

MS/003 - 25/01

Université Abou Bekr Belkaid



جامعة أبي بكر بلقايد

تلمسان الجزائر

République Algérienne Démocratique et Populaire  
Université Abou Bakr Belkaid- Tlemcen  
Faculté des Sciences  
Département d'Informatique

Mémoire de fin d'études

pour l'obtention du diplôme de Master en Informatique

*Option: Système d'Information et de Connaissances (S.I.C)*

*Thème*

**La protection de la vie privée dans  
un système de gestion d'identité.**

**Réalisé par :**

- Mr BOUMRIGA Bahmed.
- Mr BOUMRIGA Hassane.

**Présenté le 1 Juillet 2012 devant le jury composé de :**

- Mr BENAMAR Abdelkarim. (Président)
- Mr BELABED Amine. (Encadreur)
- Mme HALFAOUI A. (Examineur)
- Mme KHITRI S. (Examineur)

Année universitaire : 2011-2012

Insc... 30 SEP 2012  
Date...  
Code... 1676

Inscrit Sous le N°.  
Date le: 15 DEC 2014  
Code: 60

## Remerciement

*Nous remercions ALLAH le tout puissant de nous avoir donné le courage et la Volonté de mener à terme ce présent travail.*

*Nous remercions énormément Mr. BELABED Amine d'avoir accepté de Nous encadrer et nous lui sommes très reconnaissant pour ces précieuses aide pendant les moments difficiles.*

*Nous remercions tout les professeurs qu'il nous enseigné dans les deux années passé et le chef de département informatique et aussi mes amis pour leur soutien et patienter.*





# *Dédicace*

*Je dédie ce mémoire*

*A mes chers parents ma mère et mon père*

*Pour leur patience, leur amour, leur soutien et leurs  
encouragements.*

*A mes frères.*

*A mes amies et mes camarades.*

*Sans oublier tout les professeurs que ce soit du  
primaire, du moyen, du secondaire ou de  
l'enseignement supérieur.*

***BOUMRIGA Bahmed***



# *Dédicace*

*Je voudrais dédier le présent de travail tout spécialement à mes chers parents qui m'ont élevé et soutenu tout au long de ma vie.*

*Je dédie également ce projet à toute la famille de BOUMRIGA et tous mes chers amis Sliman et Med SALAH.*

*Je n'oublie pas un spécial dédie à ma femme et à mon fils Mohamed Zine El Abidine.*

*Enfin, je voudrais dédier ce mémoire à toutes les personnes ayant participé de loin ou de près à la réalisation de ce travail.*

*Hassane*



# Table de matières

<b>Liste des figures</b> .....	4
<b>Liste des Acronymes</b> .....	5
<b>Introduction générale</b> .....	6
 <b>Chapitre I : La protection de la vie privée</b>	
<b>I. Introduction</b> .....	7
<b>II. Définition de la vie privée</b> .....	7
II.1. La vie privée .....	7
II.2. La notion de la vie privée sur Internet .....	8
II.3. La vie privée et les lois .....	8
<b>III. La vie privée sur internet</b> .....	9
III.1. Les niveaux de sécurité .....	9
III.1.1. L'Anonymat .....	10
III.1.2. La Non-chaînabilité .....	10
III.1.3. La Pseudonymat .....	10
III.1.4. La Non-observabilité .....	11
III.2. Les principes fondamentaux de protection de vie privée : .....	11
III.3. Marché noire des données personnelle .....	12
<b>IV. Les attaques sur la vie privée</b> .....	13
IV.1. Les principes de base de la sécurité .....	13
IV.2. Les traces numériques .....	14
IV.3. Définition du vole d'identité (ou usurpation d'identité) .....	14
IV.4. La boite à outils de vole d'identité en ligne .....	14
IV.4.1. Le vol d'identité reposant uniquement sur les logiciels malveillants	15
IV.4.2. L'hameçonnage (« phishing ») .....	16
IV.4.3. Les techniques d'hameçonnage .....	17
IV.4.4. Les différentes formes de vol d'identité .....	19
<b>V. Technologies de protection de la vie privée</b> .....	21
V.1. Privacy by design .....	21

V.2. Privacy Enhancing Technologies « PET » .....	22
V.2.1. Les systèmes de communications et accès anonymes .....	23
V.2.2. Les systèmes de gestion d'identités .....	26
V.2.3. Langages de préférence en termes de vie privée et politiques d'accès	27
<b>VI. Conclusion .....</b>	<b>29</b>

## **Chapitre II : Les systèmes de gestion d'identité**

<b>I. Introduction .....</b>	<b>30</b>
<b>II. Définition .....</b>	<b>30</b>
II.1. Identité numérique .....	30
II.2. Système de Gestion d'Identités.....	31
<b>III. Justifications de besoin d'un SGI.....</b>	<b>32</b>
III.1. Les principaux avantages d'utilisation d'un SGI.....	35
III.1.1. Garantie de traçabilité et d'auditabilité.....	35
III.1.2. Réduction des coûts d'administration.....	35
III.1.3. Amélioration de l'efficacité et de la réactivité.....	35
III.1.4. Amélioration de la sécurité.....	36
<b>IV. Les Modèles de SGI.....</b>	<b>37</b>
IV.1. Définition.....	37
IV.2. Le modèle isolé.....	37
IV.3. Le modèle centralisé.....	38
IV.4. Modèle fédéré.....	39
IV.4.1. Architecture d'Identité Fédérée.....	40
IV.4.2. Échange d'attributs dans un système d'identité fédéré.....	41
IV.4.3. Exemples des modèles fédéré.....	43
IV.5. Modèle centré utilisateur .....	46
<b>V. La gestion d'identité et la vie privée .....</b>	<b>47</b>
V.1. Problématique .....	47

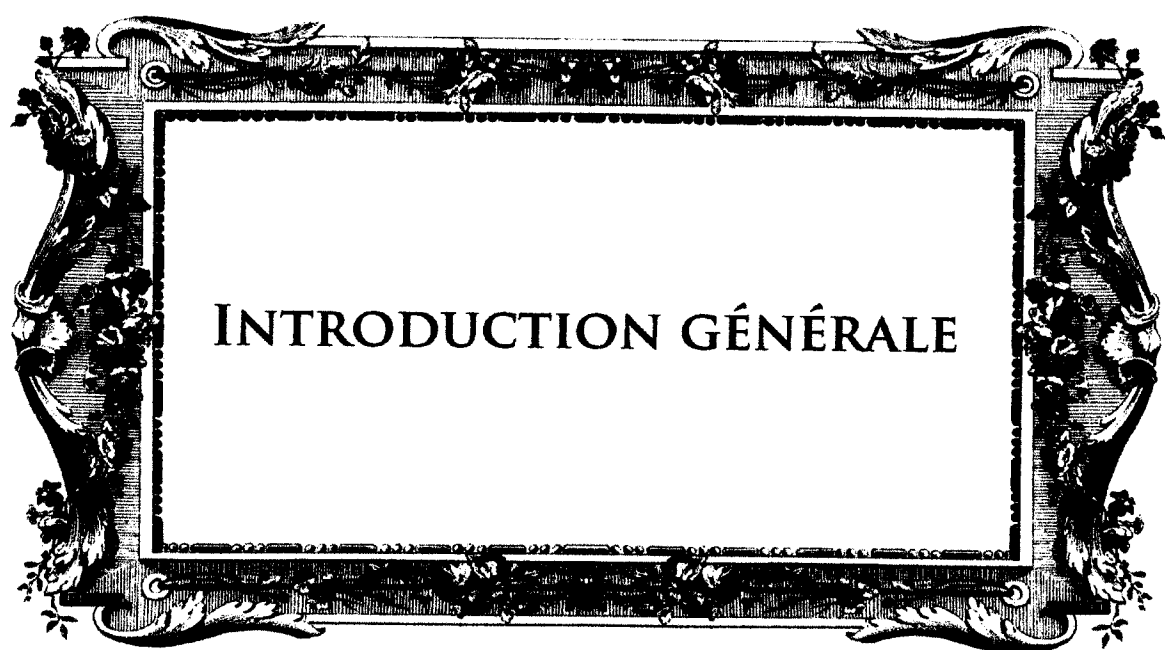
V.2. Les systèmes qui prennent en charge la vie privée.....	48
V.2.1. Microsoft's CardSpace.....	48
V.2.2. Higgins Project centré utilisateurs .....	49
V.2.3. PseudoID.....	51
<b>VI. Conclusion .....</b>	<b>52</b>
 <b>Chapitre III : Conception et implémentation</b>	
<b>I. Introduction.....</b>	<b>53</b>
<b>II. OpenID.....</b>	<b>53</b>
II.1. Mode d'utilisation .....	53
II.2. Problématique de l'OpenID.....	54
<b>III. L'approche proposée (résolution le problème d'OpenID) .....</b>	<b>55</b>
III.1. Généralité sur la solution .....	55
III.2. Architecteur et fonctionnement.....	56
III.2.1. Architecteur .....	56
III.2.2. Fonctionnements (Diagramme de séquence).....	58
<b>IV. Implémentation.....</b>	<b>62</b>
<b>V. Les outils de développements et de conceptions .....</b>	<b>65</b>
V.1. PHPmySQL (WampServer Version 2.2) .....	65
V.2. L'IDE NetBeans 6.8 .....	66
<b>VI. Conclusion .....</b>	<b>66</b>
<b>Conclusion générale.....</b>	<b>67</b>
<b>Références Bibliographique.....</b>	<b>68</b>
<b>Annexe A.....</b>	<b>70</b>
<b>Annexe B.....</b>	<b>72</b>
<b>Annexe C.....</b>	<b>74</b>

# Liste des figures

- Figure I.1** : L'ensemble des pseudonymes les réseaux.  
**Figure I.2** : L'ensemble des Non-observables dans les réseaux.  
**Figure I.3** : prix des données personnelles sur les marchés noirs.  
**Figure I.4** : Triade des besoins fondamentaux sur la sécurité.  
**Figure I.5** : Phishing visant les clients de la Société Générale, BNP Paribas, CIC et CCF.  
**Figure I.6** : MIX avec un tampon pour trois messages.  
**Figure I.7** : Utilisation de deux MIX en cascade.  
**Figure II.1** : Relation entre les éléments composants une identité numérique.  
**Figure II.2** : Composants et fonctionnalités d'un SGI.  
**Figure II.3** : Standards Existants de la gestion des entités et des droits d'accès.  
**Figure II.4** : Flux de mise à jour après la mise en place de système de gestion d'identité centralisée.  
**Figure II.5** : Modèle isolé.  
**Figure II.6** : Modèle centralisé.  
**Figure II.7** : Modèle fédéré.  
**Figure II.8** : Fédération d'identités avec pseudonymes fixes.  
**Figure II.9** : Un possible échange d'attributs dans une architecture d'identité fédérée.  
**Figure II.10** : Les composants de Shibboleth.  
**Figure II.11** : Modèle centré dans l'utilisateur.  
**Figure II.12** : CardSpace Flux Simplifié.  
**Figure II.13** : Base d'implémentation Higgins 2.0.  
**Figure II.14** : Higgins Architecteur.  
**Figure III.1** : Fonctionnement de l'OpenID.  
**Figure III.2** : Les concepts clés de la fédération d'identité.  
**Figure III.3** : Le déroulement des processus de façon globale du système proposé.  
**Figure III.4** : Le déroulement des processus de façon détaillée du système proposé.  
**Figure III.5** : Diagramme de séquence des processus du système.  
**Figure III.6** : Principe de l'algorithme de signature électronique.  
**Figure III.7** : Opération de chiffrement des données signé.  
**Figure III.8** : Navigateur de java avec intégrité des services web.  
**Figure III.9** : La page d'accueil et d'identification de fournisseur des services.  
**Figure III.10** : La page d'identification de fournisseur d'identité.  
**Figure III.11** : La page d'inscription de fournisseur d'identité.  
**Figure III.12** : La page d'authentification de fournisseur d'identité.  
**Figure III.13** : La page d'authentification de fournisseur des services.

# *Liste des acronymes*

<b>AIF</b>	: Architecture d'Identité Fédérée.
<b>AOL</b>	: America Online.
<b>CC</b>	: Cercle de Confiance.
<b>CDM</b>	: Contexte Data Model.
<b>CROWds</b>	: protocole de communication anonyme.
<b>DAC</b>	: Dictionary Access Contril
<b>FI</b>	: Fournisseur d'identité.
<b>FS</b>	: Fournisseur de Service.
<b>GMT</b>	: <i>Greenwich Mean Time</i>
<b>HTTP</b>	: Hypertext Transfer Protocol.
<b>HTTPS</b>	: Hypertext Transfer Protocol Secure.
<b>IBM</b>	: International Business Machines.
<b>IdAS</b>	: l'inférieur sont Identité Attribué Service.
<b>Identités</b>	: représentation d'une personne pour un service.
<b>IMetS</b>	: Identity Meta System
<b>IMS</b>	: Identity Management System
<b>IP</b>	: Internet Protocol.
<b>OASIS</b>	: Organisation for the Advancement of Structured Information Standards.
<b>OCDE</b>	: L'Organisation de Coopération et de Développement Economique.
<b>ONU</b>	: Organisation des Nations Unies.
<b>OV</b>	: Organisation Virtuelle.
<b>OWL</b>	: Ontologie Web Langage.
<b>PET</b>	: Privacy Enhancing Technologies.
<b>Proxy</b>	: Serveur faisant le lien entre un fournisseur de services et d'identités.
<b>P3P</b>	: Platform for Privacy Preferences.
<b>RBAC</b>	: Based Access Control.
<b>RDF</b>	: Resource Description Framework.
<b>SAML</b>	: Security Assertion Markup Language.
<b>SGI</b>	: Système de Gestion d'Identités.
<b>SOAP</b>	: Simple Object Access Protocol.
<b>SSL</b>	: Secure Sockets Layer.
<b>SSO</b>	: Single Sign-On.
<b>TOR</b>	: The Onion Router.
<b>URL</b>	: Uniform Resource Locator.
<b>VoIP</b>	: Voice over Internet Protocol.
<b>W3C</b>	: World Wide web Consortium.
<b>XML</b>	: eXtensible Markup Language.
<b>XUL</b>	: XML-based User interface Language.



INTRODUCTION GÉNÉRALE

# *Introduction générale*

Jour après jour, le nombre de transactions en ligne s'augmente.. Cela a influencé la façon dont les utilisateurs interagissent entre eux et avec d'autres organisations. Une grande partie des services de la vie quotidienne sont accédés de façon numérique. Les exigences concernant l'offre de services en ligne sont de plus en plus fortes en termes de vélocité, disponibilité, mobilité et sécurité. Les entités qui fournissent ces services sont appelées Fournisseurs de Services (FS). Les FS doivent assurer l'intégrité et la confidentialité de l'information échangée dans les transactions en ligne. Quand il s'agit d'information personnelle, les FS devraient offrir des garanties aux utilisateurs que leur information ne sont pas compromises.

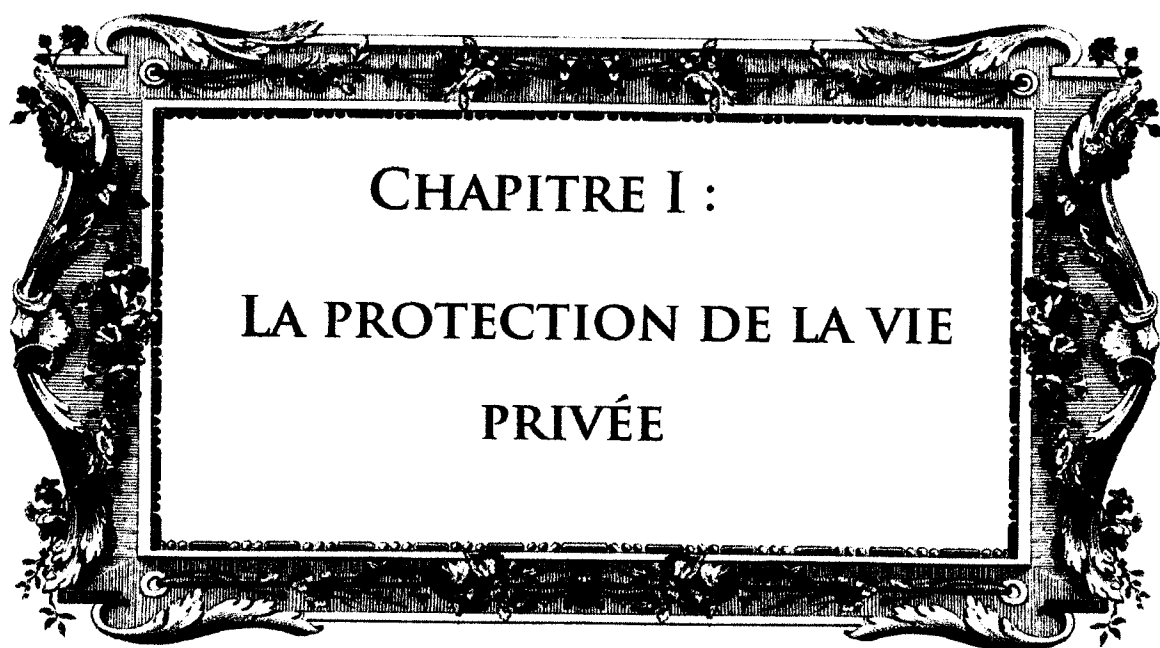
Pour chaque service fourni par les FS, les utilisateurs ont besoin d'une identité numérique qui doit être fourni par un fournisseur d'identité FI.

L'existence de multiples identités numériques représente une situation gênante aussi bien pour les utilisateurs que pour les FS. Pour les utilisateurs, il est compliqué de gérer beaucoup d'identités en accédant à plusieurs services en ligne, ils en font souvent l'amère expérience. En ce qui concerne les FS, chacun doit avoir son propre Système de Gestion d'Identités (SGI) pour gérer le cycle de vie des identités numériques, ce qui complique la collaboration quand il s'agit de fournir des services combinés.

