confidentialité à leurs usages sur la plateforme. On pourrait alors imaginer une interface adaptative dépendant de l'utilisateur, de son identité numérique et du rôle qu'il veut jouer grâce à elle.

2.1.4 RÉSEAUX SOCIAUX INFORMELS

La notion de "réseaux sociaux informels" nous a parru intéressante à étudier, dans la mesure où cette notion est essentielle au cœur de notre quotidien.

En effet, Arshad explique dans [17], qu'un réseau social informel est un concept englobant la famille, les amis, les voisins, les communautés d'intérêts, etc. pour une personne. C'est donc un ensemble de liens entre particuliers, dont chaque membre a une réelle importance pour l'ensemble réseau. Un réseau social informel dispose de trois dimensions caractéristiques :

- **Structurelle :** taille (nombre de personnes) et densité (degré de connaissance mutuelle) des rapports.
- **Interactionnelle :** nature des relations (solides ou faibles) comprenant la durabilité, l'intensité, la dispersion (facilité à établir le contact), et la réciprocité.
- **Fonctionnelle :** type de soutien. Instrumental (temps, argent, main d'oeuvre,...), émotionnel (souci, compréhension,...), informationnel, et soutien de valorisation (estime, affirmation,...).

On pourrait donc réfléchir sur la nature de la relation de l'utilisateur avec ses amis, et le degré de soutien qu'ils lui apportent (scinder le réseau social en différents réseaux sociaux informels), afin de lui proposer des listes d'amis adaptés à ces réseaux ayant pour lui un véritable sens social.

Conclusion et perspectives

Fort de ces connaissances sociologiques, nous pouvons à présent mieux cerner certains enjeux, rélféchir à leur apport dans le cadre de notre travail, et à leur implémentation concrète au sein d'une interface de paramétrage de confidentialité d'accès aux données.

Tout d'abord, nous avons vu que la notion confiance en ligne est principalement basée sur la similarité des profils. Il serait donc intéressant de classer thématiquement les données diffusées. Les thèmes peuvent être abordés de deux façons : par une classification (élaborée à priori par les concepteurs), partagée par tous les utilisateurs du système, ou par le biais de tags, choisis indépendamment par chaque utilisateur, leur permettant ainsi de réserver la diffusion de ces thèmes à certaines listes d'amis. Ces deux approches ont chacune leurs partisans (et leurs avantages), mais nous préfèrerons ici la seconde, permettant à l'utilisateur de disposer de plus de liberté concernant la classification de ses données.

Nous avons vu par la suite que de nombreux utilisateurs s'exposent sur les réseaux sociaux. Afin de leur garantir une exposition contrôlée, il faudra diriger l'interface vers la prévention et l'encouragement à créer et utiliser des listes d'amis adaptées à leur exposition. On pourrait également imaginer de suggérer aux utilisateurs de créer des tags associés à leur exposition de soi.

L'identité numérique que se crée un utilisateur peut revêtir de nombreuses formes et de nombreuses fonctions selon la volonté de celui-ci et le réseau social dans lequel il évolue. Nous pouvons désormais classer ces identités dans le repère de Cardon ; ainsi, nous pourrions adapter l'interface à ces usages afin de lui proposer une confidentialité adaptée à son identité numérique.

Les listes d'amis devraient également prendre en compte différents réseaux informels de l'utilisateur afin de l'aider dans la diffusion de ces données.

2.2 Analyse des réseaux sociaux

Les réseaux sociaux sont généralement représentés par des graphes. Il nous a donc parru important de consulter la littérature concernant l'analyse mathématique de tels graphes et disposer

ainsi de divers outils concernant la position de l'utilisateur au sein du réseau.

Guillaume Erétéo propose en 2009 dans [18] une analyse des réseaux sociaux en s'aidant de la théorie des graphes et en réfléchissant sur les apports qui pourraient provenir du web sémantique. Nous verrons donc ici différentes notions provenant de la théorie des graphes, adaptés aux graphes sociaux, et nous verrons que de nombreuses informations peuvent être utiles à l'utilisateur si on le positionne dans son graphe social à l'aide de ces différentes notions. Puis nous constaterons les apports pouvant provenir du web sémantique.

2.2.1 Théorie des graphes

Un réseau social est usuellement représenté par un graphe non orienté dont les sommets sont les membres et les arêtes les liens entre eux. Erétéo pense que la modification de cette représentation par le biais de la théorie des graphe pourrait s'avérer utile. Le graphe pourrait en effet être vu comme étant :

- **Orienté :** on pourrait représenter des relations non symétriques, par exemple la confiance (riche en sémantique), et introduire alors des notions nouvelles comme le prestige.
- **Pondéré :** les poids sur les arêtes pourraient représenter l'intensité des relations par exemple.
- **Étiquetés :** représenter différents types de relation (amis, famille, collègues,...) ou les domaine, thématiques, et points de vue concernés.
- **Multipartites :** pour modéliser les interactions utilisateur-ressources.

On manipule par la suite ces graphes via des matrices (d'incidence, d'adjacence, de Laplace,...).

2.2.1.1 Indicateurs

Pour s'aider dans l'analyse, on dispose de différents indicateurs :

- **La densité :** C'est la quantité de liens au sein d'un réseau (la cohésion). Elle est calculée relativement au nombre maximal de lignes que peut contenir la matrice représentative d'un graphe. On peut alors mener deux types d'analyse : une analyse égocentrée (influence de l'individu sur le sous-graphe auquel il appartient), ou sociocentrée (densité sur l'ensemble du graphe, et contraintes du réseau sur ses membres).

- **La centralité :** Il y a différents types de centralité :
  - de degré : les nœuds de degré les plus hauts sont centraux et représentent donc des nœuds de grand intérêt, très visibles, et doté d'un potentiel élevé à faire circuler l'information. On peut alors repérer les points dominants et les centres d'intérêt du réseau, ou encore mesurer l'influence de l'activité des nœuds (pour un graphe orienté).
  - D'intermédiarité : c'est la capacité d'un nœud à servir d'intermédiaire. Plus un nœud est intermédiaire, plus le réseau dépend de lui, et plus il a de pouvoir. La centralité d'intermédiarité constitue un indice concernant la dépendance et l'efficacité du réseau par rapport à certains nœuds. Cette information est d'autant plus intéressante si l'on considère un graphe orienté. Si deux personnes au sein d'un même groupe n'ont pas de relation (on parle alors de "trou structural"), les personnes bénéficiant d'une centralité d'intermédiarité sont alors avantagés : ils disposent en effet d'un bénéfice informationnel (accès rapide aux informations non redondantes), et ont un avantage sur le contrôle de l'information.

  - De proximité : plus les chemins reliant un nœud aux autres est court et plus il est central (car la connexion avec les autres est plus rapide). Ce critère permet de mesurer la performance des communications dans le réseau (circulation d'informations). Si le graphe est orienté, on peut évaluer la capacité d'un nœud à atteindre, ou à être atteint, par un autre.

  Pour un sommet donné, il peut être également intéressant de mesurer la centralité de ses adjacents (cela lui donne un avantage). On peut également déterminer l'influence d'un nœud sur son voisinage (centralité égocentrée).

- **Détection de communautés :** Liée à la notion de cohésion dans un groupe, la détection de communautés permet de repérer les communautés d'intérêt en ciblant les trous structuraux. On peut alors identifier par exemple les sommets intermédiaires de groupes fortement connectés, ou encore déterminer la répartition des acteurs et des activités par le biais de la contrainte du réseau (mesure de la redondance des contacts d'une personne) : plus les contacts d'une personne sont reliés entre eux, et plus son comportement est contraint. On peut alors étudier des sous-graphes tels que les composants (ensemble de nœuds connectés sans lien extérieur), les cliques (sous-graphe complet), ou les cycles (chemin revenant à son point d'origine).
  Il est également à noter que la détection automatique de communautés peut aider à la constitution des "listes d'amis". En effet, il peut être fastidieux pour un utilisateur de créer "manuellement" des groupes d'amis sur la plateforme (notamment si il a beaucoup d'amis). Or, la constution de telles listes est très importante dans la gestion de la confidentialité de ses données en ligne. Il est alors possible d'utiliser la détection de communauté de manière automatique afin de proposer à l'utilisateur des listes toutes faites. L'équipe DNET [9] de l'École Normale Supérieure de Lyon a par exemple créé une application *Facebook* baptisée "Fellows" [10] analysant la liste d'amis de l'utilisateur et lui proposant automatiquement des groupes d'amis.

L'étude de tels indicateurs peut alors se réveler fort utile dans le cadre de notre travail. On peut par exemple repérer les utilisateurs les plus influents, la cohésion de certaines parties du réseau (en utilisant la densité), les utilisateurs dont le potentiel à faire circuler l'information est plus grand (par le biais de la centralité), ou encore détecter des communautés, pour aider les utilisateurs à se constituer des groupes d'amis. On peut alors par la suite leur faire des suggestions (suggestions de groupes d'amis, prendre garde à la diffusion large de leurs informations si ils sont plus influents, etc.) de gestion de la confidentialité en se basant sur la topologie du réseau.

2.2.1.2 Structure

Newman dans [19], Watts et Strogatz dans [20], et Barabasi dans [21] définissent quelques propriétés caractéristiques d'un réseau social numérique. Principalement, on constate un effet de "petit monde", dans lequel on peut théoriquement relier toute personne à une autre par un chemin de courte distance (de l'ordre log(n), où n est la taille du réseau). Il est cependant, à mon sens, important de noter que cette caractéristique n'est pas toujours vraie en ligne. On peut en effet

9. conduit des recherches théoriques et expérimentales sur les réseaux sociaux afin de mieux appréhender leur structuration et la dynamique des processus de diffusion d'information en leur sein. L'équipe DNET est affiliée au CNRS et à l'INRIA.

10. `http://fellows-exp.com/`

trouver des ensembles d'utilisateurs complètement isolés des autres (l'exemple le plus frappant est celui de deux utilisateurs reliés l'un à l'autre, mais étant déconnectés des autres). Il est donc parfois dangereux d'extrapoler aux réseaux sociaux en ligne des enseignements sociologiques issus des réseaux sociaux réels.