L'objectif de ce mémoire est d'augmenter la protection de la vie privée des utilisateurs des systèmes de gestion d'identité par l'élimination de toutes les connexions et les transactions entre les FI et FS. tout échange entre ces deux entité se fait par l'intermédiaire du navigateur client.

Le reste du document est organisé comme suit : le premier chapitre explique d'une façon générale, le concept de la vie privée sur internet, les attaques relatives ainsi que les techniques de protection. Le 2<sup>eme</sup> chapitre présente un état de l'art sur les systèmes de gestion d'identité numérique, il concentre particulièrement sur la notion d'OpenID et les liens de confiance entre le FS et FI. le 3<sup>eme</sup> chapitre introduit l'approche proposée, une amélioration du système OpenID qui se base sur l'idée de remplacer les liens de confiance par une processus automatique implémenté au niveau de navigateur. Enfin, une conclusion générale qui résume l'ensemble de notre travail, et présente quelques perspectives.

Les annexes A, B et C présentent des informations complémentaires, éclaircissant certains aspects de notre travail.



CHAPITRE I :  
LA PROTECTION DE LA VIE  
PRIVÉE



## I. Introduction

Depuis l'apparition de l'humanité, les personnes essaient toujours de conserver leur vie privée. Avec le développement de la technologie et l'apparition des services d'internet et leur contrôle semi-total sur les domaines de la santé, de finance et de politique..., a affecté pleinement la vie privée des utilisateurs. Les personnes essaient d'être plus prudentes lors de l'utilisation de ce type de technologie.

Dans ce chapitre nous présentons une vue générale sur la vie privée sur internet, les risques relatifs, ainsi que les solutions proposées pour lutter contre la violation de la vie privée des utilisateurs.

## II. Définition de la vie privée

Cette partie définit d'une façon générale le concept de la vie privée, la notion de la vie privée sur internet, ainsi que les enjeux et les défis relatives. Nous terminons par la présentation de quelques lois relatives à cette notion.

### II.1. La vie privée

La vie privée est la capacité, pour une personne ou un groupe, de s'isoler afin de se recentrer sur sa vie et de protéger ses intérêts. Les limites de la vie privée ainsi que ce qui est considéré comme privé diffèrent selon les groupes, les cultures et les individus, bien qu'il existe toujours un certain tronc commun.

L'atteinte à la vie privée peut résulter de la diffusion d'un écrit ou d'une image concernant la personne. C'est souvent autour de contentieux liés à la diffusion d'informations par les grands médias, presse écrite et audiovisuelle que l'on a pu tenter de classer les atteintes à la vie privée.

Différentes composantes de la vie privée sont abordées dans les procès, qui correspondent aux aspects principaux de la vie :

*La vie familiale, la vie sentimentale, les loisirs, la santé, les mœurs, les convictions philosophiques et religieuses, les circonstances de la mort, le droit à l'image.*

Le numéro de sécurité sociale et les références bancaires faisaient partie de la vie privée de chacun, à l'encontre de toute personne dépourvue de motif légitime à en connaître.

Autrement dit, seules les personnes qui en ont besoin (administrations, employeurs...) sont habilitées à les connaître. <sup>[1]</sup>

Il est bien de noter qu'il y a aussi des limites de la vie privée dont on cite :

- ✓ Les activités professionnelles ne font pas partie de la vie privée.
- ✓ La personne concernée est décédée.
- ✓ Les informations révélées sont anodines. <sup>[1]</sup>

## II.2. La notion de la vie privée sur Internet

La vie privée sur Internet est différente que celle de la vie privée normale, du fait que l'Internet est un réseau international et n'importe quelle personne dans le monde peut l'accéder, En plus Internet n'est pas confidentielle à 100%, les pirates informatiques peuvent pirater des informations privées (e-mail, compte bancaire... etc.).

Les communications électroniques constituent aussi une donnée à caractère personnel toute information relative à une personne physique identifiée ou qui peut être identifiée, directement ou indirectement, par référence à un numéro d'identification ou à un ou plusieurs éléments qui lui sont propres. <sup>[2]</sup>

## II.3. La vie privée et les lois

Dans ces dernières années, le droit à la vie privée est devenu un sujet brûlant. Les gouvernements du monde entier ont réagi en promulguant de nouvelles lois pour faire face aux dangers réels ou perçus découlant des technologies numériques et informatisées ainsi que la possibilité de traiter et d'organiser des données de manières innovatrices. <sup>[3]</sup>

En réalité il y a un peu de différences entre les lois de la vie privée des pays et des organisations, pour ce la quelques exemples des lois internationales :

### ❖ Déclaration universelle des droits de l'homme, ONU, 1948 :

Art. 3 : *«Tout individu a droit à la vie, à la liberté et à la sûreté de sa personne ».*

Art.12 : *«Nul ne sera l'objet d'immixtions arbitraires dans sa vie privée, sa famille, son domicile ou sa correspondance, ni d'atteintes à son honneur et sa réputation. Toute*

*personne a droit à la protection de la loi contre de telles immixtions ou de telles atteintes ».*

❖ **Convention européenne des Droits de l'Homme et des libertés fondamentales :**

Art 8 (1950) : *«Toute personne à droit au respect de sa vie privée et familiale, de son domicile et de sa correspondance. ».*

❖ **L'OCDE**

L'Organisation de Coopération et de Développement Economique elle apparut au début de 1980 et comprenant 30 pays, Assemblée Générale de l'ONU, en décembre 1990.

Elle est publiée un Guide pour l'utilisation des données personnelles informatisées et leur transmission internationale.

On doit l'étudier de façon plus spécifique dans la section des attaques contre la vie intime des personnes.

### **III. La vie privée sur internet**

Aujourd'hui Internet est devenu un vaste domaine d'application et de communication : navigateur, messagerie, services web (commerciale, hospice, gouvernement...) ces applications touchants d'une manière ou d'une autre la vie privée des personnes, que signifie le besoin aux politiques de protection de vie privée.

La vie privée des personnes sur l'internet prend plusieurs perspectives, par rapport au niveau de protection existante (Anonymat, Pseudonymat, ...), et le degré de la menace définit les niveaux précédents.

Avant tout nous devons définir les niveaux de protection qui nous aident à comprendre plus facilement les sources de la menace et comment l'éviter.

#### **III.1. Les niveaux de protection de vie privée:**

Les niveaux de sécurité se composent d'une façon générale de deux entités (sujet, objet) et des actions qui relient entre eux.

Le sujet exécute une action sur l'objet. Le sujet peut être un émetteur comme il peut être un récepteur, il utilise l'objet comme message qui le manipule dans un réseau de communication.

Un sujet peut être un être humain (i.e : personne physique), ou une machine. [5]

### III.1.1.L'Anonymat

L'anonymat garantit qu'un utilisateur peut utiliser une ressource ou un service sans révéler son identité d'utilisateur.

### III.1.2.La Non-chainabilité

La non-chainabilité Garantit qu'un utilisateur peut utiliser plusieurs fois des ressources ou des services sans que d'autres soient capables d'établir un lien entre ces utilisations.

La non-chainabilité rend toute entité incapable de relier deux actions anonymes qui ont été menée par le même individu.

### III.1.3.La Pseudonymat

La pseudonymat Idem que anonymat, sauf que l'utilisateur peut quand même avoir à répondre de cette utilisation.

Comme présente la Figure I.1, un pseudonymat a détient ces informations par un détenteur. Qui de son tour représente une entité physique ou logique (personne ou organisation) sur les services web et qui ne doit pas être de son identité réelle. Et une entité peut avoir plusieurs détenteurs pour plusieurs utilisations pour éviter la chainabilité. [5]

Par exemple : un détenteur pour les dossiers médicaux et un autre pour les services financiers ...

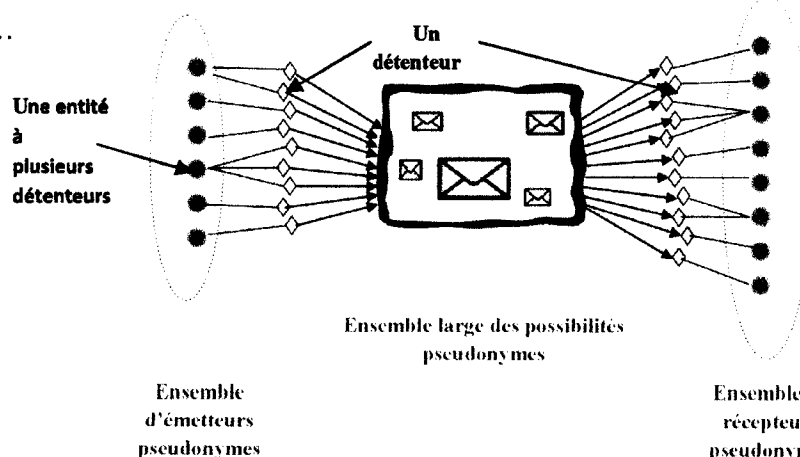


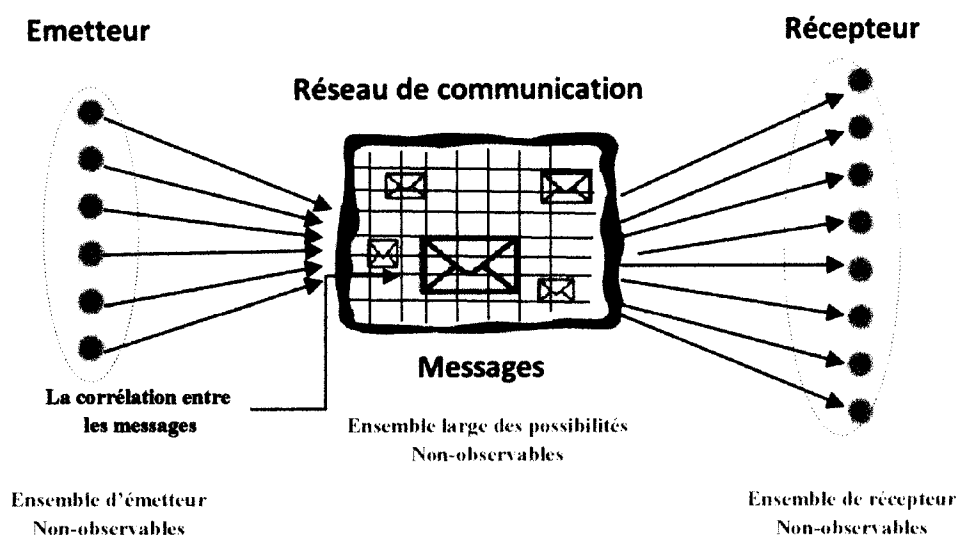
Figure I.1 : L'ensemble des pseudonymes dans les réseaux. [5]

**III.1.4. La Non-observabilité**

La Non-observabilité Garantit qu'un utilisateur peut utiliser une ressource ou un service sans que d'autres, en particulier des tierces parties, soient capables d'observer que la ressource ou le service est en cours d'utilisation.

Contrairement aux définitions précédentes, l'utilisateur peut utiliser sa vraie identité mais sans que les autres peuvent l'observer, ce qui signifie que les sources des messages sont inconnus « Bruits Aléatoires ». [5]

On remarque clairement dans la Figure I.2 en bas les corrélations entre les messages mais toujours restants non-observables et non discriminables.



**Figure I.2 :** L'ensemble des Non-observables dans les réseaux. [5]

Chaque entité qui utilise des données privées des utilisateurs doit respecter les principes suivants :

**III.2. Les principes fondamentaux de protection de vie privée**

✦ Minimisation des données :

Seule l'information nécessaire pour compléter une application particulière devrait être collectée/utilisée (et pas plus).

Application directe du critère de légitimité défini par la directive européenne sur la protection des données personnelles.

✦ **Souveraineté des données :**

Les données liées à un individu lui appartiennent, il devrait pouvoir contrôler comment elles sont disséminées.

Extension de plusieurs législations nationales sur les données médicales qui considèrent que le dossier d'un patient lui appartient, et non pas au docteur qui le crée ou le met à jour, ni à l'hôpital qui les stocke. Difficile à réaliser dans un monde ubiquitaire.

✦ **Consentement explicite :**

Avant de collecter les données personnelles d'un individu, il faut lui demander son autorisation et lui expliquer quelle utilisation sera faite de ses données.

✦ **Transparence :**

Le système ne doit pas être considéré comme une boîte noire dans laquelle l'individu doit avoir une confiance aveugle.

✦ **Imputabilité :**

L'entité qui héberge les données personnelles doit les sécuriser au meilleur de ses moyens, et le cas échéant peut être tenue responsable (par exemple devant un juge) d'un bris de vie privée.

✦ **Droit à l'oubli :**

Sur demande de l'individu, ses traces doivent être effacées.

### III.3. Marché noir des données personnelle

Il existe un marché noir des données personnelles où se vendent les fichiers de certain entreprise :

Exemple de prix de vos données personnelles :

Federal law (See 18 U.S.C. § 2706) requires law enforcement to reimburse providers like Yahoo! for costs incurred responding to subpoena requests, court orders, or search warrants. Yahoo! generally requests reimbursement when responding to legal process, except that Yahoo! maintains an exception to this policy for cases involving the abduction or exploitation of children. Yahoo! may waive reimbursement in specific cases or recognize additional exceptions to this policy in the future.

Yahoo! will seek reimbursement based on the actual time expended by Yahoo!'s compliance staff in complying with the request. The average costs related to compliance matters are listed below for your convenience. These estimates are neither a ceiling nor a floor but represent the average costs of typical searches. Time spent may vary considerably based on the wording of the request and the information available about the user. These time estimates are also based on narrowly tailored requests that do not require extensive searches in multiple databases. These estimates are not price quotes, budgets, or guarantees and should not be used for budgeting purposes. Yahoo! reserves the right to adjust its estimates and reimbursement charges as necessary.

- Basic subscriber records: approx. \$20 for the first ID, \$10 per ID thereafter
- Basic Group Information (including information about moderators): approx. \$20 for a group with a single moderator
- Contents of subscriber accounts, including email: approx. \$30-\$40 per user
- Contents of Groups: approx. \$40 - \$80 per group

Figure I.3 : prix des données personnelles sur les marchés noirs. [22]

Sécuriser l'accès aux données collectées ne résout pas forcément tous les problèmes.

**Exemple :** un employé renvoyé et en colère peut partir avec la base de données des clients.

#### IV. Les attaques sur la vie privée :

La vie privée sur internet est devenue plus insécurisée ce qui rend les informations personnelles comme des marchandises sur les services web et surtout la tierce partie publicitaire, une fois si l'identité d'un individu qui n'a pas les moyens de sécurisation suffisante un voleur d'identité peut prendre cette opportunité pour voler cette identité et l'utiliser pour des raisons non légitimes.

On va montrer les différentes méthodes d'attaque qui peuvent être utilisées contre la vie privée des individus, comme le phishing et le vole d'identité.

Avant de définir les méthodes d'attaque contre la vie privée, on va faire un petit rappel sur les principes de base de sécurité.

##### IV.1. Les principes de base de la sécurité

- La Confidentialité : protéger le contenu d'un message contre un espion qui écouterait les communications.
- L'Authentification : être capable de vérifier l'origine d'un message ainsi qu'éventuellement son intégrité.
- La Disponibilité : assurer la disponibilité d'un service système même contre un adversaire qui essayera de l'attaquer afin de le faire cracher.

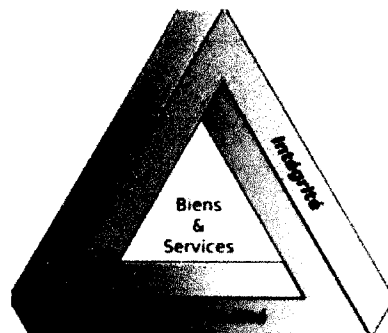


Figure I.4 : Triade des besoins fondamentaux sur la sécurité. [22]

## IV.2. Les traces numériques

- Dans la société de l'information, chaque personne laisse constamment des traces numériques qui peuvent être reliées à son identité.
- Danger : collecte des traces numériques par une entité non-autorisée pour utilisation à des malveillantes.

**Exemple :** lorsque vous consultez une page web, il est possible de relier l'adresse IP de votre ordinateur à cette page

⇒ Permet d'associer le sujet de cette page à votre identité mais aussi d'identifier votre localisation.

⇒ Brise de vie privée !!? [22]

## IV.3. Définition du vol d'identité (ou usurpation d'identité)

Le vol d'identité consiste en l'acquisition, le transfert, la possession ou l'utilisation non autorisés des informations personnelles d'une personne physique ou morale dans l'intention de commettre, ou en relation avec, des actes frauduleux ou autres délits [4].

Cette définition s'applique quel que soit le support au moyen duquel le vol d'identité est perpétré. La portée du présent document se limite toutefois au vol d'identité affectant les consommateurs. [4]

## IV.4. La boîte à outils de vol d'identité en ligne

Le vol d'identité est une activité illicite dont l'histoire remonte bien avant l'Internet. Typiquement, le vol d'identité classique était – et est encore – perpétré au moyen de techniques comme la fouille de poubelles, le vol de cartes de paiement ; le faux-semblant, l'espionnage par-dessus l'épaule, ou le vol d'ordinateur. Ces dernières années, ces agissements ont été modernisés du fait du développement rapide de l'Internet qui, comme on le verra ci-dessous, permet aux voleurs d'identité d'installer des logiciels malveillants sur les ordinateurs et d'utiliser la méthode de l'« hameçonnage », laquelle peut elle-même être perpétrée au moyen de logiciels malveillants et du spam. [4]



#### **IV.4.1. Le vol d'identité reposant uniquement sur les logiciels malveillants**

Le terme général de « logiciel malveillant » (ou « maliciel ») désigne un code ou logiciel introduit dans un système d'information afin de causer des dommages à ce système ou à d'autres systèmes, ou de les destiner à une utilisation autre que celle voulue par leurs utilisateurs légitimes. Avec l'essor de programmes malveillants furtifs comme ceux qui enregistrent les frappes de touche de clavier, ou comme les virus ou « chevaux de Troie » qui se cachent dans un système informatique et capturent secrètement des informations, le logiciel malveillant est devenu un outil technique permettant à lui seul de voler les informations personnelles des victimes.