On peut également vérifier au sein d'un réseau social, la tendance de l'homme à se socialiser (tendance au "clustering" et à une structure communautaire), le plus souvent par le biais de la transitivité : les amis de mes amis ont plus de chance de devenir mes amis. On peut aussi remarquer que le nombre de personnes ayant peu d'amis est bien plus élevé que le nombre de personnes en ayant beaucoup (par rapport à la moyenne du nombre d'amis des utilisateurs sur l'ensemble du réseau).

2.2.1.3 ALGORITHMES

Il est possible de calculer ou valider ces indicateurs et ces notions par le biais d'algorithmes [22]. On peut alors détecter les communautés via des algorithmes de clustering (hiérarchiques – agglomératifs ou séparatifs - ou heuristiques), puis valider le découpage par la mesure de 3 indices : la silhouette (mesure des propriétés d'isolation et d'hétérogénéité des clusters), l'indice de Dunn et Davies-Bouldin (calcul du nombre de clusters denses et séparés), et la modularité (différence entre la part d'arêtes intra-communautaires du réseau analysé et celle avec une répartition aléatoire des arêtes). On peut de même calculer et vérifier la centralité, ou bien la densité.

Il est donc intéressant de savoir que l'on peut s'appuyer sur des algorithmes existants, afin de pouvoir mesurer de tels indicateurs par la suite.

2.2.2 WEB SÉMANTIQUE

Le web sémantique permet aux machines de classifier automatiquement, puis d'exploiter les ressources du web de manière interopérable. Il pourrait alors être intéressant d'appliquer des propriétés du web sémantique aux réseaux sociaux en ligne afin de faire ressortir du sens à partir des informations partagées. Mika dans [23] distingue trois catégories de réseaux sociaux sur le web :

- Les réseaux sociaux inférés par le web mining (méthodes d'extractions de RS à partir de réseaux d'amis, cooccurrences de noms sur les pages web, mesure de la force de la relation, ...)
- Discussions électroniques (mail, chat, forum,...)
- Applications sociales du web 2.0 (outils de publications – wiki, blogs,... – réseaux sociaux, sites de partages, jeux collaboratifs,...)

Avec le web 2.0, de nouveaux usages du web sont apparus. Parmi eux, le "tagging" (identification des ressources par mots-clés) pourrait constituer un outil très utile pour les réseaux sociaux en ligne. On pourrait en effet réfléchir à l'introduction de "social tagging" au sein de réseaux sociaux généralistes tels que *Facebook* : une méthode de classification collaborative de ressources annotés par des tags, modélisée par des graphes tripartites dont les sommets seraient les utilisateurs, les tags, et les ressources, et les arêtes l'association d'un tag à une ressource par un acteur.

Ainsi, on pourrait proposer à l'utilisateur de classer de manière thématique les données diffusées par le biais du "tagging", puis de gérer la confidentialité associée à ces tags.

On peut alors réfléchir à la représentation sémantique d'un réseau social. L'idée serait de savoir représenter le réseau social, l'analyser, puis inférer sur les graphes sociaux par le biais d'ontologies dédiées. Erétéo propose pour cela la modélisation suivante :

- Ontologie **FOAF** (Friend Of A Friend) : décrit les personnes, les liens entre elles, ce qu'elles créent et ce qu'elles font.
- Des concepts associés aux profils (*family_name*, *nickname*, *interest*,...).
- Une propriété **Knows** pour relier les profils entre eux
- Des classes pour modéliser les usages des utilisateurs : représentation des ressources manipulées et les relations plus complexes.
- Ontologie **RELATIONSHIP** : étend *Knows* (amical, professionnel, familial,...)
- Notion de **Folksonomie** (classification des ressources par le "social tagging")

Cette modélisation s'adapte bien à notre cadre d'étude, étant donné que la gestion de la vie privée y est prise en compte (modélisation des usages, des liens entre les utilisateurs, gestion des amis par "types",...). Pour concrètement l'utiliser, il suffit d'utiliser les *microformats* (attributs HTML augmentant la sémantique).

On pourrait aussi introduire une ontologie représentant les communautés de pratiques. Pour cela, Tifous et al. propose dans [24] trois constituants principaux :

- **Engagement mutuel** (réciprocité, confiance, ouverture)
- **Entreprise commune** (pas d'objectif ou de but, simple ensemble de processus pour constituer des produits communs)
- **Répertoire partagé** (ensemble de ressources communes nécessaires à la vie en communauté)

CRITIQUE ET APPORT POUR NOTRE ÉTUDE

Ce travail présenté par Erétéo [18] offre une perspective d'analyse des réseaux sociaux par le biais de la théorie des graphes et des clés concernant l'application du web sémantique à ceux-ci. Il pourrait donc être intéressant de faire ressortir quelques notions liées à la théorie des graphes (centralité, prestige,...) et de les utiliser pour guider les choix de diffusion des informations. Des méthodes de "social tagging" pourraient également être intéressantes à manipuler : les utilisateurs pourraient "taguer" une ressource par rapport à son contenu et l'on pourrait alors lui proposer des paramètres de visibilité associés à ces tags.

Par le biais de l'analyse des graphes sociaux des communautés, on pourrait ainsi proposer des paramètres de confidentialité alternatifs, personnalisés, et évolutifs que nous détaillerons plus loin.

2.3 Un formalisme du contrôle d'accès opéré par *Facebook*

Pour appliquer toutes ces notions, il serait intéressant de disposer d'un cadre général formel de départ. Nous avons donc consulté la littérature concernant la formalisation du contrôle d'accès aux données dans les réseaux sociaux en ligne généralistes.

Fong, Anwar et Zha proposent dans [25] un modèle de contrôle d'accès formalisant et généralisant le mécanisme de gestion de confidentialité pour un réseau social généraliste de type *Facebook*. Ce modèle se veut général, et de telle sorte que *Facebook* en soit une instance. Il serait ainsi possible d'instancier d'autres réseaux sociaux à partir de ce modèle ; chacune de ces instances disposerait de mécanismes de contrôles d'accès différents.

Ce travail est, à la connaissance des auteurs en juin 2010, le seul traitant de cette problématique de modélisation des contrôles d'accès dans les réseaux sociaux. Il est représentatif de ce que l'on peut faire en terme de formalisation des interactions dans un réseau social.

Nous verrons dans un premier temps comment caractériser la gestion du contrôle d'accès de la plateforme *Facebook*, puis comment en déduire un modèle et instancier alors *Facebook* sur celui-ci. Nous aborderons enfin les limitations de ce modèle.

2.3.1 CONTRÔLE D'ACCÈS DE LA PLATEFORME *Facebook*

2.3.1.1 MÉCANISMES

D'un point de vue informel, les mécanismes de contrôle d'accès de la plateforme *Facebook* peuvent être ainsi décrits :

1. **Le profil** : représentation numérique de l'utilisateur, le profil est un espace propre à chacun, contenant des *items*, et dont l'accès est géré par son propriétaire.
2. **Les items du profil** : contenu spécifique partagé par l'utilisateur pouvant revêtir diverses formes (photo, vidéo, statut,etc.) décrivant un aspect de son identité numérique (information personnelle, professionnelle, opinions, etc.), et dont l'accès est restreint sélectivement par celui-ci.
3. **Le point d'accès**[11] : accès direct, sous la forme d'un lien, au profil d'un utilisateur. La capacité pour un utilisateur A d'atteindre le point d'accès d'un utilisateur B définit le seul moyen d'accéder à son profil, et donc aux divers items le constituant.
   NB : Les auteurs n'ont pas pris en compte le fait que l'on peut pourtant accéderà certains items d'un utilisateurs (photos, commentaires, certains messages, etc...) par le biais d'un profil tiers, c'est-à dire sans consulter son profil. C'est l'une des difficultés de l'éclatement des informations dans les réseaux sociaux.
4. **Les primitives de communication** : gouvernée par une *politique de communication*, une primitive correspond à une interaction initiée par un utilisateur envers un autre (demande d'amitié, envoi d'un message, post d'un commentaire, etc.). Pour initier une telle primitive, l'initiateur doit préalablement avoir atteint le point d'accès du récepteur.
5. **Politiques** : Associée à chaque item et à chaque primitive, une politique d'accès est fixée par l'utilisateur. Elle cible un ensemble d'utilisateurs autorisés à accéder à la ressource en question.

2.3.1.2 ATTEINTE DU POINT D'ACCÈS

Un utilisateur a deux manières d'atteindre le point d'accès d'un autre :

1. **La recherche du nom global** : l'accesseur tape le nom de l'utilisateur par le biais de la recherche sur le réseau. Une recherche fructueuse le conduira vers le point d'accès du propriétaire. Chaque utilisateur peut spécifier une politique de recherche autorisant un sous-ensemble d'utilisateurs à atteindre son point d'accès par le biais de la recherche du nom global.
2. **La traversée du graphe social** : Dans *Facebook*, chaque utilisateur est amené à se créer une "liste d'amis" correspondant à l'ensemble des points d'accès des utilisateurs ayant une relation d'amitié avec lui. Une traversée du graphe social consiste à examiner les listes d'amis des autres utilisateurs et parcourir sélectivement les différents points d'accès. Chaque utilisateur peut spécifier une politique de traversée autorisant un ensemble d'utilisateurs à consulter sa liste d'amis.

11. "Search listing" dans [25].

2.3.1.3 Accès aux items du profil

Pour accéder aux items du profil d'un utilisateur (Figure 2.5 de [25]), l'accesseur doit dans un premier temps atteindre son point d'accès (voir section précédente). Il émet alors une requête d'accès, qui sera par la suite soumise aux politiques d'accès spécifiées par le propriétaire du profil. Les items du profil seront alors affichés sélectivement selon les autorisations et interdictions que ces politiques lui accordent.

2.3.1.4 Initiation d'une primitive de communication

Pour qu'un utilisateur A puisse initier une primitive de communication envers un utilisateur B (Figure 2.6 de [25]), il doit d'abord atteindre son point d'accès (voir section précédente). À partir de là, la capacité de A à interagir d'une certaine façon avec B (par le biais d'une primitive de communication) va dépendre de la politique que B aura fixée concernant cette primitive (autorisation ou interdiction).

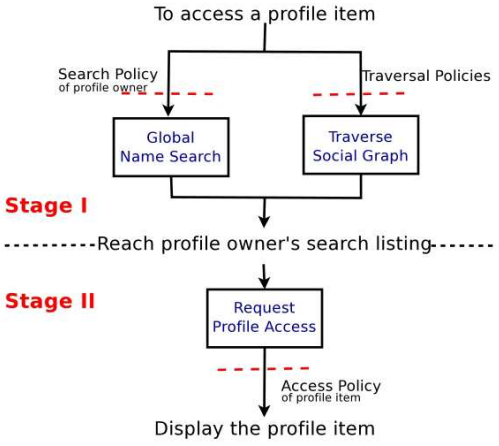


Figure 2.5 – Accéder aux items d'un profil

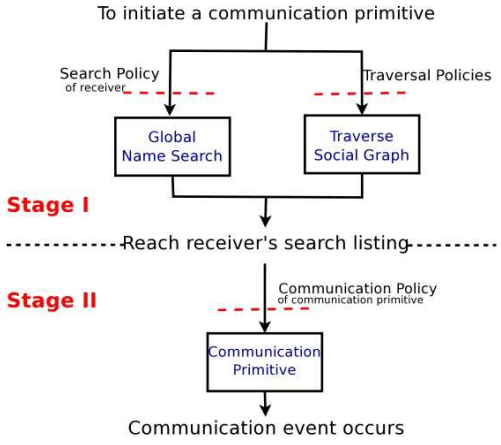


Figure 2.6 – Initier une primitive de communication

2.3.2 Modèle de contrôle d'accès

On note :

- **Sub** : Ensemble fini des utilisateurs du réseau social.
- **Obj** : Ensemble fini des identificateurs d'items (ou objets). Chaque utilisateur possède les mêmes "types" d'objets (par exemple "date de naissance", ou "photo du profil" sont des types d'objets).
- Soient u ∈ *Sub* et o ∈ *Obj*. On note **u.o** l'unique[12] objet de type o possédé par u.
  Par exemple soient u = "Jean Dupont" et o = "lieu de naissance", alors u.o correspond au lieu de naissance spécifique à Jean Dupont et est donc unique.
- Lorsque v ∈ *Sub* veut accéder à **u.o**, on appelle v l'**accesseur** et u le **propriétaire**.

12. Ici, la notion d'unicité ne limite en rien le modèle : un objet **u.o** est dit unique car il est propre à un utilisateur particulier, mais un objet de type **o** est général et peut appartenir à n'importe quel utilisateur.

2.3.2.1 MÉCANISMES D'AUTORISATION

Deux types d'informations sont nécessaires au système pour régir les droits d'accès aux objets : l'**historique de communication** (information d'ordre local) et la **topologie relationnelle** (information d'ordre global). Ces informations seront utilisées afin de définir les **politiques d'autorisation**, modélisés par des prédicats.

HISTORIQUE DE COMMUNICATION L'historique de communication est une information d'ordre local concernant deux utilisateurs. Il est défini par l'ensemble des primitives de communication initiées l'un vers l'autre depuis leur inscription sur le réseau, et sert de base aux décisions liées aux autorisations d'accès, dépendantes de la relation entre ces deux utilisateurs.

Un historique vide, par exemple, induit qu'ils n'ont pas d'amitié directe, et que l'accès tend donc à être plus restreint. En revanche, un historique contenant une requête d'ajout en ami d'un utilisateur u $\in$ *Sub* à un utilisateur v $\in$ *Sub* ; acceptation de v à cette requête induit que u et v sont amis et devraient ainsi avoir accès aux objets de l'un et l'autre (en simplifiant).

L'historique de communication peut ainsi se réduire à l'état courant de la relation entre deux utilisateurs : γ. On pourrait donc assimiler cet historique à l'"état de la relation".

Afin de garantir une certaine cohérence dans la constitution de l'historique de communication, il est nécessaire de respecter un certain protocole. Pour cela, on introduit l'**automate de communication** M $= \langle\ \Sigma,\ \Gamma,\ \gamma_0,\ \delta\ \rangle$ tel que :

- $\mathbf{\Gamma}$: Ensemble fini des primitives de communication.
- $\mathbf{\Sigma}$: Ensemble fini des états de communication.
- $\boldsymbol{\gamma_0} \in \Gamma$: état de départ.
- $\boldsymbol{\delta}$ **:** $\mathbf{\Gamma}$ **x** $\mathbb{B}$ **x** $\mathbf{\Sigma} \rightarrow \mathbf{\Gamma}$: Fonction partielle de transition (où $\mathbb{B}$ est l'ensemble des booléens : $\{0;1\}$).

NB : Pour tout couple (b,p) $\in \mathbb{B}$ x Σ, b représente le booléen identifiant l'utilisateur ayant choisi d'initier la primitive de communication p. On considère pour cela que l'ensemble Sub est totalement ordonné par la relation $\prec$, et on introduit la **fonction d'identification** $\iota_{\{u,v\}} : \{u; v\} \rightarrow \mathbb{B}$ *telle que : ($\lambda x.x = max_{\prec}$(u,v)), associant à chaque membre de la paire (u,v) d'utilisateurs un identificateur booléen unique. l'utilisateur le plus grand selon $\prec$ (identifié de façon unique par le booléen 1, via la fonction d'identification ι) est l'initiateur de la primitive.*

Ainsi, chaque nouvelle entrée dans l'historique de communication est régi par le protocole induit par l'automate garantissant alors sa cohérence. Cet historique permet alors au système de savoir dans quel état de communication se trouvent deux utilisateurs, afin d'appliquer correctement, par la suite, les différentes politiques.

Un automate de communication décrit donc le protocole nécessaire à l'établissement d'un état de communication local entre deux utilisateurs, lors de l'initiation de primitives de communication.

L'historique de communication est donc utilisé principalement pour déterminer la relation entre deux utilsateurs ("amis" ou non par exemple, pour *Facebook*), mais il peut également se révéler être un indicateur du nombre d'interactions réciproque au sein du réseau social.

A tout moment, l'état de communication entre deux utilisateurs, appelé **état de communication global**, est défini par la fonction :

His : Sub x Sub $\rightarrow \mathbf{\Gamma}$

TOPOLOGIE RELATIONNELLE La topologie relationnelle est une information d'ordre global. Pour définir une autorisation d'accès, il est parfois nécessaire de se baser sur des informations ne concernant ni l'accesseur, ni le propriétaire (par exemple, la consultation d'un objet d'un utilisateur par un autre, via un intermédiaire). Il est donc nécessaire de définir la relation binaire, symétrique [13], et irréflexive d'**amitié** par le **prédicat d'adjacence** associant à chaque état de communication local, un booléen :

$$\textbf{Adj} : \mathbf{\Gamma} \to \mathbb{B}$$

Si le prédicat est faux (0), il n'y a pas de relation d'amitié, sinon (1), il y en a une.

On induit alors de cette définition, le **graphe social** :

$$\textbf{SG(Adj,His)} = \lambda\textbf{(Adj,His)} \; . \; \langle \; \textbf{Sub, } \{\{\textbf{u,v}\} \in \textbf{Sub x Sub} \mid \textbf{Adj(His(}\{\textbf{u,v}\}\textbf{))}\} \; \rangle$$

dans lequel les nœuds sont les utilisateurs, et les arêtes les relations d'adjacence.

Par la suite, on notera :
- **V(G)**, l'ensemble des noeuds du graphe G.
- **E(G)**, l'ensemble des arêtes du graphe G (chaque élément de E(G) est un couple (u,v) de noeuds reliés par une arête)

PRÉDICATS DE POLITIQUE On peut ainsi formaliser l'accès, en se basant à la fois sur l'historique de communication et la topologie relationnelle, grâce à un prédicat de politique, défini par la fonction booléenne de type :

$$\textbf{Sub x Sub x G(Sub) x } \mathbf{\Gamma} \to \mathbb{B}$$

où G(Sub) = { Sub, E | E $\subseteq [S]^2$ } désigne l'ensemble des graphes simples ayant Sub comme ensemble des sommets.

On distingue quatre types de politique :

1. **Politique de recherche** : définit les droits d'atteinte du point d'accès.
2. **Politique de traversée** : définit les droits d'accès à la liste des amis une fois le point d'accès atteint.
3. **Politique de communication** : définit les droits d'initier chaque primitive de communication.
4. **Politique d'accès** : définit les droits d'accès à chaque objet du profil.

2.3.2.2 FORMALISATION D'UN SYSTÈME DE TYPE *Facebook*

LE SYSTÈME Soit N, le système de réseau social tel que :

$\textbf{N} = \langle \textbf{Sub, Obj, M, Adj, PS} \rangle$, où :

- **Sub** est un ensemble fini d'utilisateurs.
- **Obj** est un ensemble fini d'identificateurs d'objets tel que chaque objet du système est défini de façon unique par un couple de type *Sub x Obj*.
- $\textbf{M} = \langle \; \Sigma, \; \Gamma, \; \gamma_0, \; \delta \; \rangle$ est un automate de communication.
- $\textbf{Adj} : \Gamma \to \mathbb{B}$ est un prédicat d'adjacence.
- $\textbf{PS} = \{PS_r\}_{r \in R_N}$ est la famille d'espaces de politiques, indexés par les ressources r $\in \boldsymbol{R_N}$ = {**search ; traversal**} $\cup$ $\mathbf{\Sigma}$ $\cup$ **Obj**. Chaque PS_r est un ensemble fini de prédicats de politique (i.e. de type *Sub x Sub x G(Sub) x* $\Gamma \to \mathbb{B}$). C'est-à dire :

13. C'est un choix réducteur (ne permet pas d'introduire de définitions d'amitié assymétrique), mais cohérent avec la définition du réseau *Facebook*.

- $\boldsymbol{PS_{search}}$ = ensemble de prédicats de politique définissant la politique de recherche des utilisateurs,
- $\boldsymbol{PS_{traversal}}$ = ensemble de politiques de traversée du graphe social,
- $\boldsymbol{PS_a}$ = ensemble de politiques de communications pour a $\in \Sigma$,
- $\boldsymbol{PS_o}$ = ensemble de politiques d'accès, pour un objet de type o $\in$ *Obj*.

ETATS DU SYSTÈME A tout moment le système est dans un état S tel que :

S = ⟨His, Pol⟩, où :

- **His** est l'état de communication global.
- **Pol : Sub x $\mathcal{R}$ → $\bigcup_{r\in\mathcal{R}} \boldsymbol{PS_r}$** : spécifie, pour chaque utilisateur et chaque ressource, la politique courante.