Les voleurs d'identité emploient les logiciels malveillants selon diverses méthodes, comme les attaques mixtes ou ciblées, pour obtenir les informations personnelles de consommateurs.

##### **IV.4.1.1. Les attaques mixtes (cachées)**

La plupart de ces agissements, qui combinent plusieurs applications malveillantes dans leurs attaques, ont entraîné un changement dans le paysage des « menaces ». Les attaques de logiciels malveillants mixtes utilisent des techniques comme le piratage psychologique (« ingénierie sociale ») pour contourner les défenses en place. Une attaque mixte est constituée, par exemple, lorsque des fraudeurs incorporent un logiciel malveillant à un site Internet pourtant légitime.

##### **IV.4.1.2. Les attaques ciblées (cachées ou ouvertes)**

La plupart des attaques ciblées visent notamment à voler la propriété intellectuelle et les données appartenant à une certaine entité. Du fait que les utilisateurs à travers le monde prennent plus de mesures préventives pour protéger leurs systèmes, les attaquants abandonnent les attaques à grande échelle qui cherchent à exploiter le plus grand nombre possible de failles qu'elles rencontrent, au profit d'attaques plus ciblées. Les attaques ciblées permettent souvent à leurs auteurs de ne pas être détectés par des outils de sécurité (comme des logiciels antivirus et pare-feux) et de maintenir pendant des périodes plus longues un accès privilégié au système d'un utilisateur.

#### IV.4.2. L'hameçonnage (« phishing »)

L'hameçonnage (« phishing ») au moyen duquel les voleurs leurrent les internautes en leur envoyant des messages électroniques trompeurs ou au moyen de faux sites Internet pour les amener par la ruse à révéler leurs informations personnelles, et l'utilisation de pourriels diffusés en masse (« spam ») pour fréquemment installer des logiciels malveillants dans les ordinateurs de leurs destinataires.

##### IV.4.2.1. Contexte

Le terme anglais « phishing » a été inventé en 1996 par des pirates informatiques américains qui détournaient les comptes d'America Online (« AOL ») en soutirant les mots de passe des utilisateurs d'AOL. L'utilisation du « ph » dans cette terminologie remonte aux années 1970, avec les premiers pirates qui se livraient au « phreaking », piratage des systèmes téléphoniques.

L'hameçonnage ou phishing est aujourd'hui décrit, en général, comme une méthode de tromperie que les voleurs utilisent pour « pêcher » les informations d'identité personnelles d'utilisateurs de l'Internet peu méfiants, au moyen de messages électroniques et de sites Internet miroirs revêtant l'apparence de messages émanant d'entreprises légitimes telles que des établissements financiers ou administrations publiques. Comme elle montre la Figure I.5 ci-dessous, un exemple bien connu de phishing est le courriel prétendant émaner d'une banque dont le destinataire est client pour vérifier les identifiants de ce dernier. En France, par exemple, en 2005, une attaque d'hameçonnage a visé en même temps les clients de quatre banques.

```

De: Banque
A: ttos@hsc.fr
Objet: Societe Generale / BNP Paribas / CIC Banque / Banque CCF
Envoyé: Mon, 23 May 2006 08:53:53 +0000 X-Mailer: Microsoft Outlook Express V6.00.2900.2180
Mime-Version: 1.0
Content-Transfer-Encoding: 8bit
Content-Type: text/plain; charset=iso-3859-15

Dear Societe Generale/ BNP Paribas/ CIC Banque/ Banque CCF Member,

This email was sent by your Bank server to verify your e-mail address. You must complete this process by clicking on the link below and entering in the small window your Societe Generale/ BNP Paribas/ CIC Banque/ Banque CCF online access details. This is done for your protection - because some of our members no longer have access to their email addresses and we must verify it. To verify your e-mail address, click on the link below:

If you have Societe Generale account:
http://www.societegenerale.fr/Zev3LIomowQrUmVgFfu9YOp8z9q63599u
If you have BNP Paribas account:
http://www.bnpparibas.com/F4rjdZSKaGyKfEn8ifWcVA872icEt28j9j
If you have CIC Banque account:
http://www.cic.fr/adsDlQop65DS9rAgFUQnEDwX6tsda2Mf0w4mD3630c7v7coEz1
If you have Banque CCF account: http://www.ccf.fr/1eF6FMNPlam5MgTLGhFlbFSrrF5jZzZn59z06zm6f

```

Figure I.5 : Phishing visant les clients de la Société Générale, BNP Paribas, CIC et CCF. <sup>[4]</sup>

### VI.4.2.2. Champ d'application

L'hameçonnage consistait à l'origine en des attaques destinées à tromper la victime au moyen de messages électroniques fallacieux ou « maquillés » et de sites Internet frauduleux usurpant le nom d'une banque, d'un commerçant sur l'Internet ou d'une société de carte de crédit, afin d'amener par la ruse les utilisateurs de l'Internet à révéler leurs informations personnelles <sup>[4]</sup>. Une attaque d'hameçonnage classique par courrier électronique pouvait typiquement se décrire comme suit:

**Étape 1 :** L'hameçonneur envoie à sa victime potentielle un message électronique qui semble en apparence provenir de la banque de cette personne ou d'une autre organisation susceptible de détenir des informations personnelles. Dans cette tromperie, l'hameçonneur reproduit avec soin les couleurs, le graphisme, les logos et le langage d'une entreprise existante.

**Étape 2 :** La victime potentielle lit le message électronique et mord à l'hameçon en donnant à l'hameçonneur des informations personnelles, soit en répondant au message électronique, soit en cliquant sur un lien et en fournissant l'information au moyen d'un formulaire sur un site Internet qui a l'apparence de celui de la banque ou de l'organisation en question.

**Étape 3 :** Par le biais de ce faux site Internet ou du courrier électronique, les informations personnelles de la victime sont directement transmises à l'escroc.

### IV.4.3. Les techniques d'hameçonnage

De plus en plus, le vol d'identité se commet au moyen de logiciels malveillants ou criminels (« crimeware ») <sup>[4]</sup>. Il se propage également par le biais du spam, qui contient souvent lui-même des logiciels malveillants.

Les techniques d'hameçonnage sont de plus en plus perfectionnées et difficiles à détecter. Les formes principales sont les suivantes :

#### IV.4.3.1. L'hameçonnage reposant sur les logiciels malveillants

Bien que les logiciels malveillants ne soient pas le seul moyen par lequel un ordinateur peut être corrompu, ils offrent à l'attaquant la commodité, la facilité d'utilisation et

l'automatisme permettant de mener des attaques à grande échelle qui, sans cela, ne seraient pas possibles faute de compétences et/ou de capacité. Les types d'attaques suivants illustrent les formes que peut prendre l'hameçonnage automatisé :

➤ **Le « pharming »**

Les attaques d'hameçonnage reposant sur les logiciels malveillants peuvent revêtir diverses formes. Ces attaques adoptent souvent la technique du « pharming » (ou du « warkitting ») qui utilise le même genre d'identifiants maquillés que dans une attaque d'hameçonnage classique mais qui, en outre, redirige les utilisateurs, d'un site Internet authentique (d'une banque, par exemple) vers un site frauduleux qui reproduit l'apparence de l'original. Quand l'utilisateur connecte son ordinateur au serveur Internet d'une banque, par exemple, une résolution de nom d'hôte a lieu, qui convertit le nom de domaine de la banque (tel que « banque.com ») en une adresse IP composée de chiffres (telle que 138.25.456.562). C'est au cours de ce processus que les escrocs interviennent et changent l'adresse IP.

➤ **L'attaque de « l'homme du milieu »**

L'attaque de « l'homme du milieu » est un autre exemple d'hameçonnage reposant sur des logiciels malveillants. Cette expression décrit le processus par lequel l'hameçonneur collecte des données personnelles en interceptant le message d'un utilisateur de l'Internet qui était à destination d'un site légitime. Dans l'environnement actuel de convergence de l'Internet, deux autres techniques, reposant sur des appareils autres que des ordinateurs, sont employées depuis peu par les hameçonneurs pour voler des identités.

➤ **Le « SmiShing »**

L'hameçonnage continue de se répandre en touchant des appareils externes tels que les téléphones mobiles. En vertu de cette technique naissante, l'utilisateur d'un téléphone mobile reçoit un message de texte (un SMS) où une compagnie lui confirme qu'il a souscrit à un des ses services de rencontres, en indiquant qu'il lui sera facturé une certaine somme par jour s'il n'annule pas sa commande en se connectant au site Internet de cette compagnie. Ce site Internet est en fait corrompu et est utilisé pour voler les informations personnelles de l'utilisateur.

➤ **Le « Vishing »**

La téléphonie sur Protocole Internet (« Voice over Internet Protocol » ou « VoIP ») offre également un nouveau moyen pour dérober les informations personnelles des individus par le biais des téléphones. Dans ce cas, l'hameçonneur envoie un message électronique maquillé classique, présenté comme provenant d'une entreprise ou institution légitime et invitant le destinataire à former un numéro de téléphone. Les victimes se sentent habituellement plus en sûreté dans ce cas de figure, étant donné qu'il ne leur est pas demandé de se connecter à un site Internet où elles fourniraient leurs informations personnelles. Quand elles appellent, un répondeur automatique leur demande de saisir des informations personnelles telles qu'un numéro de compte, un mot de passe ou autre information à des fins prétendues de « vérification de sécurité ». Dans certains cas, l'hameçonneur ne recourt même pas à un message électronique et appelle à froid les consommateurs pour leur soutirer des informations financières.

**IV.4.3.2. L'hameçonnage véhiculé par le spam**

Le spam est un autre vecteur utilisé pour mener des attaques massives d'hameçonnage. Comme l'indique la Boîte à outils anti-spam de l'OCDE <sup>[4]</sup>, le spam a débuté sous la forme de messages électroniques faisant habituellement la publicité de produits ou services commerciaux. Ces dernières années, le spam a évolué, passant d'une publicité inoffensive à des messages potentiellement dangereux qui peuvent tromper le destinataire et conduire à un vol d'identité. Jusqu'à présent, le spam comportait le plus souvent du texte mais, de plus en plus, il contient des images. La société de sécurité ajoute que, alors que les expéditeurs de spam utilisaient classiquement des noms de domaine de premier niveau bien connus comme « .com », « .biz » ou « .info », ils essaient maintenant d'échapper à la détection en utilisant des noms de domaine de petits pays insulaires comme « .im » (Ile de Man britannique) ; souvent, ces noms de domaine moins connus ne sont même pas inclus dans les filtres de spam <sup>[4]</sup>.

**IV.4.4. Les différentes formes de vol d'identité**

Une fois que les voleurs d'identité ont obtenu les informations personnelles de leurs victimes, ils les exploitent de diverses façons. Dans l'édition 2006 de son *Identity Theft*

*Survey Report*, la FTC des États-Unis classe les actes de vol d'identité en trois grandes catégories :

a) Ouverture de nouveaux comptes (cartes de crédit, comptes bancaires ou emprunts) et autres types de fraude (par exemple, bénéficiaire de soins médicaux).

b) Utilisation illicite de comptes sans carte de crédit ou

c) Utilisation illicite de cartes de crédit seulement <sup>[4]</sup>. Les formes d'utilisation illicite les plus fréquentes de cartes de crédit existantes ou de comptes bancaires sans carte de crédit reposant sur le vol d'identité sont les suivantes <sup>[4]</sup> :

- ✚ Cartes de crédit (61 %).
- ✚ Comptes chèques ou comptes d'épargne (33 %).
- ✚ Services téléphoniques (11 %).
- ✚ Comptes de paiement sur Internet (5 %).
- ✚ Courrier électronique et autres comptes Internet (4 %).
- ✚ Assurances médicales (3 %).
- ✚ Autres (1 %).

Comme mentionné ci-dessus, la fraude à la carte de crédit est la forme la plus répandue d'utilisation illicite de comptes existants. Cette forme de vol d'identité a lieu lorsque le voleur d'identité obtient la carte de crédit elle-même, les numéros associés au compte, ou l'information tirée de la bande magnétique au dos de la carte. Les cartes de crédit pouvant être utilisées à distance, par exemple par le biais de l'Internet, les voleurs d'identité ont souvent la possibilité de commettre des fraudes sans être en possession matérielle de la carte de crédit de la victime.

Les voleurs d'identité se livrent également à des fraudes sur des comptes nouveaux en utilisant les informations personnelles des victimes pour ouvrir un compte, dépenser de

fortes sommes et disparaître. Souvent, les victimes ne découvrent le vol d'identité qu'au moment où un agent de recouvrement les contacte ou lorsqu'elles se voient refuser un emploi, un prêt, une voiture ou une prestation à cause de renseignements négatifs sur leur solvabilité. Dans certains cas, les voleurs d'identité déposent des chèques volés ou contrefaits, ou des chèques sans provision et retirent des espèces, causant des dommages financiers immédiats généralement de grande ampleur <sup>[4]</sup>. Bien que cette forme de vol d'identité soit moins fréquente, elle peut entraîner plus de dommages financiers, il y a moins de chances qu'on la découvre rapidement et la récupération est plus longue pour les victimes. En fait, d'après l'édition 2006 du *ID Theft Survey Report* de la FTC, 24 % des victimes d'ouverture frauduleuse de nouveaux comptes ou d'autres types de fraude ne se sont rendu compte de l'utilisation illicite de leurs informations personnelles qu'au bout de six mois, contre 3 % des victimes d'utilisation frauduleuse de cartes de crédit seulement et de comptes sans carte de crédit existants. Dans cette dernière catégorie de victimes, le délai moyen de constatation de la fraude variait d'une semaine à un mois, tandis que pour l'ouverture frauduleuse de nouveaux comptes et les autres types de fraude, il variait d'un à deux moi <sup>[4]</sup>.

## V. Les technologies de protection de la vie privée

### V.1. Privacy by design

Le concept de « Privacy by Design » : un remède à l'insuffisance des moyens actuels de protection de la vie privée.

Mais elle est aussi une démarche indispensable car il est très difficile d'améliorer la protection de la vie privée dans des systèmes qui n'ont pas été conçus avec cette exigence et elle suscite de plus en plus de travaux de recherche en informatique.

*La protection intégrée de la vie privée repose sur sept principes fondamentaux:*

- ✓ prendre des mesures proactives et non réactives, des mesures préventives et non correctives.
- ✓ assurer la protection implicite de la vie privée.
- ✓ intégrer la protection de la vie privée dans la conception des systèmes et des pratiques.

- ✓ assurer une fonctionnalité intégrale selon un paradigme à somme positive et non à somme nulle.
- ✓ assurer la sécurité de bout en bout, pendant toute la période de conservation des renseignements.
- ✓ assurer la visibilité et la transparence.
- ✓ respecter de la vie privée des utilisateurs <sup>[6]</sup>.

Parmi « Privacy by Design » on peut citer notamment :

- ✚ La conception de systèmes de péages routiers reposant sur la géo-localisation des véhicules tout en préservant la vie privée des conducteurs, en effectuant le calcul du tarif par le matériel de bord du véhicule .
- ✚ La conception de cartes d'identité blanches , ou d'accréditation anonymes, permettant de prouver de nombreux attributs, comme le fait d'être membre d'une association, d'être titulaire d'un permis de conduire, d'un droit de séjour, d'un droit de vote, etc., sans pour autant divulguer son identité.
- ✚ La conception de systèmes permettant de mieux protéger le consentement des individus avant toute divulgation de données personnelles, à travers l'assistance d'agents logiciels mettant œuvre des politiques décidées au préalable.
- ✚ La conception d'architectures permettant d'éviter la « pollution de données» et de garder la maîtrise de ses données personnelles <sup>[8]</sup>.

## V.2. Privacy Enhancing Technologies « PET »

Les PET, proposent un éventail de solutions techniques à usage individuel dont certaines sont très avancées (dans le domaine de l'anonymisation et de l'identification). Cependant, leur utilisation reste limitée.

PET constituent l'un des éléments des systèmes de protection de la vie privée avec la réglementation, l'autorégulation des acteurs économiques et le consumérisme/militantisme <sup>[9], [7]</sup>.



### V.2.1. Les systèmes de communications et accès anonymes

Il existe plusieurs types d'anonymat pour les communications (MIX, TOR, CROWds,...). L'anonymat d'émission est obtenu par un utilisateur face à un attaquant lorsque celui-ci est incapable de détecter l'émission de messages par l'utilisateur. De manière similaire, l'anonymat de réception est obtenu par un utilisateur face à un attaquant lorsque celui-ci est incapable de détecter la réception de messages par étant un *nonce* l'utilisateur. L'anonymat relationnel est obtenu par un groupe d'utilisateurs face à un attaquant lorsque l'attaquant est incapable de savoir si deux membres du groupe communiquent entre eux ou pas. De ces définitions, il découle de façon immédiate que si l'anonymat d'émission (ou de réception) est garanti à chaque utilisateur d'un groupe, le groupe possède la propriété d'anonymat relationnel.

Il est important de remarquer que les anonymats d'émission, de réception et relationnel ne procurent pas par eux-mêmes la confidentialité du contenu des messages : il faut pour cela utiliser des techniques appropriées, comme le chiffrement. À l'inverse, le chiffrement du contenu des messages en soi ne procure pas l'anonymat des communications. Ces différents types d'anonymat sont intimement liés.

Fournir des communications anonymes ne suffit pas pour obtenir un accès anonyme à un service : les messages envoyés au fournisseur de service peuvent contenir des informations identifiants, qu'il faut effacer ou transformer par un mandataire (proxy) avant qu'elles ne soient transmises au fournisseur. Cette transformation dépend de la sémantique du message (c'est-à-dire de la signification de son contenu), et la tâche peut donc être très ardue. Si le mandataire est dédié à un service spécifique, il est relativement aisé d'analyser la syntaxe des en-têtes.

Bien sûr, utiliser un seul mandataire pour accéder à un service suppose que l'on ait confiance dans ses administrateurs, puisqu'ils peuvent enregistrer des informations sensibles sur l'application comme sur les communications. La façon la plus sûre de naviguer consiste sans doute à combiner un relais d'anonymat local au niveau application avec un réseau de MIX au niveau communication. Néanmoins cette solution est coûteuse et peut parfois être remplacée par une solution basée sur un mandataire distant <sup>[9]</sup>.