On peut alors introduire des règles logiques spécifiant les conditions d'accès au point d'accès (Figure 2.7 de [25]), ou aux items (Figure 2.8 de [25]).

$$S \vdash_N u \text{ finds } u \qquad \text{(F-SLF)}$$

$$\frac{N = \langle \_, \_, \_, Adj, \_\rangle \quad G = \mathsf{SG}(Adj, His) \quad \{u, v\} \in E(G)}{\langle His, Pol\rangle \vdash_N v \text{ finds } u} \qquad \text{(F-FRD)}$$

$$\frac{\begin{array}{c}\langle His, Pol\rangle \vdash_N v \text{ finds } u' \\ N = \langle \_, \_, M, Adj, \_\rangle \quad M = \langle \_, \_, \gamma_0, \_\rangle \quad \gamma = His_{\langle\gamma_0\rangle}(\{u', v\}) \\ G = \mathsf{SG}(Adj, His) \quad \{u, u'\} \in E(G) \quad Pol(u', \mathsf{traversal})(u', v, G, \gamma)\end{array}}{\langle His, Pol\rangle \vdash_N v \text{ finds } u} \qquad \text{(F-TRV)}$$

$$\frac{\begin{array}{c}N = \langle \_, \_, M, Adj, \_\rangle \quad M = \langle \_, \_, \gamma_0, \_\rangle \quad \gamma = His_{\langle\gamma_0\rangle}(\{u, v\}) \\ G = \mathsf{SG}(Adj, His) \quad Pol(u, \mathsf{search})(u, v, G, \gamma)\end{array}}{\langle His, Pol\rangle \vdash_N v \text{ finds } u} \qquad \text{(F-SCH)}$$

FIGURE 2.7 – Règles d'atteinte du point d'accès

$$\frac{\begin{array}{c}\langle His, Pol\rangle \vdash_N v \text{ finds } u \\ N = \langle \_, \_, M, Adj, \_\rangle \quad M = \langle \_, \_, \gamma_0, \_\rangle \quad \gamma = His_{\langle\gamma_0\rangle}(\{u, v\}) \\ G = \mathsf{SG}(Adj, His) \quad Pol(u, o)(u, v, G, \gamma)\end{array}}{\langle His, Pol\rangle \vdash_N v \text{ reads } u.o} \qquad \text{(R-ACC)}$$

FIGURE 2.8 – Règles d'accès à un item

Chaque règle se lit de haut en bas ; tout ce qui se trouve en-dessous de la barre horizontale se déduit de ce qui se trouve au-dessus.

Pour *F-FRD* par exemple, il faut comprendre :

Si l'on considère le système N muni de *Adj*, le graphe social *G*, et deux utilisateurs *u* et *v* reliés par une arête dans *G* (ils sont donc amis),

Alors on déduit que la connaissance du doublet ⟨*His, Pol*⟩, permet de conclure que *v* peut trouver *u*.

Le système change d'état par le biais d'un ensemble de règles de transition $\mathcal{T}_N$ tel que :

$$\begin{array}{lcll} \mathcal{T}_N & \ni & t & ::= & \mathsf{com}(v, u, a) & \text{for } u, v \in Sub, a \in \Sigma \\ & & & | & \mathsf{pol}(u, r, P) & \text{for } u \in Sub, r \in \mathcal{R}_N, P \in PS_r \end{array}$$

Ainsi, soit l'état de communication global *His* change (l'utilisateur v initie une primitive de communication a envers l'utilisateur u), soit un utilisateur u modifie la politique P d'une de ses ressources r. Ces règles sont définies ainsi :

$$\frac{\begin{array}{c} u \neq v \qquad \langle His, Pol\rangle \vdash_N v \text{ finds } u \\ N = \langle \_, \_, M, Adj, \_\rangle \qquad M = \langle \_, \_, \_, \delta\rangle \qquad G = \mathsf{SG}(Adj, His) \\ \gamma = His(\{u,v\}) \qquad b = \iota_{\{u,v\}}(v) \qquad \gamma' = \delta(\gamma, b, a) \\ Pol(u,a)(u,v,G,\gamma) \qquad His' = His[\{u,v\} \mapsto \gamma'] \end{array}}{\langle His, Pol\rangle \xrightarrow{\mathsf{com}(v,u,a)}_N \langle His', Pol\rangle} \text{(T-COM)}$$

$$\frac{N = \langle \_, \_, \_, \_, PS\rangle \qquad P \in PS_r \qquad Pol' = Pol[(u,r) \mapsto P]}{\langle His, Pol\rangle \xrightarrow{\mathsf{pol}(u,r,P)}_N \langle His, Pol'\rangle} \text{(T-POL)}$$

2.3.2.3 CONCLUSION

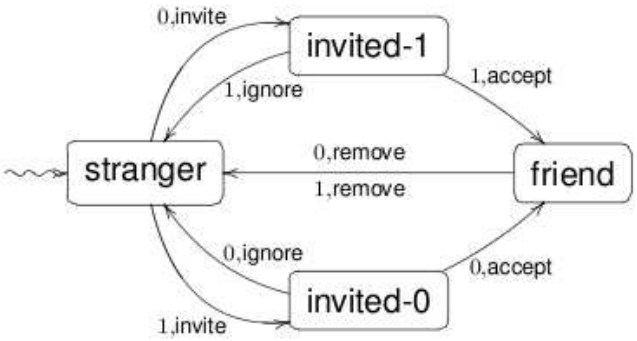
Nous disposons ainsi d'un modèle de contrôle d'accès pour un réseau social de type *Facebook*, dans lequel les utilisateurs peuvent régir par des politiques spécifiques l'accès à leurs ressources. Pour cela le système se base sur des informations d'ordre local (historique de communication, suivi selon un certain protocole déterminé par l'automate de communication), ou global (topologie relationnelle, graphe social). Ce modèle propose donc une gestion d'accès aux données efficace, prenant en compte toutes les informations disponibles sur le réseau afin de s'adapter au mieux aux besoins des utilisateurs. À partir de ce modèle, on peut instancier des réseaux sociaux tels que *Facebook*, avec ses propres règles et propriétés.

2.3.3 INSTANCIATION DE *Facebook*

Le réseau social *Facebook* peut être instancié par le système $\mathcal{FB}$**lite** suivant :

$\mathcal{FB}$**lite = ⟨Sub, Obj , M, Adj , PS**⟩, où :

- *Sub* est l'ensemble fini de tous les identifiants utilisateurs.
- *Obj* est l'ensemble des noms d'item du profil (*Basic-Information*, *Contact-Information*, *Personal-Information*, *Status-Updates*, *Wall-Posts*, *Education-Info*, *Work-Info*, etc.).
- $M = \langle \Sigma, \Gamma, \gamma_0, \delta\rangle$ est l'automate de communication tel que :
  - $\Sigma = \{$invite, accept, ignore, remove$\}$.
  - $\Gamma = \{$stranger, invited-1, invited-0, friend$\}$.
  - $\gamma_0 =$ stranger.
  - δ tel que :



- *Adj* $= (\lambda\ \gamma.\ \gamma =$ friend)
- *PS* est tel que :
  - $PS_{traversal} =$ {no-one, only-me, only-friends, friends-of-friends, everyone}, dont les prédicats sont ainsi définis :

Policy	**Semantics**
no-one	$\bot$
only-me	$\lambda(u, v, G, \gamma)\,.\,u = v$
only-friends	only-me $\vee\ (\lambda(u, v, G, \gamma)\,.\,\{u, v\} \in E(G))$
friends-of-friends	only-friends $\vee$ $(\lambda(u, v, G, \gamma)\,.\,(\exists v' \in Sub\,.\,\{u, v'\} \in E(G) \wedge \{v', v\} \in E(G)))$
everyone	$\top$

  - PS_{search} aurait pu être défini de la même façon que $PS_{traversal}$, à ceci près que lorsqu'un utilisateur v envoie une invitation à devenir ami à un utilisateur u, le point d'accès de v devient disponible pour u. Nous définirons donc PS_{search} ainsi :
    $PS_{search} =$ P $\vee$ owner-invited — P $\in PS_{traversal}$, où :
    owner-invited $= (\lambda$(u, v, G, γ) . (u $\prec$ v $\wedge\ \gamma =$ invited-1) $\vee$ (v $\prec$ u $\wedge\ \gamma =$ invited-0)) est un prédicat retournant vrai si et seulement si u a envoyé une invitation à v.
  - PS_o, pour un certain o $\in$ *Obj*, peut être défini de la même façon que $PS_{traversal}$, à ceci près que, lorsque u envoie une invitation à devenir ami à v, certains objets de u, par exemple "Basic-Information", devient accessible à v. Nous écrirons alors $PS_{Basic-Information} = PS_{search}$. De même pour les autres objets.
  - PS_a, pour tout a $\in \Sigma$, est défini ainsi :

a	PS_a
accept	{everyone}
ignore	{everyone}
remove	{everyone}
invite	{no-one, friends-of-friends, everyone}

2.3.4 DES LIMITATIONS TROP IMPORTANTES

Cette instanciation admet beaucoup de limitations. En effet, ce n'est là que la modélisation très basique du réseau : le modèle ne tient pas compte des groupes, réseaux, communautés, ou des listes d'amis spécifiées par l'utilisateur. De plus l'ensemble *Obj* est loin d'être complet, les messages, "*j'aime*", commentaires, et toute autres interactions autres que des demandes d'amitiés ne sont pas prises en compte par le protocole de communication. Nous pourrions, bien entendu, les modéliser et étendre alors ce modèle, mais cela demanderai un temps dont nous ne disposons hélas pas.

Enfin, le "monde extérieur" comme par exemple la recherche par le biais d'un moteur de recherche autre que celui du réseau social, n'est pas non plus modélisé. En effet, sur *Facebook*, chaque utilisateur peut spécifier d'une part qui peut le trouver via le moteur de recherche *Facebook*, et qui du "monde extérieur" peut le trouver via un moteur de recherche classique (*Google* par exemple). Ce dernier n'est, ici, pas pris en compte.

2.3.5 CRITIQUE ET APPORT POUR NOTRE ÉTUDE

Ce formalisme pourrait constituer une base solide afin de remonter vers l'élaboration d'un système de contrôle d'accès aux données, en introduisant alors des éléments nouveaux correspondant aux études menée plus haut. On pourrait alors imaginer étoffer ce formalisme en comblant les

limitations (enrichissement des primitives de communication par exemple), en élaborant des politiques alternatives basées sur des notions sociologiques étudiées, en modélisant des apports nouveaux inhérents au sujet du stage, etc. et ce par le biais de la logique des prédicats.

Malheureusement, nous n'aurons pas le temps de le faire dans le cadre de notre travail, le modèle admettant de trop grandes limitations compte tenu du temps disponible.

Conclusions et perspectives

Ces travaux permettent de saisir les enjeux techniques des mécanismes de contrôle d'accès de la plateforme *Facebook*, et de mener à bien une réflexion sur les éventuels ajouts et modifications à apporter à ceux-ci afin de répondre au mieux aux problématiques inhérentes à notre travail. Ce travail livre un formalisme possible, basé sur la logique des prédicats et le lambda-calcul ; il ne fournit que les bases de la modélisation, mettant de côté de nombreux aspects. Il pourrait cependant être intéressant de l'enrichir, et disposer alors d'une modélisation fidèle de la gestion des accès du réseau *Facebook* sur laquelle on pourrait introduire de nouveaux mécanismes liés aux travaux menés précédemment et modifier en profondeur les existants.

Conclusion et perspectives générales

Ce travail d'étude m'a permis d'établir une cartographie de différents domaines inhérents au sujet du stage, constituant une fondation pour mon travail ultérieur.

Un dialogue avec de nombreux spécialistes de *Télécom Bretagne*, mon organisme d'accueil, – entre autres sociologues, informaticiens et mathématiciens – a été mené durant l'élaboration de cette partie. Il a été, par ailleurs, pour moi très enrichissant et constructif de discuter avec eux de cet ensemble de travaux, me permettant à la fois de l'améliorer et de bien en saisir les tenants et les aboutissants, mais également d'en assurer la validité scientifique.

L'aspect principal en jeu dans l'établissement d'une relation de confiance étant centré sur la similarité des profils [13], il serait intéressant de proposer aux utilisateurs de classer les informations qu'ils diffusent selon leur contenu, via des tags. Chaque utilisateur pourrait donc associer à chaque thème des listes d'utilisateurs (pas forcément ses amis directs). Les personnes présentes dans cette liste pourraient voir en priorité (apparition plus probable sur leur "fil d'actualité" pour *Facebook* par exemple), ou exclusivement (selon le choix de l'utilisateur) ces informations. Chaque nouvelle demande d'amitié serait une opportunité pour lui de placer directement le demandeur dans une ou plusieurs liste(s) ; il pourrait cependant le faire à tout moment. Le système pourrait également proposer à l'utilisateur des suggestions, basées sur la similarité de leurs profils (via la fonctionnalité "Liens d'amitié" sur *Facebook* par exemple). Il aurait ainsi une meilleure maitrise de la diffusion de ses données.

Si l'utilisateur veut s'exhiber, il doit être conscient de l'impact engendré par cette exposition, et doit pouvoir alors le faire en toute connaissance de cause. Le système devra donc être centré sur la prévention concernant l'impact de la diffusion des informations : nombre d'utilisateurs susceptibles de voir l'information, schémas, graphiques, messages du système insistant sur la création de listes,... La notion de listes d'amis sera également centrale. On pourra relier ce travail à celui effectué sur la confiance en intégrant une liste de confiance sur le thème "exposition". Les informations les plus consultées statistiquement par les utilisateurs souhaitant établir une relation d'amitié devront également être mis en avant.

L'utilisateur devrait également être incité par le système à créer des listes de réseaux sociaux informels : amis proches, collègues de travail, famille, etc. afin de pouvoir diffuser des informations à caractère personnel le mieux possible.

Il serait également intéressant de classer les usages de l'utilisateur sur le repère établi par Dominique Cardon, via un questionnaire par exemple. On pourrait alors adapter certains aspects de l'interface à ses usages, mettant en avant certains paramètres, par exemple. L'utilisateur aurait alors une maîtrise adaptée de l'évolution de son identité numérique.

Enfin, il me faudrait réfléchir à un outil s'appuyant sur divers algorithmes pour analyser le graphe social de chaque utilisateur et en retirer un maximum d'informations liées à la théorie des graphes, afin de proposer une confidentialité adaptée à l'usage qu'il a de son identité numérique sur la plateforme, ou encore à la place qu'il occupe au sein du réseau (centralité). Il serait également intéressant de lui suggérer des listes d'amis en étudiant les communautés d'utilisateurs qui l'entourent.

Nous avions pensé adapter l'ensemble de ces notions au cœur d'un formalisme d'accès aux données. Nous avons donc étudié le formalisme proposé par Fong, Anwar et Zha dans [25]. Cependant, de par l'instanciation trop basique proposée pour *Facebook* et le peu de temps dont je dispose pour accomplir les objectifs du stage, il ne me sera pas possible de m'appuyer sur ce formalisme pour le travail d'élaboration d'une interface de paramétrage de confidentialité. Il est en revanche intéressant de noter qu'il existe ; il pourrait, je pense, à force d'enrichissements (comblant les trop grandes limitations), constituer une réelle base pour un futur travail de modélisation plus approfondi, et servirait alors pour élaborer des paramètres de gestion de confidentialité de manière personnalisée en s'appuyant sur les différentes notions étudiées ici.

Fort de cette base, il va à présent me falloir remonter vers un système aidant l'utilisateur à gérer les droits d'accès aux données pour un réseau généraliste tel que *Facebook*, répondant aux critères d'ergonomie précédemment étudiés et s'appuyant sur ces nouveaux apports.

Chapitre 3

Vers un système interactif de gestion de la confidentialité pour une plateforme de réseautage social généraliste

Introduction

Nous allons voir ici comment appliquer cette étude à l'élaboration d'une plateforme de réseautage social en ligne offrant à l'utilisateur un système de gestion de sa confidentialité adapté.

L'objectif de cette dernière partie est de présenter l'évolution de notre travail au sein d'une démarche progressive partant de l'amélioration de l'existant pour arriver, finalement, à un système interactif de gestion de la confidentialité basé sur notre étude préliminaire.

3.1 Élaboration de scénarios d'usage

Dans un premier temps, il est nécessaire de bien cibler les besoins de l'utlisateur, son contexte, les tâches qu'il peut être amené à réaliser, ou encore les informations nécessaires pour le faire.

Nous avons donc réfléchi à un ensemble de scénarios d'usage d'un réseau social, pour différents types d'utilisateurs, la manière de les réaliser sur notre cas d'étude (*Facebook*), et ce qui nous semble être la façon idéale de le faire.

NB : On supposera ici que l'utilisateur a déjà formé des listes d'amis pertinentes et correspondant à ses attentes, grâce par exemple à l'outil "Fellows", ou bien manuellement.

3.1.1 Mes amis forment un seul et même groupe indistinguable

Exemples : Uniquement de la famille proche, un groupe d'amis rencontré à l'étranger, le réseau professionnel, etc.

SCÉNARIO 1 *Ma liste d'amis est uniquement professionnelle. Je souhaite donc partager mes coordonnées professionnelles avec mes contacts, mais aussi avec leurs contacts à eux : mon profil professionnel pourrait en intéresser d'autres ! Je ne partage mes coordonnées personnelles qu'à certains contacts afin qu'ils puissent me contacter le soir concernant une affaire par exemple.*

⇒ DANS FACEBOOK :
- On place la confidentialité de "Adresse", "Autre téléphone" (correspondant ici au téléphone professionnel), et le champ correspondant à notre adresse électronique professionnelle sur "Amis et leurs amis", et le reste de la page "coordonnées" à "personnaliser", puis "montrer à certains amis".

NB : on ne peut pas spécifier plusieurs adresses postales (une personnelle, une professionnelle par exemple) dans Facebook.

⇒ DANS L'IDEAL :
- Pouvoir spécifier toutes les informations que l'on souhaite (autant d'adresses postales que l'on souhaite par exemple) et pourquoi pas même proposer d'introduire des champs personnalisés (comme adresse-étage-numéro de l'appartement-... C'est l'utilisateur qui choisit) et avoir la possibilité de spécifier la confidentialité de chacun de ses champs un par un.
- Au lieu de passer par la page de gestion, il serait appréciable d'avoir un widget de confidentialité à côté du champ correspondant sur notre profil ; ainsi, on associe directement le niveau de diffusion souhaité non pas grâce au nom du champ (comme cela est fait dans l'interface), mais par son contenu (risque d'erreur largement diminué).

3.1.2 MON PROFIL EST CELUI D'UNE ASSOCIATION, D'UNE ENTREPRISE, D'UN ÉVÈNEMENT RÉGULIER,...

Exemples : Un bar, une troupe de théâtre, une entreprise locale,...

SCÉNARIO 2 *Je suis une association. Mon but étant de toucher le plus de monde possible, je souhaite que toutes mes informations soient publiques. En revanche, je ne v eux pas que l'on puisse m'identifier sur une photo ou une vidéo (par souci de crédibilité). De plus, je ne souhaite aucune dérive (un message sur mon mur par exemple concernant un membre de mon association à titre privé) ; je ne veux donc pas que l'on puisse publier sur mon mur. En revanche, je suis d'accord pour que l'on commente mes publications, mais uniquement si ce sont mes amis.*

⇒ DANS FACEBOOK :
- On place toutes les informations de la page "personnalisez vos paramètres" à "Tout le monde", soit manuellement un par un (conseillé pour être plus sûr et surtout d'avantage conscient de l'impact), soit par le biais du paramètre prédéfini "Tout le monde" sur la page d'accueil de l'interface.
- Sur la page "entrer en contact", on place les paramètres concernant la liste d'amis, la formation, l'emploi, la ville actuelle et d'origine, et les centres d'intérêts, activités et connexions sur "Tout le monde"
- On active la "Recherche publique" (page "applications et sites web") et la "recherche sur *Facebook*" (page "entrez en contact")
- On place le paramètre "photos et vidéos dans lesquels je suis identifié" sur "moi seulement". Cela n'empêche pas mes amis de me "taguer", mais si ils le font il n'y a que moi qui pourrait voir ce tag (il faut quand même le savoir)

- On désactive le paramètre "mes amis peuvent publier sur mon mur"
- On place le paramètre "commenter mes publications" à "amis seulement"

⇒ DANS L'IDEAL :
- Avoir un paramètre explicite "m'identifier sur une photo/vidéo"
- Regrouper les informations personnelles de la page "personnalisez vos paramètres" avec celles de la page "entrez en contact"

3.1.3 Mon profil est personnel : il n'y a que des amis ou connaissances personnelles et des membres de ma famille.

Scénario 3 *Je publie une photo concernant mes amis. Cela m'est égal qu'un "ami d'ami" puisse la consulter, mais je ne souhaite pas qu'il puisse la commenter.*

⇒ DANS FACEBOOK :
- Au moment de publier, je clique sur le cadenas, et sélectionne "Amis seulement".
- Dans les paramètres de confidentialité, je place "commenter mes publications" sur "Amis seulement".