➤ Les MIX

Les premiers relais utilisés pour l'anonymat de communication furent décrits par David Chaum en 1981. Ces relais, qu'il appela MIX, opèrent de la façon suivante :

- ils n'acceptent que des messages de taille fixe.
- ils déchiffrent les messages entrants avec leur clé privée.
- ils attendent d'avoir reçu un certain nombre de messages.
- puis réordonnent les messages avant de les réémettre.

L'objectif est de fournir un anonymat relationnel entre le groupe des utilisateurs émetteurs d'un message, et le groupe des utilisateurs récepteurs d'un message vis-à-vis d'un attaquant qui observerait tous les messages arrivant et sortant du MIX <sup>[11]</sup>.

**Fonctionnement d'un MIX**

Quand un utilisateur *A* veut envoyer un message *M* à un utilisateur *E*, il ajoute l'adresse de *E* au message et chiffre le tout avec  $K_{MIX}$  (clé publique du MIX). On note ce message chiffré  $K_{MIX}(R_A, M, E)$ ,  $R_A$  Valeur aléatoire utilisée une seule fois.

Quand le MIX reçoit le message chiffré, il le déchiffre avec sa clé privée, élimine le *nonce*  $R_A$ , et met en attente le résultat *M,E*. Quand le MIX a mis en attente un nombre donné de messages il les envoie dans un ordre aléatoire aux adresses de destination.

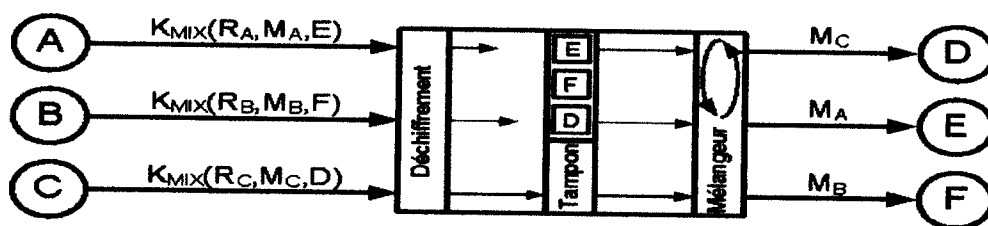


Figure I.6 : MIX avec un tampon pour trois messages. <sup>[11]</sup>

S'il y a peu de trafic, les délais introduits par un MIX de ce type peuvent être excessifs. Pour diminuer les délais de transit, de faux messages peuvent être envoyés au MIX, voire insérés par le MIX lui-même pour assurer que les tampons se remplissent assez fréquemment <sup>[11]</sup>.

➤ **Le « Vishing »**

La téléphonie sur Protocole Internet (« Voice over Internet Protocol » ou « VoIP ») offre également un nouveau moyen pour dérober les informations personnelles des individus par le biais des téléphones. Dans ce cas, l'hameçonneur envoie un message électronique maquillé classique, présenté comme provenant d'une entreprise ou institution légitime et invitant le destinataire à former un numéro de téléphone. Les victimes se sentent habituellement plus en sûreté dans ce cas de figure, étant donné qu'il ne leur est pas demandé de se connecter à un site Internet où elles fourniraient leurs informations personnelles. Quand elles appellent, un répondeur automatique leur demande de saisir des informations personnelles telles qu'un numéro de compte, un mot de passe ou autre information à des fins prétendues de « vérification de sécurité ». Dans certains cas, l'hameçonneur ne recourt même pas à un message électronique et appelle à froid les consommateurs pour leur soutirer des informations financières.

**IV.4.3.2. L'hameçonnage véhiculé par le spam**

Le spam est un autre vecteur utilisé pour mener des attaques massives d'hameçonnage. Comme l'indique la Boîte à outils anti-spam de l'OCDE <sup>[4]</sup>, le spam a débuté sous la forme de messages électroniques faisant habituellement la publicité de produits ou services commerciaux. Ces dernières années, le spam a évolué, passant d'une publicité inoffensive à des messages potentiellement dangereux qui peuvent tromper le destinataire et conduire à un vol d'identité. Jusqu'à présent, le spam comportait le plus souvent du texte mais, de plus en plus, il contient des images. La société de sécurité ajoute que, alors que les expéditeurs de spam utilisaient classiquement des noms de domaine de premier niveau bien connus comme « .com », « .biz » ou « .info », ils essaient maintenant d'échapper à la détection en utilisant des noms de domaine de petits pays insulaires comme « .im » (Ile de Man britannique) ; souvent, ces noms de domaine moins connus ne sont même pas inclus dans les filtres de spam <sup>[4]</sup>.

**IV.4.4. Les différentes formes de vol d'identité**

Une fois que les voleurs d'identité ont obtenu les informations personnelles de leurs victimes, ils les exploitent de diverses façons. Dans l'édition 2006 de son *Identity Theft*

*Survey Report*, la FTC des États-Unis classe les actes de vol d'identité en trois grandes catégories :

a) Ouverture de nouveaux comptes (cartes de crédit, comptes bancaires ou emprunts) et autres types de fraude (par exemple, bénéficiaire de soins médicaux).

b) Utilisation illicite de comptes sans carte de crédit ou

c) Utilisation illicite de cartes de crédit seulement <sup>[4]</sup>. Les formes d'utilisation illicite les plus fréquentes de cartes de crédit existantes ou de comptes bancaires sans carte de crédit reposant sur le vol d'identité sont les suivantes <sup>[4]</sup> :

- ✦ Cartes de crédit (61 %).
- ✦ Comptes chèques ou comptes d'épargne (33 %).
- ✦ Services téléphoniques (11 %).
- ✦ Comptes de paiement sur Internet (5 %).
- ✦ Courrier électronique et autres comptes Internet (4 %).
- ✦ Assurances médicales (3 %).
- ✦ Autres (1 %).

Comme mentionné ci-dessus, la fraude à la carte de crédit est la forme la plus répandue d'utilisation illicite de comptes existants. Cette forme de vol d'identité a lieu lorsque le voleur d'identité obtient la carte de crédit elle-même, les numéros associés au compte, ou l'information tirée de la bande magnétique au dos de la carte. Les cartes de crédit pouvant être utilisées à distance, par exemple par le biais de l'Internet, les voleurs d'identité ont souvent la possibilité de commettre des fraudes sans être en possession matérielle de la carte de crédit de la victime.

Les voleurs d'identité se livrent également à des fraudes sur des comptes nouveaux en utilisant les informations personnelles des victimes pour ouvrir un compte, dépenser de

fortes sommes et disparaître. Souvent, les victimes ne découvrent le vol d'identité qu'au moment où un agent de recouvrement les contacte ou lorsqu'elles se voient refuser un emploi, un prêt, une voiture ou une prestation à cause de renseignements négatifs sur leur solvabilité. Dans certains cas, les voleurs d'identité déposent des chèques volés ou contrefaits, ou des chèques sans provision et retirent des espèces, causant des dommages financiers immédiats généralement de grande ampleur <sup>[4]</sup>. Bien que cette forme de vol d'identité soit moins fréquente, elle peut entraîner plus de dommages financiers, il y a moins de chances qu'on la découvre rapidement et la récupération est plus longue pour les victimes. En fait, d'après l'édition 2006 du *ID Theft Survey Report* de la FTC, 24 % des victimes d'ouverture frauduleuse de nouveaux comptes ou d'autres types de fraude ne se sont rendu compte de l'utilisation illicite de leurs informations personnelles qu'au bout de six mois, contre 3 % des victimes d'utilisation frauduleuse de cartes de crédit seulement et de comptes sans carte de crédit existants. Dans cette dernière catégorie de victimes, le délai moyen de constatation de la fraude variait d'une semaine à un mois, tandis que pour l'ouverture frauduleuse de nouveaux comptes et les autres types de fraude, il variait d'un à deux moi <sup>[4]</sup>.

## V. Les technologies de protection de la vie privée

### V.1. Privacy by design

Le concept de « Privacy by Design » : un remède à l'insuffisance des moyens actuels de protection de la vie privée.

Mais elle est aussi une démarche indispensable car il est très difficile d'améliorer la protection de la vie privée dans des systèmes qui n'ont pas été conçus avec cette exigence et elle suscite de plus en plus de travaux de recherche en informatique.

*La protection intégrée de la vie privée repose sur sept principes fondamentaux:*

- ✓ prendre des mesures proactives et non réactives, des mesures préventives et non correctives.
- ✓ assurer la protection implicite de la vie privée.
- ✓ intégrer la protection de la vie privée dans la conception des systèmes et des pratiques.

- ✓ assurer une fonctionnalité intégrale selon un paradigme à somme positive et non à somme nulle.
- ✓ assurer la sécurité de bout en bout, pendant toute la période de conservation des renseignements.
- ✓ assurer la visibilité et la transparence.
- ✓ respecter de la vie privée des utilisateurs <sup>[6]</sup>.

Parmi « Privacy by Design » on peut citer notamment :

- ✚ La conception de systèmes de péages routiers reposant sur la géo-localisation des véhicules tout en préservant la vie privée des conducteurs, en effectuant le calcul du tarif par le matériel de bord du véhicule .
- ✚ La conception de cartes d'identité blanches , ou d'accréditation anonymes, permettant de prouver de nombreux attributs, comme le fait d'être membre d'une association, d'être titulaire d'un permis de conduire, d'un droit de séjour, d'un droit de vote, etc., sans pour autant divulguer son identité.
- ✚ La conception de systèmes permettant de mieux protéger le consentement des individus avant toute divulgation de données personnelles, à travers l'assistance d'agents logiciels mettant œuvre des politiques décidées au préalable.
- ✚ La conception d'architectures permettant d'éviter la « pollution de données» et de garder la maîtrise de ses données personnelles <sup>[8]</sup>.

## V.2. Privacy Enhancing Technologies « PET »

Les PET, proposent un éventail de solutions techniques à usage individuel dont certaines sont très avancées (dans le domaine de l'anonymisation et de l'identification). Cependant, leur utilisation reste limitée.

PET constituent l'un des éléments des systèmes de protection de la vie privée avec la réglementation, l'autorégulation des acteurs économiques et le consumérisme/militantisme <sup>[9], [7]</sup>.

### V.2.1. Les systèmes de communications et accès anonymes

Il existe plusieurs types d'anonymat pour les communications (MIX, TOR, CROWds,...). L'anonymat d'émission est obtenu par un utilisateur face à un attaquant lorsque celui-ci est incapable de détecter l'émission de messages par l'utilisateur. De manière similaire, l'anonymat de réception est obtenu par un utilisateur face à un attaquant lorsque celui-ci est incapable de détecter la réception de messages par étant un *nonce* l'utilisateur. L'anonymat relationnel est obtenu par un groupe d'utilisateurs face à un attaquant lorsque l'attaquant est incapable de savoir si deux membres du groupe communiquent entre eux ou pas. De ces définitions, il découle de façon immédiate que si l'anonymat d'émission (ou de réception) est garanti à chaque utilisateur d'un groupe, le groupe possède la propriété d'anonymat relationnel.

Il est important de remarquer que les anonymats d'émission, de réception et relationnel ne procurent pas par eux-mêmes la confidentialité du contenu des messages : il faut pour cela utiliser des techniques appropriées, comme le chiffrement. À l'inverse, le chiffrement du contenu des messages en soi ne procure pas l'anonymat des communications. Ces différents types d'anonymat sont intimement liés.

Fournir des communications anonymes ne suffit pas pour obtenir un accès anonyme à un service : les messages envoyés au fournisseur de service peuvent contenir des informations identifiants, qu'il faut effacer ou transformer par un mandataire (proxy) avant qu'elles ne soient transmises au fournisseur. Cette transformation dépend de la sémantique du message (c'est-à-dire de la signification de son contenu), et la tâche peut donc être très ardue. Si le mandataire est dédié à un service spécifique, il est relativement aisé d'analyser la syntaxe des en-têtes.

Bien sûr, utiliser un seul mandataire pour accéder à un service suppose que l'on ait confiance dans ses administrateurs, puisqu'ils peuvent enregistrer des informations sensibles sur l'application comme sur les communications. La façon la plus sûre de naviguer consiste sans doute à combiner un relais d'anonymat local au niveau application avec un réseau de MIX au niveau communication. Néanmoins cette solution est coûteuse et peut parfois être remplacée par une solution basée sur un mandataire distant <sup>[9]</sup>.

### ➤ Les MIX

Les premiers relais utilisés pour l'anonymat de communication furent décrits par David Chaum en 1981. Ces relais, qu'il appela MIX, opèrent de la façon suivante :

- ils n'acceptent que des messages de taille fixe.
- ils déchiffrent les messages entrants avec leur clé privée.
- ils attendent d'avoir reçu un certain nombre de messages.
- puis réordonnent les messages avant de les réémettre.

L'objectif est de fournir un anonymat relationnel entre le groupe des utilisateurs émetteurs d'un message, et le groupe des utilisateurs récepteurs d'un message vis-à-vis d'un attaquant qui observerait tous les messages arrivant et sortant du MIX <sup>[11]</sup>.

#### Fonctionnement d'un MIX

Quand un utilisateur  $A$  veut envoyer un message  $M$  à un utilisateur  $E$ , il ajoute l'adresse de  $E$  au message et chiffre le tout avec  $K_{MIX}$  (clé publique du MIX). On note ce message chiffré  $K_{MIX}(R_A, M, E)$ ,  $R_A$  Valeur aléatoire utilisée une seule fois).

Quand le MIX reçoit le message chiffré, il le déchiffre avec sa clé privée, élimine le *nonce*  $R_A$ , et met en attente le résultat  $M, E$ . Quand le MIX a mis en attente un nombre donné de messages il les envoie dans un ordre aléatoire aux adresses de destination.

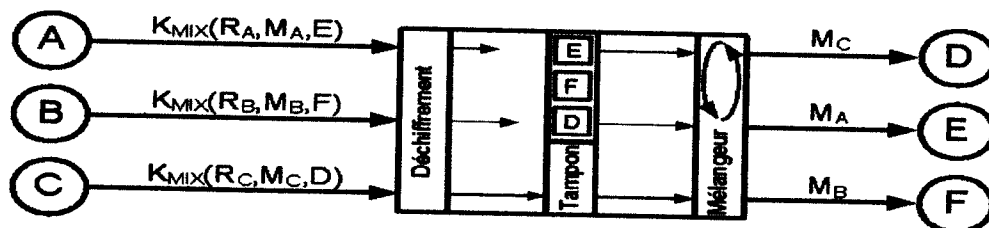


Figure I.6 : MIX avec un tampon pour trois messages. <sup>[11]</sup>

S'il y a peu de trafic, les délais introduits par un MIX de ce type peuvent être excessifs. Pour diminuer les délais de transit, de faux messages peuvent être envoyés au MIX, voire insérés par le MIX lui-même pour assurer que les tampons se remplissent assez fréquemment <sup>[11]</sup>.



### Utilisation de plusieurs MIX

Quand un utilisateur veut utiliser plusieurs MIX les uns après les autres il va appliquer plusieurs couches de chiffrement pour qu'elles soient traitées et retirées à raison d'une par MIX utilisé. Par exemple, un utilisateur  $A$  qui veut envoyer un message  $M$  à un utilisateur  $U$  va chiffrer une première fois le message pour le MIX le plus proche du destinataire, qu'on nomme MIX2 et obtenir  $K_{MIX2}(R_A, M, U)$ . Il va ensuite rechiffrer ce message à l'intention d'un autre MIX(MIX1) précédant MIX2 dans le chemin de  $A$  à  $U$ , pour obtenir :

$$K_{MIX1}(R_A, K_{MIX2}(R_A, M, U), MIX2)$$

La Figure I.7 illustre l'exemple donné ci-dessus pour deux MIX en cascade.

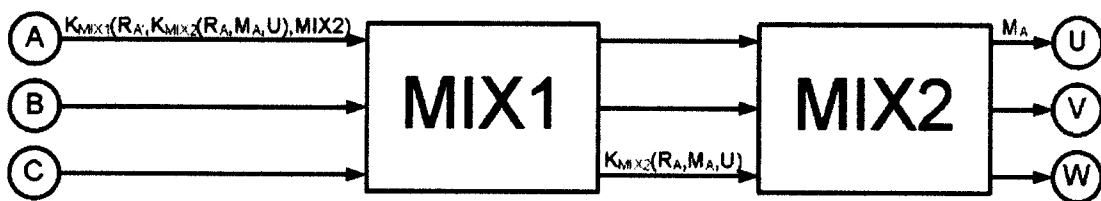


Figure I.7 : Utilisation de deux MIX en cascade. <sup>[11]</sup>

#### ➤ Tor

Tor (The Onion Router) est un réseau de routeurs interconnectés, comprenant des nœuds d'entrée et des nœuds de sortie. L'utilisateur se connecte à un nœud d'entrée, et ses messages transitent par un certain nombre de routeurs (suivant un chemin aléatoire) avant de sortir du réseau Tor par un nœud de sortie et d'être transmis au destinataire final de manière conventionnelle.

Pour utiliser Tor, l'usager doit installer un logiciel de proxy spécial sur sa machine et configurer ses logiciels (principalement son navigateur web, mais possiblement tout logiciel utilisant le protocole TCP/IP) pour l'utiliser. N'importe qui peut utiliser le réseau Tor sans enregistrement préalable, mais l'accès aux routeurs d'entrée et de sortie, dont la liste est publique, peuvent être bloqués dans les pays pratiquant la censure (c'est le cas en Chine) ou bien par des services n'acceptant pas les communications anonymes. D'autre part, l'utilisation de Tor ralentit énormément les activités réseau. Il est par conséquent à réserver à des tâches ponctuelles particulièrement sensibles et/ou nécessitant peu de bande passante <sup>[8]</sup>.

➤ **CROWds**

Est un protocole de communication anonyme qui protège l'anonymat de l'expéditeur d'un message en le routant de manière aléatoire vers des groupes d'utilisateurs similaires, et l'idée principale de cette protocole est cacher l'origine d'un message en le faisant se fondre dans la foule.

Le système particulièrement adapté pour les réseaux du type pair-à-pair et l'initialisation de crowds c'est chaque nouvel utilisateur s'enregistre en tant que membre d'un groupe (appelé "Crowd") en contactant le responsable du groupe, quand un utilisateur rejoint un groupe, tous les membres du groupe en sont notifiés. Le responsable du groupe est aussi chargé de la distribution des clés symétriques assurant la confidentialité entre paires de nœuds.

L'innocence probable de cette protocole c'est l'adversaire est incapable de prédire avec plus de 50% de confiance, le nœud qui est l'initiateur d'un message, chaque nœud apparaît comme ayant pu ou non être l'expéditeur d'un message (et donc il est probablement innocent).

La notion d'anonymat de crowd est dépend de la taille du groupe ("crowd") et de la probabilité  $p_f$ . Plus la probabilité  $p_f$  est grande, plus l'anonymat de l'expéditeur est protégé mais aussi plus la longueur moyen du chemin généré sera longue <sup>[12]</sup>.

### V.2.2. Les systèmes de gestion d'identités<sup>1</sup>

Pour qu'une personne puisse protéger sa vie privée, il est important de cacher ou de réduire autant que possible les liens entre cette personne et les actions et données correspondantes. Par exemple, si une personne est le seul utilisateur d'un ordinateur connecté à Internet avec une adresse IP fixe, il est possible pour un observateur d'associer à cette personne toutes les informations émises depuis cette adresse IP : l'adresse IP peut alors être considérée comme un identifiant unique, c'est-à-dire qu'il est propre à une seule personne. De tels identifiants uniques permettent d'établir un lien fort entre différentes actions indépendantes réalisées par la même personne, ou entre des ensemble d'informations liées à la même personne. C'est donc une menace directe contre la vie privée, et en particulier vis-à-vis de la troisième exigence des critères communs présentés précédemment (unlinkability).

---

<sup>1</sup> Identité : représentation d'une personne pour un service.

Un moyen pour réduire les risques d'établissement de tels liens consiste à utiliser des communications anonymes et des accès anonymes aux services. Mais bien souvent c'est insuffisant, puisque pour obtenir un service personnalisé, l'utilisateur doit se faire reconnaître avec une identité. Cette fois encore, si une personne accède à plusieurs services sous la même identité, il est possible d'établir un lien entre ces accès. Aussi est-il souhaitable d'avoir des identités virtuelles (ou pseudonymes) multiples pour accéder à des services multiples <sup>[10]</sup>.

**Exemples :** Windows Live ID, Single Sign-On (SSO), OpenID...

### V.2.3. Langages de préférence en termes de vie privée et politiques d'accès

Les informations de sécurité doivent être échangées entre les organisations (fournisseurs d'identité et fournisseurs de services). Ces informations incluent la preuve de l'authentification de l'utilisateur auprès de son organisation de rattachement, la méthode d'authentification utilisée, la date et l'heure de l'authentification, etc. Ceci, permet aux fournisseurs de services d'appliquer des politiques de sécurité dépendamment du contexte d'authentification de l'utilisateur. Ainsi, les droits d'accès d'un utilisateur à une ressource diffèrent selon le niveau de sécurité lié à la méthode d'authentification utilisée.

L'information de sécurité (preuve d'authentification et les accréditations d'un utilisateur), échangée entre les sites partenaires, peut être exprimée en utilisant des standards tels que Security Assertion Markup Language (SAML) ou bien des spécifications telles que Web Services-Federation (WS-Federation).

Exemple d'un langage : Security Assertion Markup Language (SAML).

#### V.2.3.1. Security Assertion Markup Language « SAML »

L'OASIS<sup>2</sup> a défini le standard SAML, [SAML] basé sur le langage XML<sup>3</sup> (pour l'échange de données d'authentification et d'autorisation entre les domaines de sécurité distribués). SAML permet aux partenaires de générer des assertions concernant l'identité, les attributs et les droits d'une entité (utilisateur, machine)

<sup>2</sup> Organisation for the Advancement of Structured Information Standards.

<sup>3</sup> eXtensible Markup Language.

et, les transférer à d'autres entités (organisations, applications, etc.). Il définit la syntaxe et les règles pour demander, créer, communiquer et utiliser les assertions SAML au format spécifié. Les assertions SAML, encapsulées dans des messages SOAP<sup>4</sup> et transférées par le biais du protocole HTTP, permettent à la fédération de surpasser les limites imposées par les différences entre les infrastructures déployées chez les différents partenaires.

✦ SAML 2.0, dans sa conception, fournit des solutions pour l'authentification unique (SSO), interaction de Services Web en plus de la fédération d'identité, nous allons décrire ces trois solutions:

- **Authentification unique (SSO)** : fournit une solution multi domaine en proposant une sémantique et un protocole standardisés pour le transfert de l'information concernant un utilisateur entre les domaines.
- **Fédération d'identité** : permet à des organisations partenaires de s'entendre sur (et établir) l'existence d'un identificateur partagé unique pour faire référence à un utilisateur et l'exploitation de cet identificateur pour propager les informations requises le concernant chaque fois qu'il demande un accès au service.
- **Services Web** : offre une modularité permettant aux assertions d'être utilisées dans des contextes dépassant ceux de SAML <sup>[12]</sup>.

#### V.2.3.2. The Platform for Privacy Preferences « P3P »

Le standard P3P, promu par le W3C, permet aux internautes de reconnaître automatiquement en se connectant à un site la politique de protection des données du site. D'une protection « individuelle » ou « marchande » (confiée à des prestataires), on passe, avec le P3P, à une forme de protection collective <sup>[7]</sup>.

##### ➤ *La spécification P3P 1.0*

La spécification P3P 1.0 définit la syntaxe et la des politiques de confidentialité P3P et les mécanismes permettant d'associer les politiques aux ressources web. Les politique P3P consistent en déclarations utilisant le vocabulaire P3P afin d'exprimer des pratiques P3P touchant à la vie privée. Les politiques P3P appellent également des éléments du

---

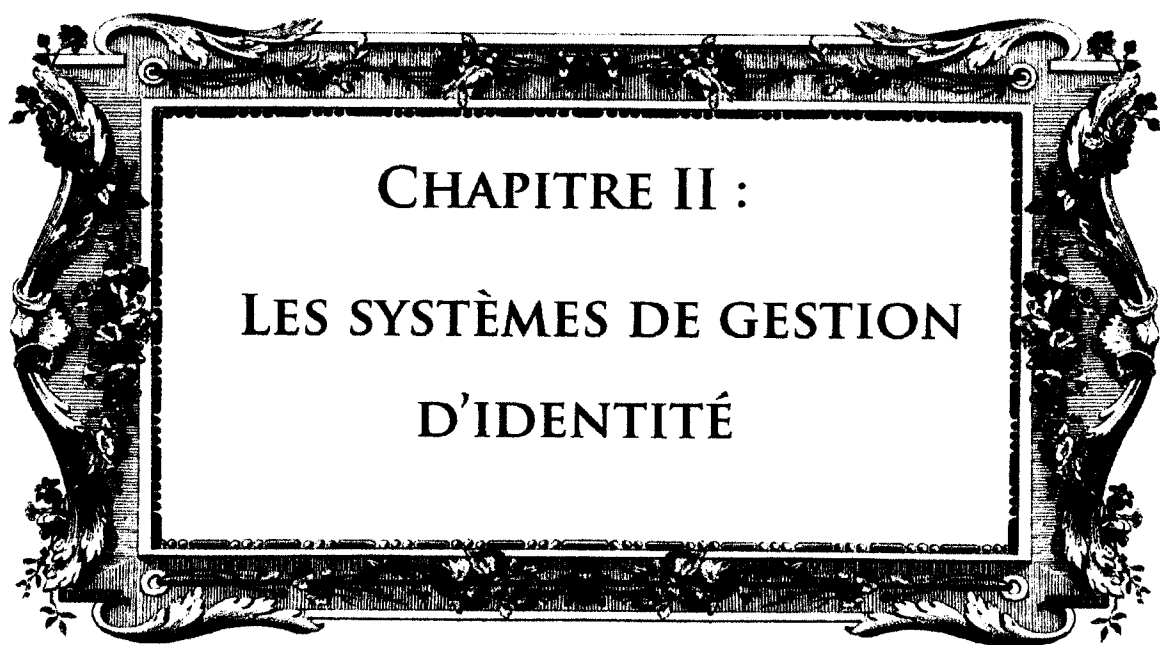
<sup>4</sup> Simple Object Access Protocol.

schéma de données de base de P3P, un jeu standard d'éléments de données que tout agent utilisateur P3P devrait reconnaître. La spécification P3P comprend un mécanisme permettant de définir de nouveaux éléments de données et ensemble de données et un mécanisme simple autorisant l'extension du vocabulaire de P3P <sup>[13]</sup>.

#### **VI. Conclusion :**

Il existe plusieurs méthodes et propositions pour augmenter la sécurité de la vie privée dans les services d'internet. Malheureusement il n'existe pas une standardisation mais que des propositions. Ces propositions ne sécurisent pas la totalité des services web mais elles essayent de fournir plus d'intimité et plus de protection aux vies et aux informations des utilisateurs contre les attaques probables par les voleurs d'identité.

L'une des méthodes les plus prometteuses dans ce domaine sont les systèmes de gestion d'identité (Identity Management system idMs) le prochain chapitre explique en détail ce type de système.



CHAPITRE II :  
LES SYSTÈMES DE GESTION  
D'IDENTITÉ

## I. Introduction

Plusieurs solutions et approches sont proposées pour protéger la vie privée des personnes dans les web services. Parmi ces solutions on trouve les systèmes de gestion d'identité, qui devient un sujet d'actualité et un champ de recherche très actif dans le domaine de sécurité des informations personnelles et de protection de la vie privée sur l'internet.

Dans ce chapitre nous présentons les besoins qui ont conduit à ce type de systèmes, leurs principaux types et caractéristiques, ainsi que quelques exemples de tels systèmes. Nous terminons le chapitre par une conclusion qui discute les défis de ce type de systèmes ainsi que leur faiblesse.

## II. Définition

Pour définir un système de gestion d'identité il faut définir premièrement le terme identité numérique :

### II.1. Identité numérique

Une identité numérique peut être définie comme un ensemble de données numériques qui représentent de façon unique une entité dans un domaine d'application. Dans ce contexte, une entité peut être une personne, un ensemble de personnes, une organisation, un processus ou un dispositif, c'est-à-dire, tout objet capable de faire une transaction. Les éléments composant une identité numérique sont nommés Attributs, lesquels peuvent être assignés, intrinsèques à l'entité ou dérivés. Certains attributs distinguent de manière unique une identité numérique dans un contexte d'espace de noms: ils sont connus comme identifiants. Généralement, un identifiant est utilisé pour réaliser une authentification (c.-à.-d. Valider l'identité). L'ensemble des éléments servant de preuve à l'authentification est appelé « *credentials* ». Un « *credential* » peut se présenter sous la forme d'un mot de passe, d'une réponse à un défi (quelque chose que l'on sait), d'une information fournie par une carte à puce ou un certificat numérique (quelque chose que l'on a), ou d'une information dérivée des caractéristiques de la personne, comme l'empreinte digitale, l'iris ou le timbre de la voix (ce que l'on est). La Figure II.1 montre les relations qui existent entre les éléments composant une identité numérique <sup>[17]</sup>.

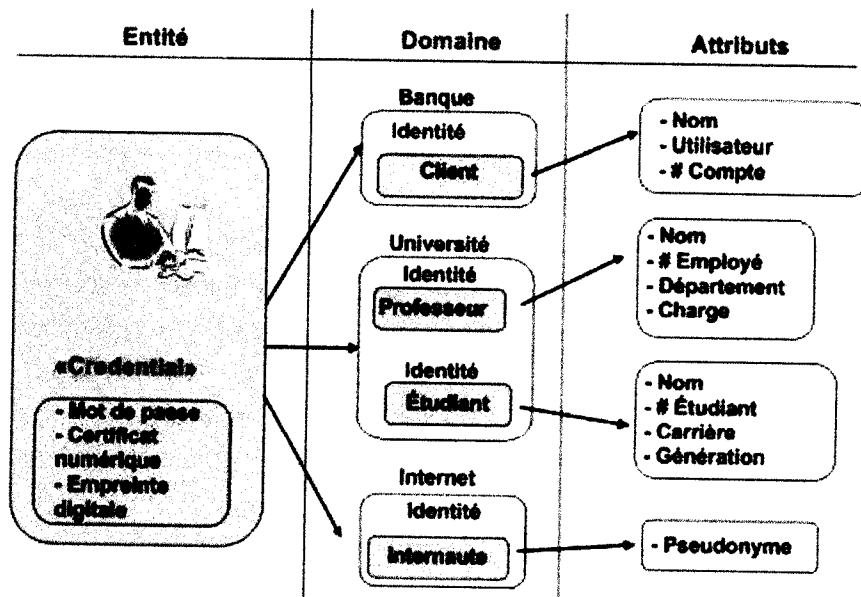


Figure II.1 : Relation entre les éléments composants une identité numérique. [17]

Dans la Figure II.1, l'entité a plusieurs identités numériques nécessaires pour interagir avec différents domaines d'application. Un domaine d'application est un contexte où une même identité numérique est valide, par exemple, une entreprise, un hôpital, un club sportif, une université ou l'Internet. On constate qu'une entité peut avoir plusieurs identités dans un même domaine d'application. Par exemple, à l'université, un professeur peut avoir simultanément l'identité de professeur et celle d'un étudiant dans le cas où il suit un cours en formation continue. Chaque identité est composée d'un ou plusieurs attributs et un « credential » associé.

## II.2. Système de Gestion d'Identités

On définit un système de Gestion d'identités (SGI) comme l'ensemble des processus permettant de gérer le cycle de vie d'une identité numérique, c'est-à-dire, sa création, sa manipulation et sa fin de vie. Le SGI s'occupe aussi des composants opérationnels, lesquels gèrent les différents aspects de la sécurité, c'est-à-dire, le processus d'authentification, le contrôle d'accès et l'audit. La Figure II.2 montre les composants du cycle de vie et les composants opérationnels d'un SGI [17].



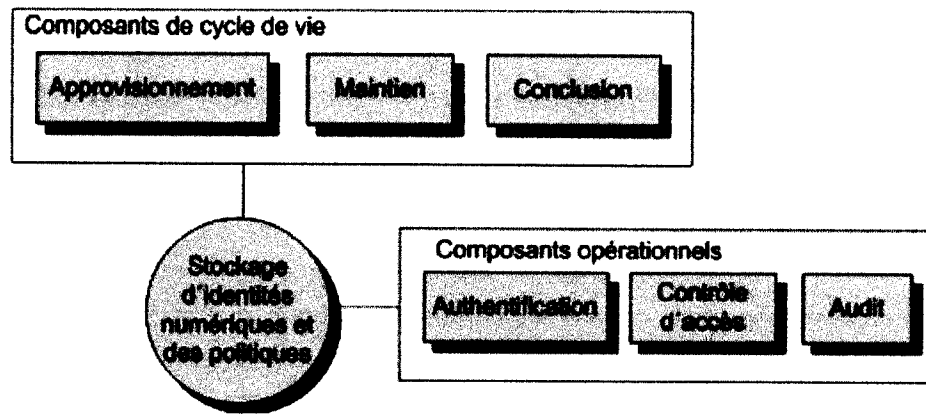


Figure II.2 : Composants et fonctionnalités d'un SGI. <sup>[17]</sup>

Il existe plusieurs modèles de SGI :

IV.4. *Modèle isolé, Modèle centralisé, Modèle centré utilisateur, Modèle fédéré, Role Based Access Control (RBAC), Dictionary Access Control (DAC), ...*

### III. Justifications de besoin d'un SGI

Dans cette partie nous allons expliquer les besoins d'utilisation des systèmes de gestion des entités, leur rôle ainsi que la manière dont ces systèmes répondent à des besoins de sécurité, de vie privée.

Plus un citoyen a des activités en ligne, plus il s'expose à des risques de sécurité ou de violation de sa vie privée. Les technologies d'identification permettront aux citoyens de gérer ces risques et deviendront donc un élément essentiel de la communication Internet, quel que soit son objectif. C'est pourquoi l'adoption de ces technologies aura des répercussions directes sur l'adoption de l'ensemble des services de la société de l'information <sup>[14]</sup>.

Si nous avons besoin d'un référencement d'un nouvel utilisateur dans un Système d'Information, nous pouvons identifier les actions suivantes :

- ✦ embauche dans une entreprise ⇒ Référencement dans le système de gestion de la paie.
- ✦ attribution d'un bureau ⇒ référencement dans la base du service logistique.
- ✦ attribution d'un numéro de téléphone ⇒ Référencement dans la base téléphonique.
- ✦ attribution d'un ordinateur, d'un identifiant et d'un mot de passe pour accéder au réseau ⇒ Référencement dans le système bureautique.

- ✚ attribution d'un badge pour l'accès aux locaux ⇒ Référencement dans le système de gestion des badges.
- ✚ droits d'accès à un restaurant d'entreprise ⇒ Référencement dans la base du restaurant d'entreprise.
- ✚ droits d'accès sur une application ⇒ Référencement dans la base de l'application.
- ✚ etc.

Dans la majorité des entreprises, ces opérations font appel à des annuaires qui ne sont ni compatibles entre eux, ni synchronisés (cf. Figure II.3 ci-dessous). Ainsi pour un nouvel utilisateur il faut saisir plusieurs fois les mêmes informations dans des systèmes différents par des personnes différentes et il en va de même en cas de modification d'une information. Cette mise à jour est parfois très longue ou que partiellement réalisée.