⇒ DANS L'IDEAL :
- On ne peut pas choisir qui peut commenter les publications une par une, on pourrait donc imaginer un widget en plus du cadenas, permettant de définir les personnes ayant le droit ou non de commenter la publication en cours.

3.1.4 Cas généraux de types d'utilisateurs ayant diverses listes d'amis "classiques" (amis, famille, collègues,...)

Scénario 4 *Je souhaite commenter la publication d'un ami, mais je ne veux pas que mon réseau professionnel puisse voir ce commentaire (et donc accéder à la publication).*

⇒ DANS FACEBOOK :
- C'est impossible car la visibilité des commentaires que l'on fait soi-même n'est pas paramétrable et dépend uniquement du niveau de confidentialité associé à la publication de l'ami en question.

⇒ DANS L'IDEAL :
- Disposer d'un paramètre "voir mes commentaires"
- Pourquoi pas un widget a chaque fois que l'on commente définissant sa visibilité ?

Scénario 5 *Je souhaite que mon numéro de téléphone personnel ne soit visible que par mon cercle privé et mon numéro professionnel seulement par mon cercle professionnel.*

⇒ DANS FACEBOOK :
- On place les champ "téléphone" correspondants sur les paramètres "Montrer ce contenu à" la liste Famille pour l'un et "Montrer ce paramètre à" la liste "Professionnel".

⇒ DANS L'IDEAL :

- Même si, dans *Facebook*, on peut définir autant de numéros que l'on souhaite, il n'y en a que 2 qui sont visibles sur le profil... Il serait donc appréciable, dans un premier temps, d'avoir tous les numéros renseignés disponibles sur notre profil, et dans un deuxième temps de pouvoir spécifier la confidentialité de chacun de ces numéros.
- Pourquoi pas un widget placé juste à côté du numéro de téléphone (sur le profil ou sur la page de saisie) afin de spécifier directement sa confidentialité ?

Nous disposons donc à présent d'une vue d'ensemble de besoins concernant plusieurs catégories d'utilisateurs. Il s'agit à présent de réfléchir à l'intégration au sein d'une plateforme de réseautage social d'un ensemble de fonctionnalités permettant de répondre au mieux à ces besoins, et ce grâce à l'étude menée au chapitre précédent.

3.2 Un ensemble de travaux préliminaires d'amélioration de l'existant

Les besoins étant ciblés et la problématique située, nous avons alors orienté notre travail sur un ensemble d'outils pouvant concrètement aider un utilisateur dans sa démarche de gestion de confidentialité.

Dans un premier temps, à partir de la critique menée dans [8] de notre cas d'étude (la plateforme *Facebook*) selon les critères de Bastien et Scapin [9] (voir chapitre 1), nous avons donc réfléchi à l'amélioration de l'interface de gestion de la confidentialité présente dans *Facebook*.

3.2.1 Programmation HTML

Un remodelage complet de l'interface (disponible à l'adresse : `http://pageperso.univ-brest.fr/~e20505489`) a donc été mené, corrigeant, dans la mesure du possible, les exigences de Bastien et Scapin :

- Les **regroupements** ont été complètement revus : les paramètres ont été groupés par type (*Identité numérique*, *Paramètres professionnels*, *Vos amis et vous*, *Activité sur Facebook*, *coordonnées*,...), et selon si les informations sont partagés par l'utilisateur ou par les autres. Il est également possible désormais, par le biais de boîtes à cocher, de sélectionner des paramètres et de les fixer au niveau de confidentialité désiré de façon groupée.
- le **retour d'informations** : programmant hors-ligne, il était difficile d'élaborer quelque action résultant d'une confirmation suite à un changement de paramètre. Cependant, en haut de chaque page, figure *l'amplitude de la diffusion*, à savoir le nombre d'amis de l'utilisateur, le nombre d'amis de ses amis, et le nombre d'amis et "réseaux" (pages d'entreprise, d'école, etc.).
- par rapport à la **charge de travail**, l'utilisateur est toujours conscient de la page où il se trouve, via un ensemble de liens en haut de chaque page. Il peut ainsi passer d'un onglet à l'autre d'un simple click et n'a ainsi pas besoin de passer par la page principale pour naviguer.

Ce premier travail nous a ainsi permis de disposer d'une référence de base respectant les exigences de Bastien et Scapin en terme d'ergonomie. Nous ne proposons donc ici qu'une première évolution, non pas du système en lui-même, mais de l'interface.

3.2.2 UNE INTERFACE ALTERNATIVE INTÉGRÉE

Dans un second temps, nous avons complètement repensé l'interface de gestion de la confidentialité de *Facebook* (voir figure 3.1) :

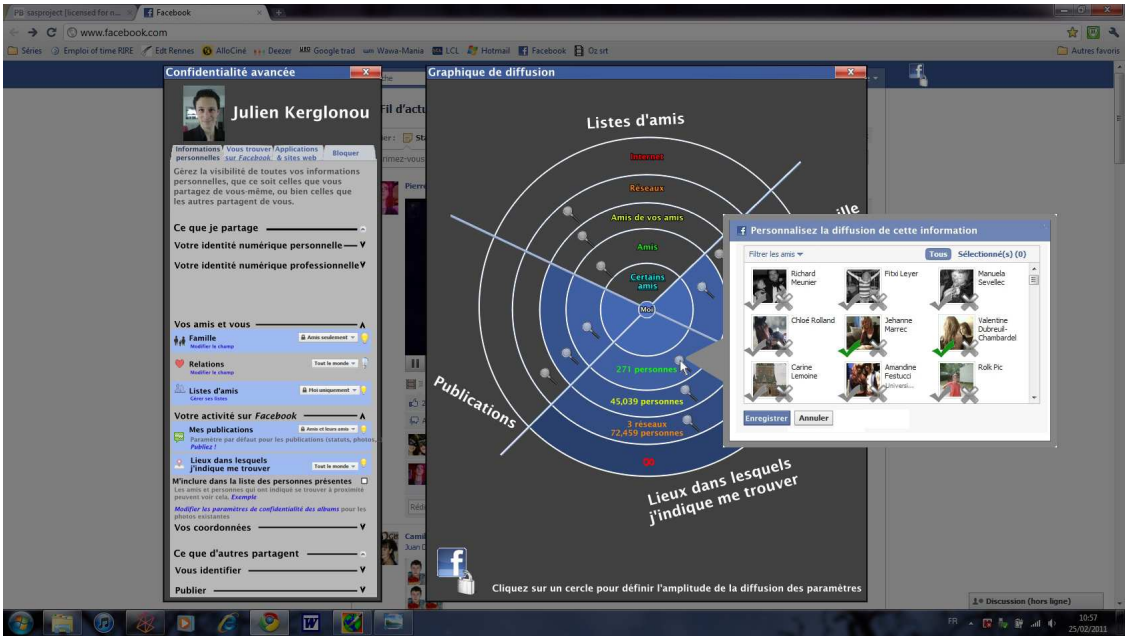


FIGURE 3.1 – Une interface intégrée

L'utilisateur a la possibilité d'accéder à cette interface via une icône (représentant le logo de *Facebook* orné d'un cadenas) présente en haut à droite de n'importe quelle page.

L'interface est partagée en deux fenêtres : une première, à gauche, présente l'ensemble des informations à paramétrer, réparties sur quatre onglets correspondant aux quatre menus de l'interface d'origine. Une seconde fenêtre, au centre, contient une représentation schématique de l'impact de la diffusion d'un ou plusieurs paramètre(s) sous la forme d'un disque. Ce disque représente donc l'étendue de la diffusion : plus on s'éloigne du centre ("Moi"), plus grand est le nombre de personnes ayant accès au(x) paramètre(s) sélectionné(s).

À chaque paramètre est associé à une icône représentative, et une ampoule (correspondant à une case à cocher) permettant de le sélectionner (en bleu). Une fois un ou plusieurs paramètre(s) sélectionné(s), l'impact de la diffusion est présenté à l'utilisateur par le biais du disque, partagé en autant de part qu'il y a de paramètres sélectionnés. L'utilisateur peut alors cliquer sur l'un des disques pour changer la valeur de confidentialité du paramètre (ou il la sélectionne directement dans le menu déroulant).

En cliquant sur la loupe dans les cercles, il visualise les personnes concernées par la diffusion. Il peut alors personnaliser directement la diffusion, ami par ami ou liste par liste (l'icône de validation verte : seules personnes autorisés à voir cette information ; la croix rouge : personnes interdites d'accès).

Chaque paramètre est également explicité clairement (en gris dessous).

Ce travail constitue une version finale d'amélioration de l'interface présente dans *Facebook*. Il pourrait se concrétiser par l'élaboration d'une application *Facebook*, ou encore d'un module pour navigateur web (module *FireFox* par exemple).

3.3 Élaboration d'un système interactif

Fort de cette base, il nous faut à présent nous pencher sur la mise en application de l'étude menée au chapitre précédent. Nous allons donc présenter dans cette partie les possibilités techniques permettant de concrétiser ces travaux.

3.3.1 Placement des utilisateurs dans le repère de Cardon

Comme nous l'avons vu dans le chapitre précédent, D. Cardon a proposé une décomposition de l'identité numérique de l'utilisateur selon un repère [16]. On pourrait donc théoriquement placer l'utilisateur sur ce repère, ciblant ainsi son "type de profil", afin d'amener le système à l'aider dans la gestion de l'accès à ses données.