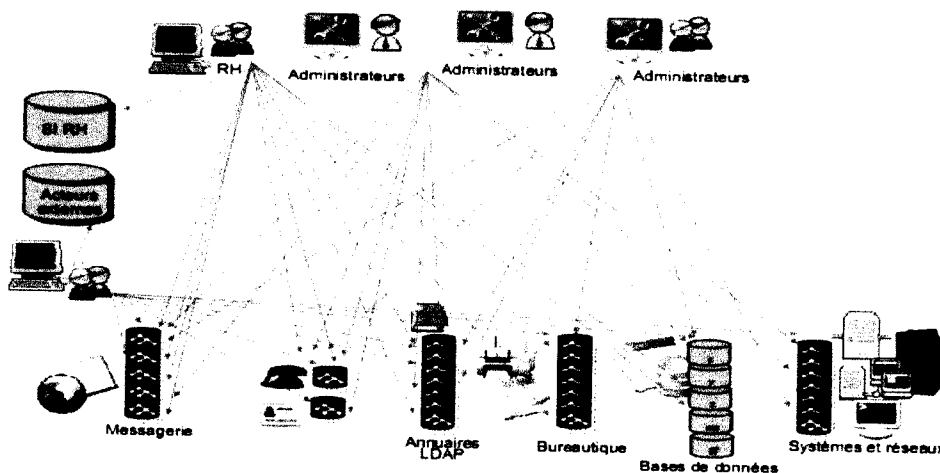


Figure II.3 : Standards Existants de la gestion des entités et des droits d'accès. [15]

L'absence de gestion globale des identités et des droits d'accès peut générer de nombreux problèmes, parmi lesquels :

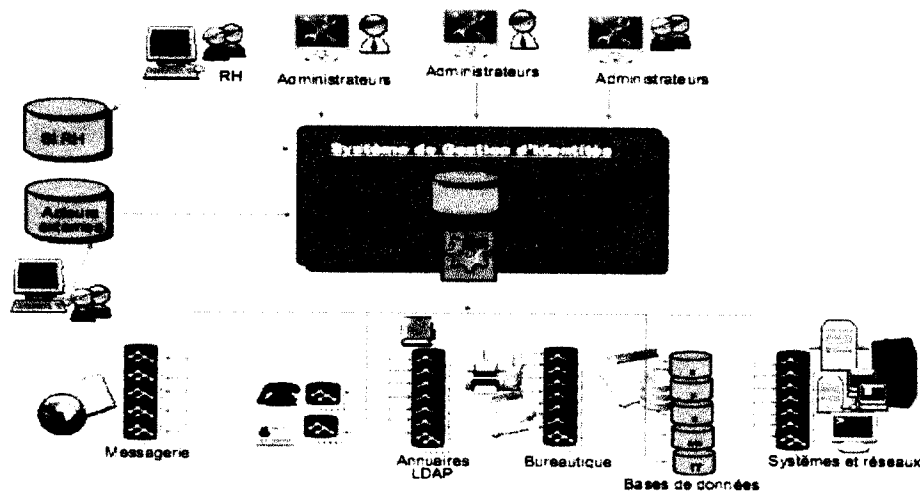
- ✚ la perte de productivité due aux délais d'obtention des droits d'accès.
- ✚ une charge importante d'administration (multiplication des administrateurs, réinitialisation des mots de passe, etc.).
- ✚ l'impossibilité de tracer les actions d'administration des droits et d'en contrôler la cohérence et la pertinence.
- ✚ la difficulté d'auditer les accès aux ressources.

- ✚ des entorses au principe de séparation des tâches.
- ✚ le non respect des contraintes légales et/ou réglementaires (par exemple au travers d'un mauvais paramétrage des règles de gestion).

La justification d'un projet de gestion des identités et des droits d'accès reposera sur les améliorations suivantes :

- ✚ garantie de traçabilité et d'auditabilité afin de répondre aux obligations légales et/ou réglementaires.
- ✚ repositionnement des « propriétaires fonctionnels » au centre du débat.
- ✚ réduction des coûts d'administration.
- ✚ amélioration de l'efficacité et de la réactivité.
- ✚ amélioration de la sécurité (adéquation des droits aux besoins métier).

L'ensemble des flux présentés dans la figure précédent (Figure II.3) va pouvoir être représenté de la manière suivante après la mise en place d'une gestion centralisée des identités (cf. Figure II.4)



**Figure II.4 :** Flux de mise à jour après la mise en place de système de gestion d'identité centralisée. <sup>[15]</sup>

Les informations sont mises à jour dans le référentiel central qui alimente ensuite automatiquement les annuaires ou bases de sécurité des différents environnements.

### III.1. Les principaux avantages d'utilisation d'un SGI

#### III.1.1. Garantie de traçabilité et d'auditabilité

Les principales lois et réglementations impliquant le SI et ayant des impacts directs sur les aspects traitant de la sécurité (en particulier traçabilité et auditabilité) « aspects juridiques ».

D'une manière générale, ces lois et règlements « imposent » au Système d'Information des exigences de :

- ✦ continuité d'activité.
- ✦ de séparation des tâches : par exemple, une même personne ne doit pas à la fois commander une fourniture ou prestation et valider sa réception.
- ✦ de traçabilité et d'auditabilité : permettant de valider « qui a fait quoi » au sein du système d'information, et « qui a habilité qui ».
- ✦ de respect de la vie privée.

Déroger à ces exigences peut entraîner un risque juridique pour les responsables de l'entreprise.

#### III.1.2. Réduction des coûts d'administration

Un système de gestion des identités et des droits d'accès permet d'alléger la charge de travail de l'équipe de « support informatique » (administration, help desk). Cet allègement résulte d'une part de l'automatisation de tâches de gestion de comptes (réduction du nombre d'administrateurs) et d'autre part de la diminution du nombre d'appels d'utilisateurs (perte ou oubli de nombreux mots de passe, relance de demandes d'accès, etc.).

Le système de gestion des identités peut permettre aux utilisateurs la gestion directe de certains aspects de leur profil (par exemple le mot de passe, l'adresse, les numéros de téléphone, etc.).

#### III.1.3. Amélioration de l'efficacité et de la réactivité

Un système de gestion des identités et des droits d'accès permet de réduire le nombre d'interventions humaines par une automatisation de la propagation des droits sur les

différents environnements concernés. La conséquence est à la fois une réduction des délais de mise à disposition des droits d'accès et une réduction des sources d'erreur (prise en compte systématique de tous les besoins liés à l'activité de l'utilisateur, garantie de cohérence dans les droits attribués).

Les gains générés concernent à la fois les utilisateurs internes (gain de productivité) et externes (amélioration de la qualité du service et de l'image de l'entreprise). Sur un autre plan, lors d'une fusion ou d'une acquisition, il faut fournir le plus rapidement possible un accès aisé aux ressources rassemblées d'entreprises auparavant autonomes. Là encore, une solution de gestion des identités et des droits d'accès aidera à relever ce défi au travers d'un service d'intégration des informations multi-plates-formes permettant de connecter les systèmes de chaque entreprise à la plupart des systèmes (nouveaux ou préexistants) de la nouvelle entité.

#### **III.1.4. Amélioration de la sécurité**

Un système de gestion des identités et des droits d'accès permet de renforcer la sécurité. Une telle approche conduit à établir des liens entre toutes les applications, bases de données et annuaires en s'appuyant sur des notions de rôle et de profil. Cette solution offre un point unique de gestion des règles de sécurité pour l'ensemble des systèmes concernés. Elle permet de créer simplement des règles d'accès et de sécurité, en cohérence avec la Politique de Sécurité des Systèmes d'Information et les besoins métier, puis de les propager automatiquement à tous les systèmes de l'entreprise.

La gestion centralisée des identités permet d'éliminer une source considérable d'erreurs d'administration pouvant causer des failles de sécurité d'accès au SI de l'entreprise. Elle permet également de résilier complètement et immédiatement les droits d'accès sur l'ensemble des systèmes lorsque des salariés ou personnels extérieurs quittent l'entreprise ou changent d'affectation et supprimer ainsi les comptes « fantômes ».

En mettant en place des processus maîtrisés d'habilitation, le système permet d'impliquer les responsables métiers dans le circuit d'habilitation et de ne plus laisser au seul administrateur technique la maîtrise des droits d'accès <sup>[15]</sup>.

#### IV. Les Modèles de SGI

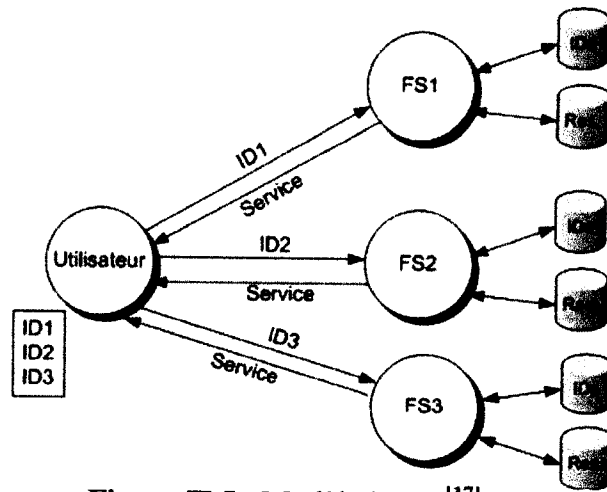
V.4. Les SGI ont beaucoup évolué. A ces débuts, chaque département dans une organisation gérait ses propres identités sans aucune interaction facile possible (modèle isolé). Plus tard, des solutions centralisées ont été implémentées afin de n'avoir plus qu'à gérer qu'une seule identité de l'utilisateur (modèle centralisé). Aujourd'hui, il est nécessaire de gérer plusieurs identités dans un environnement distribué et collaboratif, à cette fin, quelques modèles ont été proposés : Modèle isolé, Modèle centralisé, Modèle fédéré, Modèle centré utilisateur.

##### IV.1. Définitions

- ✚ **Utilisateur** : Entité qui est représentée par une identité numérique et capable de réaliser une transaction.
- ✚ **Fournisseur de Service (FS)** : Entité qui fournit un service aux utilisateurs. il s'agit usuellement d'un site Web ou d'un service Web « Web Service ».
- ✚ **Fournisseur d'identité (FI)** : Entité qui gère l'identité numérique de l'utilisateur et qui exécute le processus d'authentification.

##### IV.2. Le modèle isolé

Dans ce modèle, chaque FS a la responsabilité de gérer l'identité de chacun de ses utilisateurs. Le FS déploie sa propre SGI en prenant en compte la complexité et les fonctionnalités définies par l'organisation. Il s'avère très difficile, pour les FS, d'intégrer ces SGI afin de fournir des services coordonnés. De même, ce modèle devient assez lourd pour l'utilisateur lorsque le nombre d'identités à gérer augmente. En ce qui concerne le respect de la vie privée. Le FS a le contrôle total de l'information des utilisateurs. Ceux-ci ont peu ou aucun contrôle sur les données personnelles gardées par le FS. La Figure II.5 montre les interactions entre un utilisateur et les FS pour le processus d'authentification et la fourniture de services.

Figure II.5 : Modèle isolé. <sup>[17]</sup>

#### IV.3. Le modèle centralisé

Ce modèle repose sur le stockage unique des identités numériques. L'utilisateur peut s'authentifier avec tous les FS en utilisant la même identité. Il est donc assez simple d'implémenter la fonctionnalité de « Single Sign On » où l'utilisateur, une fois authentifié par le FI, peut accéder à plusieurs FS, et ce, sans authentification supplémentaire. Le modèle centralisé ne nécessite aucune expertise préalable de l'utilisateur du fait de la grande simplicité d'accès à plusieurs services avec une seule identité. Cependant, ce modèle ne manque pas d'inconvénients, parmi lesquels on peut souligner sa vulnérabilité du fait que le stockage centralisé d'identités représente un point unique de défaillance. Il ne passe pas à l'échelle lorsque le nombre d'identités devient très grand. La Figure II.6 montre comment un utilisateur peut utiliser la même identité pour accéder à plusieurs FS.

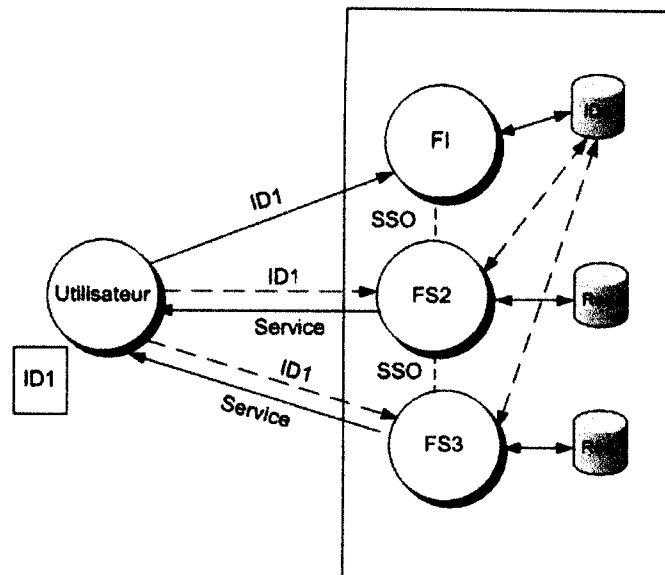


Figure II.6 : Modèle centralisé. [17]

#### VI.4. Modèle fédéré

Dans le modèle fédéré, les identités sauvegardées dans les différents FS sont liées au travers de pseudonymes. Les entités qui composent la fédération forment un Cercle de Confiance (CC) en établissant des relations de confiance avec des accords commerciaux et une plateforme technologique commune. Le modèle fédéré définit des services d'identité tels que le « Single Sign On », la fédération d'identités et l'échange d'attributs. Tout comme dans le modèle précédent, une fois l'utilisateur authentifié avec le FI, il peut avoir accès aux services fournis par d'autres FS sans authentification supplémentaire. Comme le stockage est distribué, il n'a pas de point unique de défaillance, ni de limitation en principe quant au facteur d'échelle. Par contre, dans ce modèle, l'utilisateur n'a pas de contrôle sur ses données personnelles et il n'a aucune garantie quant au respect de sa vie privée. La Figure II.7 montre les éléments d'un système fédéré et les relations entre eux.



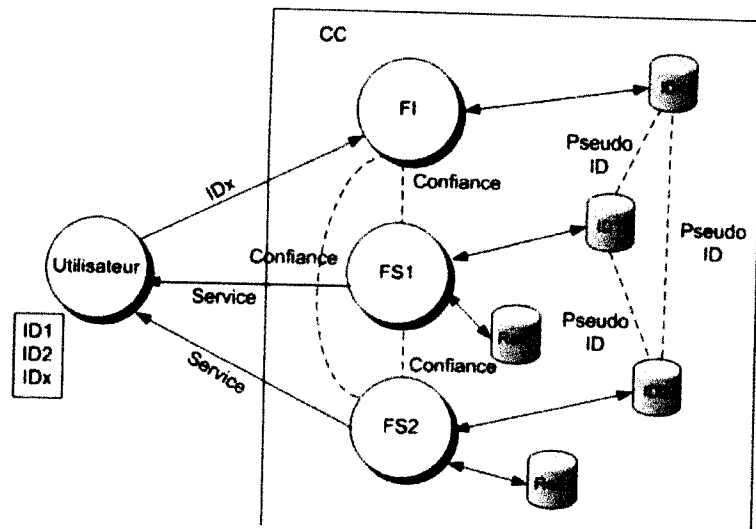


Figure II.7 : Modèle fédéré. [17]

#### IV.4.1. Architecture d'Identité Fédérée

Une Architecture d'Identité Fédérée (AIF) est composée d'un ensemble d'organisations qui ont établi des relations de confiance entre elles afin d'échanger des données de manière sûre, tout en préservant l'intégrité et la confidentialité de l'information. Dans ces travaux, on s'intéresse aux données personnelles. Le FI dans l'AIF gère l'information d'identité de l'utilisateur et fait le processus d'authentification. Il peut y avoir un ou plusieurs FI dans le Cercle de Confiance. L'AIF doit accomplir les fonctionnalités suivantes du point de vue des utilisateurs, fournisseurs d'identités et fournisseur de services :

- ✦ « Single Sign On, » : SSO permet aux utilisateurs de s'authentifier avec un FI et d'accéder aux services fournis par plusieurs FS sans avoir besoin de s'authentifier à nouveau.
- ✦ Fédération d'identités : il s'agit de l'union de deux identités numériques au travers d'un pseudonyme pour implémenter les services de SSO et d'échange d'attributs.
- ✦ Échange d'attributs : Le FS peut demander des attributs additionnels au FI pour fournir des services personnalisés.

Quand les identités sont fédérées, un identifiant est créé pour chaque couple de FI et FS dans le but de lier les deux identités. L'identifiant peut être dynamique, c'est-à-dire, il est créé à chaque nouvelle session de l'utilisateur ou il peut être fixe pendant une

longue période de temps. Un identifiant de type pseudonyme permet de préserver l'identité réelle de l'utilisateur et de mieux respecter sa vie privée. L'accord commercial établit la manière dont les identités sont fédérées, c'est-à-dire, la structure du pseudonyme, si l'identificateur est permanent ou dynamique, et les attributs échangés. La Figure II.8 montre la fédération d'identités qui utilise un pseudonyme fixe. Le tableau des identités de FI1 montre comment l'identité ID1 est associée à l'identifiant aléatoire 65ER4589 quand elle est fédérée avec l'identité ID2 de FS1.

Simultanément, le tableau des identités de FS1 montre la relation existante entre l'identité locale ID2 et l'identificateur 65ER4589. Le pseudonyme a une couverture locale : le FI et le FS ne connaissent que le compte local et le pseudonyme. Quand deux entités ont besoin d'interagir pour échanger des informations d'identité, ils utilisent le pseudonyme pour la référencer.

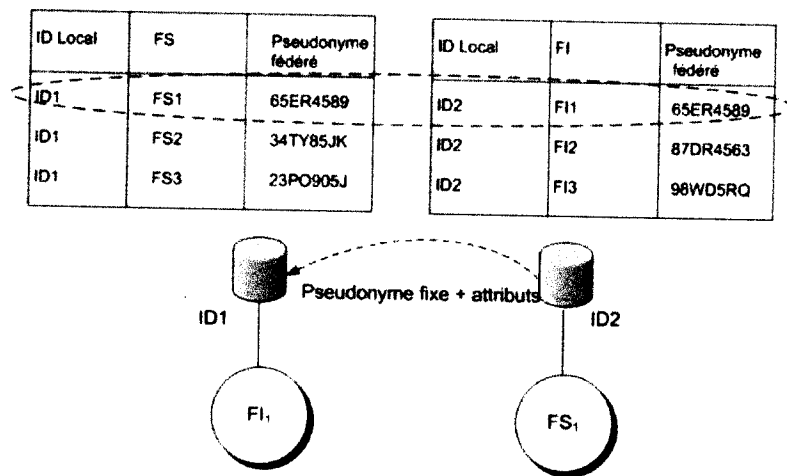


Figure II.8 : Fédération d'identités avec pseudonymes fixes. [17]

#### IV.4.2. Échange d'attributs dans un système d'identité fédéré

Dans le système d'identité fédéré, l'échange de données personnelles (attributs) peut se produire entre n'importe quelle entité du cercle de confiance (utilisateurs, FS et FI). Quand l'utilisateur participe à l'échange d'attributs, il peut décider à quel destinataire les fournir et sous quelles conditions. Malheureusement, le scénario le plus répandu et celui qui pose un vrai risque de non respect de la vie privée a lieu quand le FS demande les attributs au FI afin de personnaliser et ainsi d'améliorer

le service fourni. Dans ce cas, l'utilisateur ne peut pas contrôler ses données personnelles publiées par le FI.

La Figure II.9 montre un flux possible d'informations pendant le processus d'authentification et d'échange d'attributs entre le FI et le FS dans l'AIF. Dans ce scénario, le FI et le FS détiennent chacun des données personnelles liées à l'utilisateur, celles-ci sont fédérées par un pseudonyme. Le processus commence quand :

1. l'utilisateur s'authentifie avec le FI en suivant n'importe quelle méthode d'authentification définie par lui.
2. Si l'authentification réussit, le FI donne à l'utilisateur un jeton d'identité.
3. Avec l'information d'authentification et un pseudonyme qui est employé pour accéder aux services fournis dans le cercle de confiance. L'utilisateur demande un service du FS et présente le jeton donné par le FI.
4. le FS valide le jeton d'authentification avec le FI.
5. puis traduit le pseudonyme en l'identité locale afin de pouvoir le service correspondant.
6. Si le service demandé a besoin d'attributs supplémentaires qu'il n'a pas, ils sont demandés au FI.
7. la pseudo-identification est employée pour référencer l'utilisateur. Les attributs sont envoyés au FS.
8. le service est fourni.

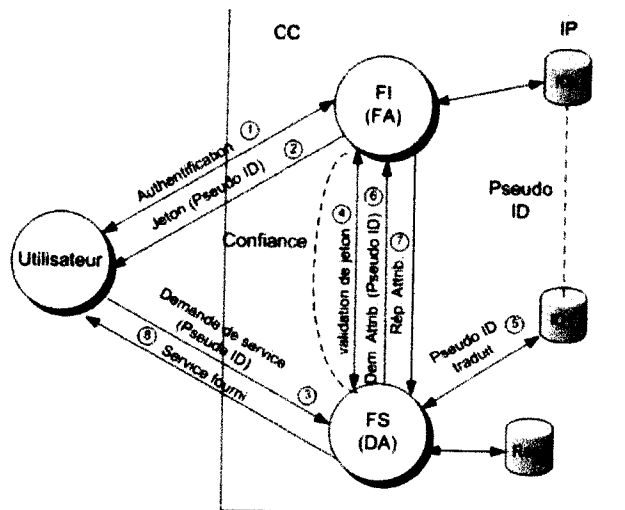


Figure II.9 : Un possible échange d'attributs dans une architecture d'identité fédérée. [17]

#### IV.4.3.Exemples des modèles fédérés

Les cas d'utilisation dans cette partie expliquent comment la fédération peut fournir une parfaite expérience des utilisateurs finaux en authentifiant une fois pour de multiples applications.

##### IV.4.3.1 Shibboleth

Shibboleth est développé depuis 2001 par Internet2 et désigne à la fois une norme et un produit libre. Shibboleth est un système Web SSO qui permet de mettre en œuvre des fédérations d'identité dans le cadre des OV<sup>1</sup>. Shibboleth est une extension de SAML offrant un langage standardisé pour l'échange de l'information de sécurité entre des domaines distribués hétérogènes. En effet, Shibboleth enrichit les fonctionnalités de fédération d'identités de SAML en facilitant pour un ensemble de partenaires la mise en place de deux fonctionnalités importantes, la délégation d'authentification et la propagation d'attributs (Figure II.10).

Shibboleth propose une architecture qui permet aux membres d'une OV de s'authentifier une fois auprès de leur organisation mère, puis d'atteindre différents

<sup>1</sup> Organisation Virtuelle (OV), est une alliance temporaire d'organisations, autour d'une structure réseau, où la mutualisation de ressources matérielles, logicielles, humaines permet d'atteindre un objectif commun.

systèmes et services de l'OV, physiquement distribués entre les sites, sans pour autant avoir besoin de procéder à chaque fois à une nouvelle authentification. L'extension majeure étant de pouvoir interconnecter des services à la base indépendants structurellement et juridiquement.

De plus, Shibboleth donne la possibilité, à une organisation tierce de guider les décisions relatives à des accès que pourrait faire un utilisateur, en considérant la valeur des attributs que lui a accordé l'organisation mère dont il dépend.

#### IV.4.3.1.1 Architecture

Shibboleth identifie six modules de base (Figure II.10) qui interagissent afin de mettre en œuvre la fédération d'identité entre des domaines de sécurité distribués. Du côté FI, Shibboleth définit deux modules:

- ✓ le service de handles responsable du traitement de la fonctionnalité SSO
- ✓ l'autorité d'attributs responsable du traitement (affectation et gestion) des attributs des utilisateurs au sein d'une organisation.

Du côté FS, Shibboleth définit trois modules:

- ✓ le module SHIRE responsable de l'affectation d'un identifiant à un utilisateur que l'FI et le FS peuvent utiliser pour faire référence à lui.
- ✓ le module SHAR responsable de la récupération des attributs d'un utilisateur.
- ✓ le module ShibAuthz responsable de la prise de décision d'autorisation d'accès à une ressource demandée.

Finalement, le module WAYF permet d'orienter l'utilisateur vers son fournisseur d'identités afin de s'authentifier.

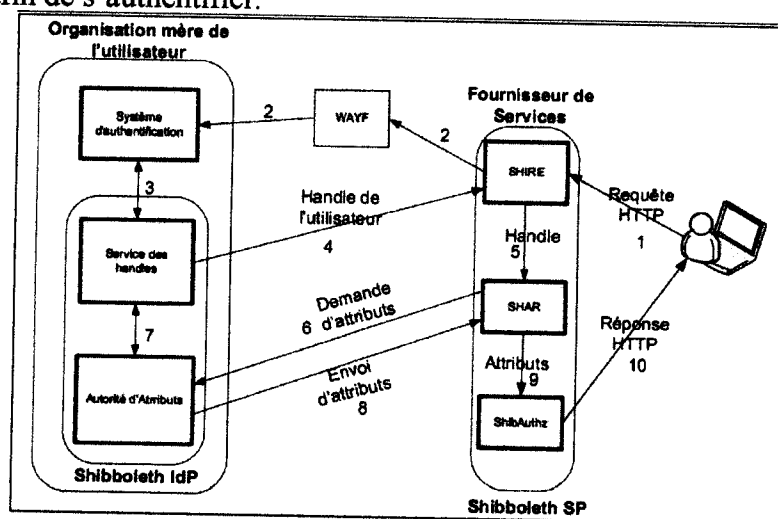


Figure II.10 : Les composants de Shibboleth.

#### IV.4.3.2. Liberty Alliance

Connu sous le nom de « Project Liberty », est un projet qui réunit des acteurs des mondes industriel, informatique, bancaire et gouvernemental sous la forme d'un consortium. L'objectif est de définir un ensemble de spécifications de protocoles de fédération d'identité et de communication entre services web.

Le projet « Liberty Alliance » a été initié par la société Sun en septembre 2001 afin de proposer une alternative au projet Passport de Microsoft. Depuis sa création, ce projet a été rejoint par plusieurs centaines d'acteurs des mondes industriel, informatique, bancaire et gouvernemental tel que : France Telecom, Intel, Nokia, Sony, Vodafone ....

Le but de ce consortium est de fournir une infrastructure de standards visant à fédérer la gestion d'identités électroniques entre plusieurs services ou systèmes en respectant la vie privée de l'utilisateur.

##### IV.4.3.2.1. Objectifs

Les principaux objectifs de ce consortium sont les suivants :

- ✚ Laisser la possibilité à chaque client de contrôler toutes les informations qui le concernent en réseau (son profil). En effet, toute modification d'attributs d'un utilisateur doit être précédée par son accord : « Liberty Alliance » permet de coupler les exigences d'une authentification forte avec le respect de la vie privée des usagers.
- ✚ Réduire le coût d'administration pour les utilisateurs (contrôle de son Profil Personnel) et les entreprises.
- ✚ Créer une infrastructure de gestion de l'identité en réseau qui peut supporter, même à long terme tous les moyens de connexion à Internet.
- ✚ Fournir un standard ouvert de SSO qui permet à un utilisateur d'accéder à des services proposés par différents fournisseurs tout en utilisant un compte unique <sup>[18]</sup>.

IV.4.3.3. OpenID

OpenID est un système d'authentification décentralisé qui permet l'authentification unique, ainsi que le partage d'attributs. Il permet à un utilisateur de s'authentifier auprès de plusieurs sites (devant prendre en charge cette technologie) sans avoir à retenir un identifiant pour chacun d'eux mais en utilisant à chaque fois un unique identifiant OpenID. Le modèle se base sur des liens de confiance préalablement établis entre les fournisseurs de services et les fournisseurs d'identité (*OpenID providers*). Il permet aussi d'éviter de remplir à chaque fois un nouveau formulaire en réutilisant les informations déjà disponibles. Ce système permet à un utilisateur d'utiliser un mécanisme d'authentification forte [19].

On va l'étudier de façon plus spécifique dans le prochain chapitre.

IV.5. Modèle centré utilisateur

Ce modèle a été proposé dans le but de donner aux utilisateurs plus de contrôle sur ses données personnelles. Il peut en effet sélectionner le FI qui lui convient et choisir l'identité à utiliser pour accéder aux différents FS. Les FS n'établissent pas de relation de confiance entre eux pour fournir des services à l'utilisateur. Ce modèle permet à l'utilisateur de mettre en place son propre FI dans son ordinateur ou son portable. Parmi les défauts de ce modèle, on trouve la difficulté d'intégrer les FS pour fournir des services coordonnés, et le fait que le dispositif de l'utilisateur devient un point unique de défaillance. La Figure II.11 montre les éléments d'un système centré utilisateur et les relations entre eux.

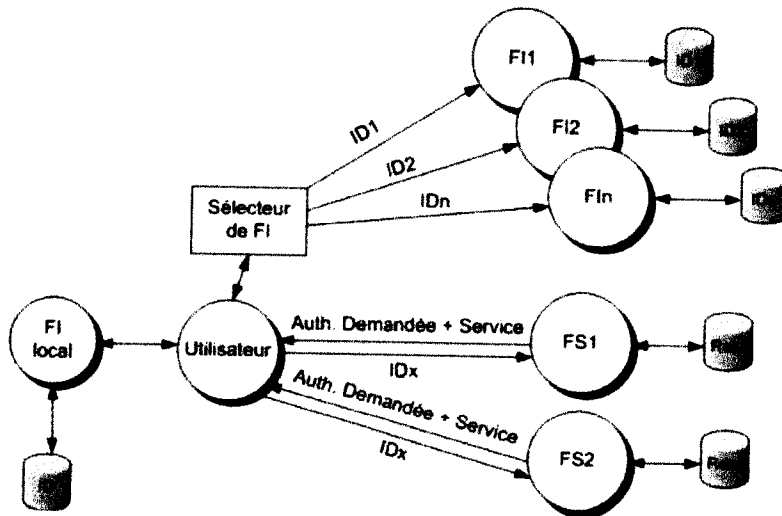


Figure II.11 : Modèle centré dans l'utilisateur. [17]

## V. La gestion d'identité et la vie privée

La gestion des identités est un domaine technologique récent dont le but est d'aider les utilisateurs de l'Internet à gérer leurs relations avec les fournisseurs de services, notamment en prouvant qu'une personne est autorisée à accéder à des ressources spécifiques (par exemple, un compte client). Ces technologies ont un impact majeur sur la vie privée et peuvent être conçues de manière à faciliter le traçage et la surveillance centralisée de toutes les activités en ligne et hors ligne d'un individu, ou à réduire à un strict minimum les données à caractère personnel qui sont divulguées à des secondes ou tierces parties, permettant ainsi aux individus de bénéficier du même niveau de protection de la vie privée sur l'Internet que dans le monde réel <sup>[16]</sup>.

### V.1. Problématique

Les **données personnelles** d'un individu englobent une quantité non-négligeable et importante d'informations plus ou moins nominatives (nom, prénom, âge, sexe, lieu de résidence, loisirs préférés, pseudo, n°client, etc.), force est de constater que bon nombre de personnes ignorent précisément de quoi il s'agit, mais aussi par qui et dans quel but des **fichiers** sont créés.

S'il est aisé d'imaginer que nous sommes tous fichés par l'Etat et les organismes qui lui sont rattachés (sécurité sociale, fisc, police à travers la carte nationale d'identité, la préfecture lors de l'établissement de la carte grise, le Pôle emploi, le médecin, etc.), par son employeur, par des associations indépendantes (club de sport, association à laquelle on fait un don, forum de discussion ou chat, etc.) ou encore par des sociétés commerciales (banque, assureurs, téléphonie, fichiers clients des commerces, etc.), on imagine moins être fichés par des sociétés que l'on ne connaît pas. Les fournisseurs de services, que ce soit l'administration ou des opérateurs de service sont amenés à collecter et stocker des informations personnelles dans le but de fournir le service.

Pour assurer ces informations qui ne soient pas à la main des mauvaises sociétés publicitaire ou des revendeurs, et de garder toujours aux clients le choix et le seul responsable de ces informations personnelle, et aussi pour faciliter le style des



inscriptions et d'abonnements avec ces différents services, nous définissons des protocoles et des systèmes de protection et de gestions d'identité.

**V.2. Les systèmes qui prennent en charge la vie privée**

Dans cette section on va décrire un peu d'IMetSs (identité méta systemes), comme CardSpace et Higgins, comme prometteur IMS (OpenID).

Les utilisateurs sont les seuls qui utilisent IMetSs. Les FIs sont des entités de confiance qui fournissent et garantissent les informations et les identités des utilisateurs. Les FSs sont les entités qui ont besoin de dépenser les informations et les identités des utilisateurs fournis par FIs. Par exemple, un site web d'une magasin en ligne (FS) ont besoin de connaître l'identité d'utilisateur et les informations de paiement dépendent de la banque d'utilisateurs (FI) avant que la transaction soit complétée.

**V.2.1. Microsoft's CardSpace**

Windows CardSpace est un composant de Microsoft .NET Framework version 3.0 (ex-WinFX) qui apporte une expérience utilisateur cohérente requise par le métasystème d'identité. Windows a été spécialement renforcé pour lutter contre toute altération et mystification. Et ce dans le but de préserver les identités digitales de l'utilisateur et d'en assurer la maîtrise par ce dernier.

L'autorité de garantir les identités et les informations des utilisateurs (appelé aussi « Managed cards »). Les flux générales de CardSpace système ce représentent dans la figure suivants :

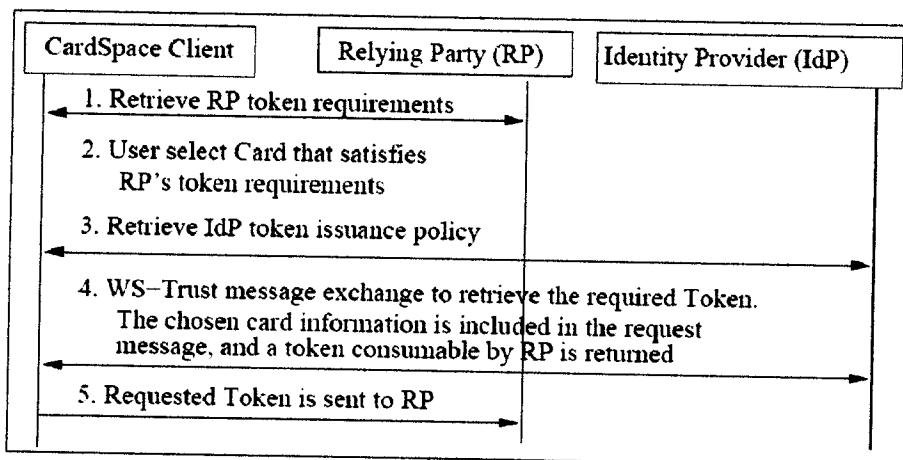


Figure II.12 : CardSpace Flux Simplifié.

### V.2.1.1. Principes

Les données sont stockées localement par un gestionnaire des mots de passe (Windows Live ID) sur un serveur (annuaire d'entreprise ou serveur compatible WS-Federation).

Windows CardSpace est livré avec le .NET Framework 3.0. Il est donc installé par défaut avec Windows Vista et disponible en téléchargement pour Windows XP et Windows Server 2003. Une mise à jour pour .NET 3.5 existe.

La communauté Open Source Eclipse a pris exemple sur CardSpace pour le framework « Higgins » supporté par IBM et Novell.

Ce framework s'est ouvert aux technologies concurrentes telles que CardSpace et OpenID.

### V.2.1.2. Sécurité

CardSpace est basé sur fichier d'échange XML et un protocole sécurisé compatible Web Service ( WS-Security, WS-Trust, WS-MetadataExchange et WS-SecurityPolicy).

CardSpace a accès à ces données (professionnelle, personnelle, administrative, adresse) par un code pin et dans un mode sécurisé pour contrer les chevaux de Troie ou les keyloggers.