Pour se faire, il suffirait donc de concevoir un système qui, par le biais d'un questionnaire (comme cela est fait dans [15]) ou bien d'une analyse de ses actions au sein de la plateforme, identifierai la nature de l'identité numérique de l'utilisateur et lui proposerai alors des suggestions adaptées à son profil. Par exemple, si l'utilisateur a tendance à partager beaucoup de photos personnelles avec ses amis, le système pourrait s'adapter et focaliser alors d'avantage l'interface sur la gestion du partage des photos (par le biais de messages "pop-up" par exemple).

Si la position de l'utilisateur sur le repère montre qu'il a tendance à s'exposer, il serait intéressant que le système lui suggère de créer une liste d'amis "exhibition", ou "personnes sûres" afin de maîtriser au mieux cette exposition.

On pourrait également penser un système analysant l'évolution de l'utilisateur sur ce repère, lui proposant alors de configurer certains paramètres concernant des contenus qu'il partage plus fréquemment. Par exemple, si l'utilisateur se met tout à coup à publier beaucoup plus de photos qu'auparavant, le système peut l'informer de l'impact de la diffusion de photos et lui suggérer alors de paramétrer cette diffusion de façon plus fine (par listes d'amis par exemple).

Une vision plus "intrusive", mais qu'il est néanmoins intéressant d'évoquer, serait de positionner les amis de l'utilisateur sur le repère (profils plus "actifs" que d'autres par exemple) et de proposer ainsi à l'utilisateur de les positionner dans des listes prédéfinies. Le système pourrait également analyser leurs évolutions (comportements inhabituels, identités numériques de plus en plus étendue,...) et de l'en informer pour l'aider à décider de partager avec eux ou non.

3.3.2 Les tags

L'une des principales innovations de l'étude présentée au chapitre précédent est l'implémentation de tags qui, associés aux informations publiées sur la plateforme, permettrait à l'utilisateur de gérer la confidentialité de celles-ci non plus de manière globale en fonction des listes d'utilisateurs, mais de manière locale et thématique.

On pourrait donc imaginer, à la manière de plateformes telles que *Youtube* ou certains blogs, une liste de tags, élaborés collaborativement par l'ensemble des utilisateurs, associés à chaque contenu publié (que ce soit des publications ponctuelles – photos, vidéos, liens,... – ou bien les caractéristiques

de l'identité numérique de l'utilisateur – Coordonnées, goûts musicaux ou cinématographique,...). A chaque tag serait associé une liste d'utilisateurs autorisés à consulter le contenu associé (liste de confiance), permettant ainsi à l'utilisateur d'avoir une meilleure maitrise du partage qu'il mène sur la plateforme. Cet accès serait alors géré par le biais de l'I.H.M.

Mis à part certains tags élaborés par défaut représentant des thèmes répandus, il est évident que le système ne peut pas élaborer ces tags de lui-même. On peut donc s'interroger sur les personnes autorisés à associer certains tags à des contenus publié par leurs amis. Bien entendu, le propriétaire du contenu publié est le premier à pouvoir y associer une liste de tags, mais on pourrait également imaginer que l'accesseur lui suggère un ensemble de tags (que l'utilisateur peut alors accepter ou refuser). On aurait donc un système de partage collaboratif permettant au réseau d'amis de l'utilisateur de l'aider dans son partage, et si besoin de l'avertir.

Un autre usage de ces tags serait la configuration du "fil d'actualité". Dans *Facebook*, ce "fil" récapitule les contenus partagés par les amis de l'utilisateur au sein d'une page en perpétuelle évolution, au gré des interactions et des publications. À l'heure actuelle, il est possible de configurer ce fil d'actualité par filtrage : l'utilisateur peut *masquer* "l'actualité" de certains amis ou de certaines applications (c'est-à-dire en interdire l'affichage). Il pourrait être alors très intéressant de pouvoir filtrer cet affichage par thème ! Imaginons qu'un de mes amis sur la plateforme, un collègue par exemple, publie très fréquemment des choses concernant sa passion pour la moto. Sachant que je n'ai guère d'intérêt pour ce sujet, je peux en devenir agacé. Mais voilà, cet ami et moi travaillant dans le même domaine, il peut parfois publier concernant un intérêt commun et cela pourrait alors m'intéresser ! Sur *Facebook*, je n'ai pas le choix, soit je le "masque", et je choisis alors de ne plus voir ses publications quelles quelles soient, soit je ne fais rien et décide alors de tout voir. Avec les tags, ce serait bien plus simple : je configure mon fil d'actualité de telle sorte que le tag "moto", associé donc aux publications qui ne m'intéressent pas, soit masqué. Je ne vois donc plus que les publications qui m'intéressent. Cela peut s'avérer très pratique.

3.3.3 Analyse du graphe social

Nous avons vu plus haut qu'il pourrait être intéressant d'appliquer des mécanismes mathématiques provenant de la théorie des graphes directement sur les graphes sociaux des utilisateurs afin de faire alors ressortir quelques notions inhérentes, intéressantes à utiliser pour guider l'utilisateur.

Ainsi, à l'aide d'algorithmes divers présentés dans [22], on pourrait imaginer un système évaluant certains indicateurs pour chaque utilisateur :

- La cohésion serait évaluée par le calcul de **La densité**, permettant au système de mener alors une analyse égocentrée ou sociocentrée, dégageant ainsi l'influence de l'utilisateur sur le réseau ou encore certaines contraintes induites par celui-ci. Le système pourrait alors informer l'utilisateur des sous-graphes pour lesquels il est influent (et donc lui faire prendre conscience de l'impact potentiel de la diffusion vers ce sous-graphe) ; le repérage de certaines contraintes sur le réseau (par exemple un sous-graphe dont les membres seraient peu connectés entre eux) pourrait permettre au système d'en informer l'utilisateur afin de lui faire prendre conscience que la diffusion d'une information vers celui-ci pourrait avoir une grande importance.
- L'évaluation de la **centralité** pourrait permettre au système de lui proposer des suggestions diverses : par exemple, si l'utilisateur est influent, il pourrait lui proposer un avertissement comprenant le nombre d'utilisateurs touchés par la diffusion d'une information, servant donc de mise en garde. Il pourrait également l'informer des utilisateurs représentant des nœuds centraux, et donc d'un potentiel élevé à faire circuler une information personnelle.

– En **détectant des communautés**, pourrait permettre au système de proposer des listes d'amis prédéfinies pour l'utilisateur, à l'image de l'outil "Fellows", lui épargnant ainsi un travail laborieux.

Nous pourrions également imaginer une plateforme proposant à l'utilisateur de se relier aux autres utilisateurs par d'autres liens qu'une simple relation symétrique d'amitié, induisant alors un graphe social bien plus riche en informations. L'utilisateur pourrait par exemple être ami avec un autre, sans pour autant que l'autre soit ami avec lui, induisant ainsi une notion de prestige (graphe orienté), ou spécifiant un niveau de confiance envers lui (graphe pondéré). Il pourrait également, à l'image du réseau social distribué *Diaspora* [1], proposer à l'utilisateur de placer ses amis dans des "aspects" (des listes d'amis obligatoires) ; le graphe social serait alors étiqueté, et permettrait une diffusion des informations plus sûre.

3.3.4 Aides visuelles et suggestions

Pour bien comprendre l'impact provoqué par la publication d'informations personnelles sur une plateforme de réseautage social, il est important de guider l'utilisateur. Pour cela, il serait intéressant de doter le système de messages, graphiques, ou schémas lui montrant par exemple le nombre d'utilisateurs concernés par une publication. Lui proposer cette information sous forme graphique, ou même textuelle (inscription du nombre d'amis, d'amis des amis,...) est à mon sens très utile pour lui faire prendre conscience de l'impact, et donc de l'importance, de la gestion d'accès à ses informations (comme cela a été fait dans le travail présenté plus haut – voir figure 3.1).

Par la suite, il est important de le guider. On pourrait par exemple imaginer que le système se base sur la similarité des profils (notion de confiance), ou bien même la quantité d'interactions mutuelles, pour proposer à l'utilisateur de partager avec plus de sécurité à un ensemble de personnes.

Il serait enfin utile de doter le système d'un ensemble de listes d'amis prédéfinis correspondant à certains réseaux informels dans lesquels l'utilisateur pourrait se retrouver (amis proches, famille, collègues,...) : il aurait ainsi plus de facilité à placer directement ses nouveaux "amis" dans ces listes, qui pour lui paraissent naturelles car en lien direct avec son quotidien.

3.4 Vue d'ensemble des moyens à disposition pour répondre aux problématiques

3.4.1 Travail restant à réaliser

Durant la fin du stage (mois de juin 2011), il nous reste donc à identifier toutes les informations retirées de ce travail (et présentées au cours de ce rapport) entrant en jeu dans la mise en oeuvre d'un tel système. Il va nous falloir synthétiser tout le travail mené précédement, en extraire l'ensemble des informations nécéssaires à l'élaboration d'un tel système, et réfléchir à leur organisation. On pourra alors s'appuyer sur divers scénarios, et cas d'étude.

3.4.2 Travail *in fine*

Travaillant dans le cadre d'une plateforme de réseau social **en ligne**, les moyens mis à notre disposition pour concrétiser ce travail sont les différents outils disponibles par le biais de la programmation web.

1. `https://joindiaspora.com/`

Notre travail s'orientant dans le domaine des interfaces, nous devons piocher dans la littérature correspondante afin de disposer d'une vision globale de ces outils.

Frédéric Cavazza résume dans [26] qu'une "interface est composée de différents éléments (cadres, items de navigation, moteur de recherche, carrousel, etc.). Ces composants sont en quelque sorte l'alphabet du langage des interfaces, et il n'en existe qu'un nombre limité à partir desquels est composée la très grande majorité des interfaces."

Theresa Neil (conceptrice d'interfaces à l'Université du Texas, Austin) liste dans [27] 30 composants de base pour créer des interfaces riches (champ de saisie avec auto-complétion, carrousel, graphiques, accordéon, boite de sélection multiple, sélectionneur de date, fenêtre nodale, menu flottant, module glissé-déposé, tableau à filtrage dynamique, indicateur de statut, loupe, jauge, aide contextuelle, raccourci clavier, info-bulle géante, module d'édition en ligne, barre de progrès, notation, glissière, tableau dynamique, éditeur WYSIWYG,...). Une vue d'ensemble de ces composants est disponible sur le blog de Frédéric Cavazza [2].

Cet ouvrage très riche pourrait servir de support afin de rendre de notre travail concrêt. Nous sélectionnerions les outils les plus adaptés à nos attentes et tâcherons de réfléchir à leur mise en œuvre concrète.

Conclusion

Suite à l'analyse menée au chapitre précédent, nous avons pu extraire un ensemble de fonctionnalités, utilisables sur un réseau social en ligne de type généraliste, permettant à l'utilisateur de contrôler l'accès à ses données de manière plus fine. La plateforme devra donc disposer des fonctionnalités suivantes :

- Des listes de tags, représentant des thèmes, auxquelles seraient associé un ensemble de listes d'amis de l'utilisateur. Il pourra sélectionner plusieurs listes pour le même thème.
- Des suggestions, basées sur la similarité des profils (ou la quantité d'interactions mutuelles), permettant à l'utilisateur de mieux gérer ses listes d'amis.
- Des messages, graphiques, ou schémas montrant à l'utilisateur quel impact sera engendré par la diffusion de ses données.
- La proposition de création d'une liste d'amis "exhibition" (ou autre appellation assimilée) pour pouvoir s'exposer devant un ensemble d'utilisateurs contrôlé.
- Des listes d'amis prédéfinies correspondant à des réseaux sociaux informels répandus (amis proches, famille, collègues,...).
- Des tags prédéfinis correspondants à des thèmes répandus lors de partages sur des réseaux sociaux : vacances, quotidien, soirées, famille,...
- Un étude personnalisée concernant la nature de l'identité numérique de l'utilisateur, par le biais d'un questionnaire, ou d'une analyse de son activité sur le réseau. On pourra alors le placer sur le repère de Cardon [16] et lui proposer des suggestions correspondant à son type de profil lui permettant de mieux gérer son identité numérique.
- Une analyse du graphe social de l'utilisateur basée sur les algorithmes présentés dans [18] extrayant des indicateurs (centralité, densité, détection de communautés, etc.) permettant au système de lui proposer des suggestions, comme des listes d'amis (comme le fait l'application *Fellows*), ou des informations, par exemple sa position ou son influence au sein du réseau.

2. http://www.interfacesriches.fr/2011/05/27/une-bibliotheque-de-composants-dinterfaces-riches/

Un tel système pourrait s'inscrire dans l'élaboration de nouveaux réseaux sociaux généralistes présentés comme étant une alternative à *Facebook*, comme *Diaspora*, ou encore les travaux menés par *Altly*[3], dont les objectifs sont en phase avec ce travail.

3. `http://altly.com/`, présentés sur le blog `http://blog.altly.com/`

Conclusion générale et perspectives

L'avènement récent des plateformes de réseautage social en ligne est à l'origine d'un engouement considérable. De plus en plus d'utilisateurs sont inscrits sur de tels services à travers le monde, échangeant et partageant des données – à caractère personnel ou professionnel – les uns avec les autres. La quantité de données partagées étant de plus en plus grande, et le nombre d'utilisateurs ne cessant d'augmenter, la question de la gestion de la confidentialité est placée au premier plan des interrogations, la détention d'informations concernant un utilisateur par un tiers pouvant s'avérer lourd de conséquences pour lui.

Avec ses plus de 700 millions d'utilisateurs, la plateforme de réseautage social généraliste *Facebook* est au cœur de cette révolution sociale. Nous avons donc décidé d'en faire notre cas d'étude. Nous avons vu qu'il est difficile pour un utilisateur de gérer le contrôle d'accès aux données qu'il introduit sur la plateforme tout en en gardant la maîtrise au long de son activité au sein du réseau, et de celle des autres. Les risques d'exposition non désirée envers d'autres utilisateurs sont grands ; la difficulté principale est liée au manque d'ergonomie de l'I.H.M. lui permettant de gérer ce contrôle d'accès au sein du réseau, notamment le manque de guidage, ou encore l'absence d'indicateurs concernant l'amplitude de la diffusion de ses informations, l'empêchant alors de bien saisir l'impact de la diffusion de ses données.

Notre travail était donc ici de comprendre comment aider l'utilisateur dans la gestion de l'accès aux données personnelles qu'il introduit au sein du réseau afin de diminuer ses risques d'exposer, de façon indésirable, certaines informations envers des utilisateurs. Le tout devra être fait en suivant certains critères d'ergonomie présentés dans [9] et devra lui permettre de mieux saisir l'impact de la diffusion de ses données.

Cette problématique est vaste, et peut s'appuyer sur de nombreux domaines d'investigation. Nous avons tenté d'en préciser le contour au travers de quelques travaux qui nous ont parus pertinents : sociologie, analyse des graphes sociaux pour les utilisateurs, ou encore formalisation.

Mettant de côté le formalisme pour des raisons techniques (limitations trop importantes, modélisation trop basique, et peu de temps restant), nous avons pensé placer les utilisateurs sur le repère de Cardon [16] (par le biais d'un questionnaire ou d'une analyse de ses actions au sein de la plateforme) dans le but de lui apporter un ensemble d'aides et de suggestions adaptées à son identité numérique. Nous avons également pensé axer la gestion d'accès aux données du système en les identifiant par des tags, représentant des thèmes et auxquels seraient associés des listes de "confiance" (listes d'amis vers qui partager les données en question). Une analyse de son graphe social par le biais de l'évaluation de divers indicateurs provenant de la théorie des graphes pourrait également permettre de fournir à l'utilisateur un ensemble de suggestions (comme par exemple des listes d'amis grâce à la détection de communauté) et d'informations diverses concernant sa position ou encore son influence au sein du réseau : il serait alors d'avantage conscient de son impact sur la plateforme. Enfin, la mise en évidence par le biais de divers graphiques ou schémas, de l'amplitude de diffusion de ses informations lui permettrait d'en saisir mieux le poids.

Mon travail restant durant le mois de juin 2011 consistera donc à réfléchir à la manière d'élaborer l'interface d'un tel système, notamment en repérant les informations utilisables et leur organisation générale.

Bibliographie

[1] V. Glad, « Un anglais tue sa femme à cause d'un statut facebook », 2008. http://www.20minutes.fr/article/264630/High-Tech-Un-Anglais-tue-sa-femme-a-cause-d-un-statut-Facebook.php.

[2] Rédaction, « Un homme pourrait perdre son emploi à cause de facebook », 2008. http://www.gentside.com/facebook/un-homme-pourrait-perdre-son-emploi-a-cause-de-facebook_art2706.html.

[3] M. Rees, « En arrêt maladie pour dépression, ne souriez pas sur facebook », 2009. http://www.pcinpact.com/actu/news/54251-nathalie-blanchard-profil-facebook-assurance.htm.

[4] S. Wasserman et K. Faust, « Social network analysis. methods and applications », Cambridge University Press, 1994.

[5] L. Leca, « services sociaux en ligne, "social graph" et expérience utilisateur », 2007. http://www.kobaye.net/web/services-sociaux-en-ligne-social-graph-et-experience-utilisateur.

[6] D. Maniez et A. Froge, « Intégrer la dimension éthique et le respect de la déontologie », (Université Lyon 2), 2008.

[7] Facebook, « Privacy policy », 22 décembre 2010.

[8] J. Kerglonou, « Gestion de la confidentialité dans un réseau social – etude bibliographique », (UBO, Telecom Bretagne), 2010.

[9] C. Bastien et D. Scapin, « Ergonomic criteria for the evaluation of human-computer interfaces. », 1993. RT no. 156, INRIA.

[10] Afnor, « Management du risque. approche globale », 2002. AFNOR.

[11] P. Sztompka, « Trust : A sociological theory », Cambridge University Press, 1999.

[12] M. Deutsch, « Cooperation and trust. some theoretical notes », (Nebraska Symposium on Motivation), Nebraska University Press, 1962. Jones, M.R. ED.

[13] J. Golbeck, « Trust and nuanced profile similarity in online social networks », (New York, NY, USA), ACM Transactions on the Web (TWEB), 2009. Vol. 3 Issue 4.

[14] J. Golbeck, « Computing and applying trust in web-based social networks », (University of Maryland, College Park, MD, USA.), Ph.D., 2005.

[15] C. Aguiton, D. Cardon, A. Castelain, P. Fremaux, H. Girard, F. Granjon, C. Nepote, Z. Smoreda, D. Trupia et C. Ziemlicki, « Does showing off help to make friends ? experimenting a sociological game on self-exhibition and social networks », rap. tech., faberNovel, Laboratoire SENSE Orange Labs, FING, 2009.

[16] D. Cardon, « Le design de la visibilité : un essai de typologie du web 2.0 », 2008. http://www.internetactu.net/2008/02/01/le-design-de-la-visibilite-un-essai-de-typologie-du-web-20/.

[17] I. Arshad, « L'interaction entre les réseaux sociaux informels et les organismes communautaires formels pour aider les particuliers à gérer les transitions dans leurs parcours de vie », 2011. Horizons.

[18] G. Erétéo, « Analyse des réseaux sociaux et web sémantique : un état de l'art », July 2009. Agence Nationale de la Recherche.

[19] M. Newman, « Scientific collaboration networks. shortests paths weighted networks, and centrality », Phys Rev 64 : 016132, 2001.

[20] D. Watts et S. Strogatz, « Collective dynamics of "small-world" networks. », Nature 393 (6684) : 409–10, 1998.

[21] R. Albert, H. Jeong et A.-L. Barabási, « Internet : Diameter of the world-wide web », sept. 1999. Nature, Volume 401, Issue 6749, pp. 130-131.

[22] N. Bolshakova et F. Azuaje, « Cluster validation techniques for genome expression data », Signal processing, 2003.

[23] P. Mika, « Ontologies are us : A unified model of social networks and semantics », vol. 3729, The Semantic Web. Proceedings of the 4th International Semantic Web Conference, ISCW 2005, November 6-10 2005.

[24] A. Tifous, A. E. Ghali, R. Dieng-Kuntz, A. Giboin, C. Evangelou et G. Vidou, « An ontology for supporting a community of pratice », K-CAP'07, 2007.

[25] M. Anwar, Z. Zhao et P. W. Fong, « An access control model for facebook-style social network systems », technical report, University of Calgary, 2010.

[26] F. Cavazza, « Une bibliothèque de composants d'interfaces riches », 2011. http ://www.interfacesriches.fr/.

[27] T. Neil, « Designing web interfaces : Principles and patterns for rich interactions 1st », O'Reilly Media, Inc., 2009.