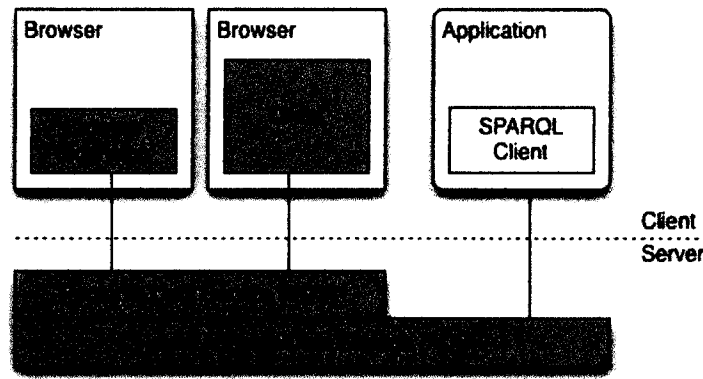
### V.2.2. Higgins Project centré utilisateurs

Higgins est une open source identité framework, au début développée par le projet de Fondation Eclipse (Eclipse 2008) qui permette aux utilisateurs et applications d'intégrer l'identité, profile, et les informations des relations sociales à travers les sources des données multiple et protocoles.

Depuis la réalisation de Higgins version 1.0 au février 2008, plusieurs producteurs commerciale sont basé sur Higgins qui sont annoncé par Novell, CA et IBM. La version courant de Higgins 2.0 à un nouveau code qui est basé sur l'implémentation des PDS

Comme on voit dans la Figure II.13 Higgins suit le model de gestion d'identité centré utilisateur qui avait comme but principal :

- Fournir une expérience sécurisée est consisté à l'authentification de la base d'utilisateurs sur la carte d'identité (I-cards).



2.0.103

Figure II.13 : Base d'implémentation Higgins 2.0.

- Construire une infrastructure confiance qui permette aux personnes de partager ces informations personnelles au temps que leur vie privée est protégée.
- De définir plusieurs fournisseurs plug-ins qui permettent aux développeurs de créer et d'adapter aux systèmes hérités, protocoles et types format dans le sens d'intégrer le système d'identité.
- Pour fournir les plug-ins il faut l'adaptés pour supporter les différentes sources des données.

#### V.2.2.1. Architecteur et opération de Higgins

La Figure II.14 démontre que l'architecteurs d'Higgins à trois couches, les composent de l'inférieur sont Identité Attribué Service (IdAS) la couche qui fournit l'interopérabilité et la portabilité à travers silos de donnée identité. Les composent intermédiaire sont Identité Service (IdS) la couche qui supporte différent type de I-cards. Les composants de la couche supérieur contient l'application d'utilisateur tout avec les éléments opérationnelle de l'architecteur : Identité Selectors (IdSs), FI et Relaying Parties (FSs).

IdAS utiliser Contexte Data Model (CDM) pour fournir les données abstraction qui permettent la probabilité à travers les sources des données hétérogènes. Le model de donnée exprimé par les contextes qui sont des ensembles qui représente les personnes, groupes, organisations, objets, etc. le contexte est exprimé par les ontologies utilisant un langage standard comme Resource Description Framework (RDF) et Ontologie Web Langage (OWL). Chaque entité dans un contexte attribué par un seul ou des valeurs structuré. Le contexte fournit des cartes de source de données au contexte correspondant.

IdS est la couche responsable pour manipuler différents jetons de sécurité, Higgins 1.0 supporte SAML et Nom d'Utilisateur/Mot de passe (UN/PW) jetons de sécurité, les versions suivantes (Higgins 1.1 ...) supportent des jetons de sécurité additionnelles tels que Kerberos, X.509 et Indemix. Une I-carte est une pièce d'identité qui fournit n'importe quel type d'information personnelle. Higgins est compatible avec CardSpace de Microsoft et OpenID I-card protocole. [17]

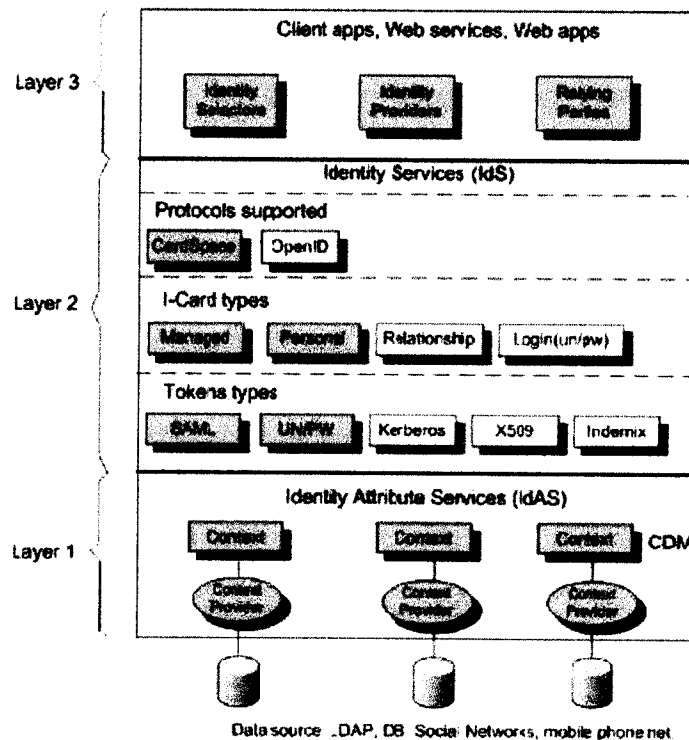


Figure II.14 : Architecture de Higgins. [17]

La couche supérieure accordée avec les applications et les fonctionnalités du Higgins. Les applications compatibles avec Higgins peuvent être client/serveur. Web ou web services. Identité sélecteur est une application centrée utilisateur pour la création, sélection, partagées et managées les I-Cartes qui représente l'identité dans différentes relations et contextes. Le poste Id sélecteur à sécuriser le jeton qui est validé par le FS. FI génère un jeton sécurisé qui contient les informations personnelles, telle que les informations réalisées et sélectionnées par l'utilisateur à FS via FI.

### V.2.3. PseudoID

PseudoID est modélisé d'être un seul sens, consistant, et non-chainable système de login fédéré. Il consiste d'un jeton de service utilisé au temps de setup, et un fournisseur

d'identité privée utilisée pour sign-ons. L'utilisateur a un compte avec un jeton de service, qui doit être persistant, l'identité « réelle » comme un email adresse. Au temps de setup, les logs d'utilisateur sur le service de jeton utilisé un système d'authentification familière, comme le nom et le mot de passe.

Après l'utilisateur demande un jeton d'accès de service de jeton qui est engagé au pseudonyme désiré. Au moment de connecter au FS, l'utilisateur doit présenter ce jeton au FI. L'FI va vérifier l'authentification du jeton et retourner le pseudonyme d'utilisateur au FS.

Pour être non-chainable, le jeton d'accès doit être généré comme si les deux jeton de service et FI sont compromis, l'identité « réel » d'utilisateur avec le jeton de service ne peut pas être chaîné à leur pseudonyme dans différents FSs. PseudoID accomplit cet propriété à l'aide d'utilisation Blind signature qui va mentionner au prochain chapitre.

## **VI. Conclusion**

Les solutions précédentes peuvent sécuriser la vie privée mais pas d'une façon totale, il reste toujours des liaisons entre les Fournisseurs des Services et les Fournisseurs d'Identités (les liens de confiance), ce qui permet à ce dernier de faire des liaisons entre les services et les clients, et connaître leurs centres d'intérêts, leurs préférences et habitudes. Dans les cas extrêmes les informations personnelles des clients peuvent être vendues à d'autres tiers sans son accord.

Le prochain chapitre présente notre solution pour un système de gestion d'identité. La solution proposée essaie d'éliminer les éventuelles liaisons entre le fournisseur d'identité et le fournisseur de service, ce qui augmente le niveau de confidentialité et protège mieux la vie privée des utilisateurs.

## I. Introduction

Dans ce chapitre nous décrivons les détails de conception et d'implémentation d'un système de gestion d'identité. Ce système est une amélioration des solutions existantes en termes de protection de la vie privée des utilisateurs.

Nous commençons par l'introduction de la problématique des solutions existantes, en particulier l'OpenID, le système le plus utilisé. Après, nous expliquerons notre approche pour avoir une vision claire sur son fonctionnement et son but. En fin, nous décrivons en détail les tests d'utilisation, et les outils de développement.

## II. OpenID

### II.1. Mode d'utilisation

#### La création d'un compte

Chaque utilisateur est identifié par un URI, qu'il acquiert auprès de son fournisseur d'identité OpenID. Le mode de fonctionnement est le suivant :

L'utilisateur choisit un fournisseur d'identité OpenID, par exemple VeriSign PIP ou Google

L'utilisateur choisit un nom d'utilisateur (Davide pour l'exemple), un mot de passe et crée son compte. L'utilisateur peut renseigner certaines informations le concernant. Ce compte aura pour identifiant OpenID : Davide.pip.verisignlabs.com. Le profil de l'utilisateur est également disponible à cette adresse.

#### L'Utilisation du compte

Le mode de fonctionnement est le suivant :

L'utilisateur Davide désire, par exemple, accéder au service UnServiceWeb qui requiert que l'utilisateur soit authentifié.

- L'utilisateur entre son identifiant : Davide.sub.domain.com.
- Le fournisseur de service (UnServiceWeb dans notre cas) contacte le fournisseur d'identité et ils créent ensemble un secret partagé.
- L'utilisateur sera redirigé vers le site du fournisseur d'identité et entre son mot de passe pour s'authentifier.
- Le fournisseur de service (UnServiceWeb) avertit l'utilisateur qu'il souhaite accéder à certains attributs de son profil OpenID (par exemple, son nom, prénom, et son adresse de courriel), il accepte ou non cette demande.
- A l'étape finale, l'utilisateur sera redirigé vers le fournisseur de service avec une preuve cryptographique de son identité (créée grâce au secret partagé préalablement établi) fournie par le fournisseur d'identité.

Tant que la session de l'utilisateur est active, il pourra être reconnu automatiquement sur les autres sites utilisant OpenID grâce au mécanisme de l'authentification unique. <sup>[19]</sup>

Le schéma suivant représente le fonctionnement de l'OpenID étape par étape entre les trois concepts (utilisateur, fournisseur d'identité FI, service consommateur FS) :

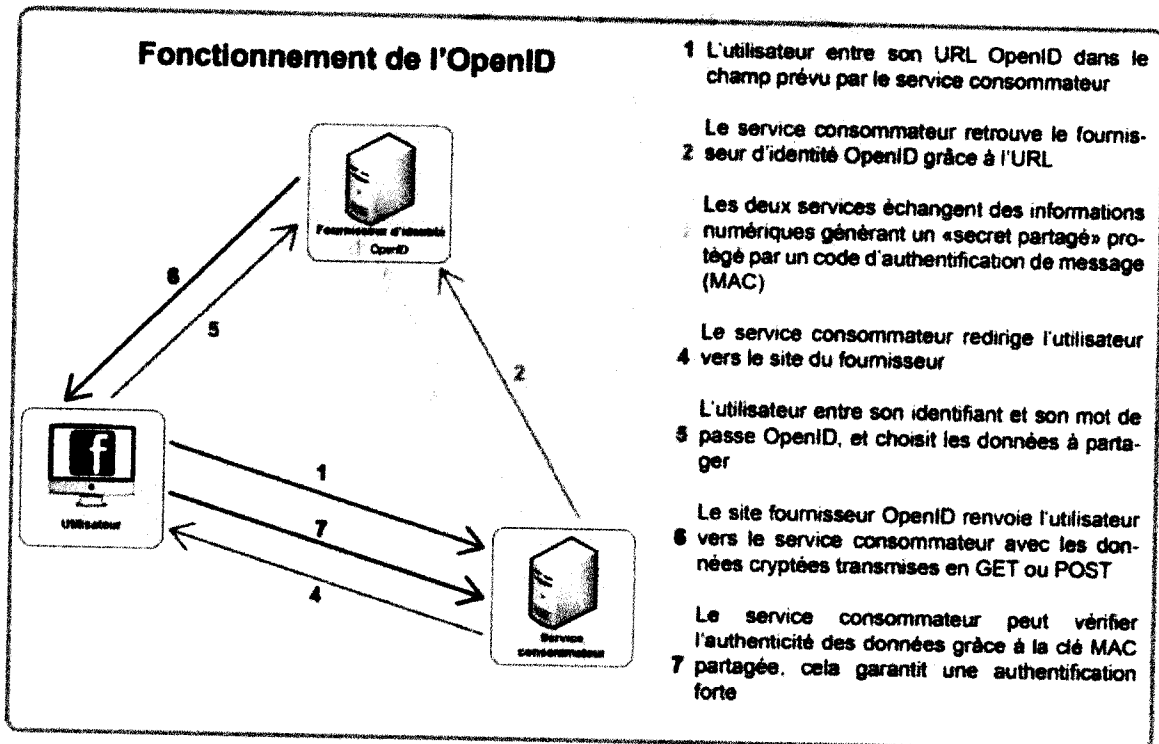


Figure III.1: Fonctionnement de l'OpenID.

## II.2. Problématique de l'OpenID

la figure III.2 suivant représente le domaine de confiance entre FS et FI. Un domaine de confiance désigne l'ensemble des fournisseurs d'identité, des fournisseurs de services et les liens de confiance établies entre eux (pour leur permettre l'accès contrôlé et sécurisé aux différents services) <sup>[20]</sup>.

La sécurité du système d'identifiant unique OpenID repose sur les liens de confiance qui existent entre les différents acteurs intervenants dans le mécanisme d'authentification. Ce la peut induit plusieurs risques et menaces sur la vie privée des utilisateurs, parmi ces risques on cite :

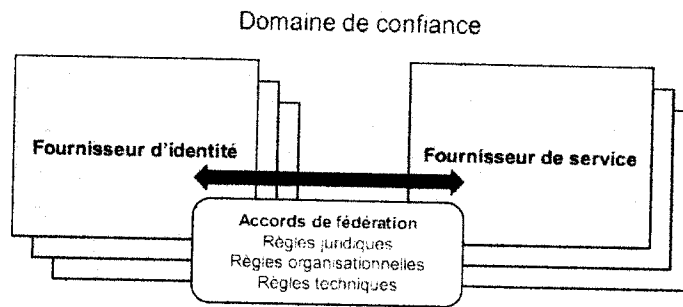


Figure III.2 : Les concepts clés de la fédération d'identité.

#### 👇 Le risque de hameçonnage

Une des faiblesses possibles d'OpenID est le risque d'hameçonnage, c'est-à-dire de détournement de l'utilisateur vers un autre site imitant son fournisseur OpenID habituel. Ce qui trompe l'utilisateur et donnera son mot de passe et d'autres informations personnelles au pirate.

#### 👇 Le risque des fournisseurs

Le Fournisseur de service doit connaître tout les fournisseurs d'identité de chaque utilisateur grâce à la liaison entre le fournisseur de service et d'identité (domaine de confiance).

Par conséquent le fournisseur d'identité peut connaître tous les sites Web qu'un utilisateur accède, les services utilisés, ses préférences et ces habitudes, ce la peut menacer pleinement sa vie privée.

### III. L'approche proposée

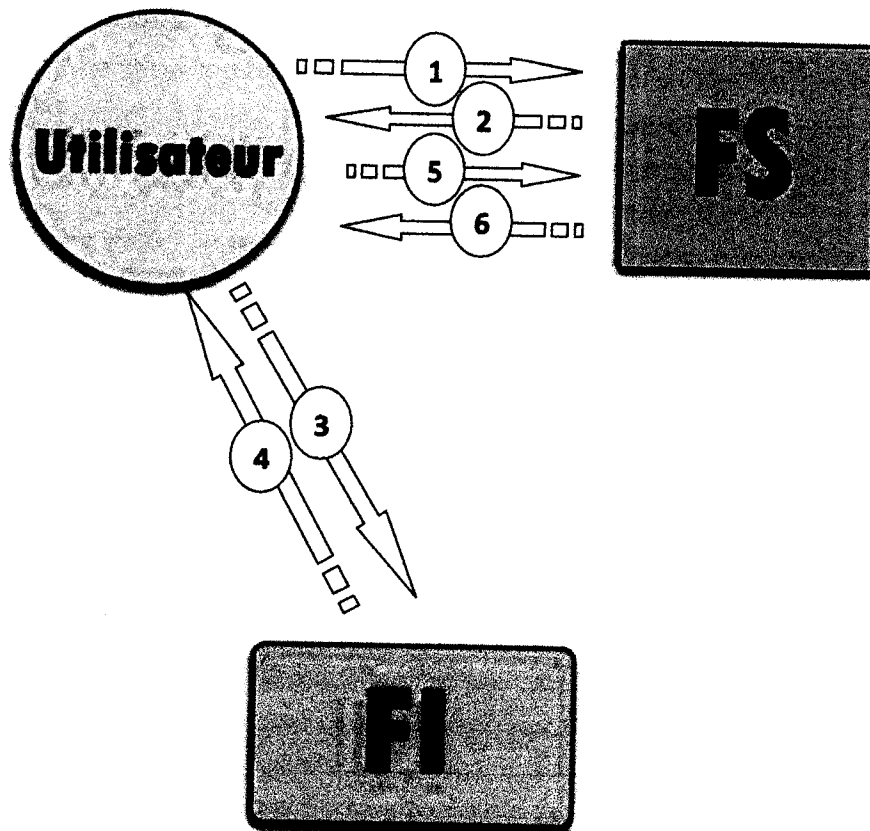
La solution proposée est une amélioration de l'architecture OpenID décrite dans la section précédente. Ce paragraphe introduit les détails de conceptuels de cette solution.

#### III.1. Généralité sur la solution

Le but de notre application est de résoudre le problème de sécurité de l'OpenID dont sa conception se repose principalement sur un lien entre les fournisseurs d'identité FIs et les fournisseurs des services FSs. Ce lien pose plusieurs problèmes et risques pour la sécurité et la vie privée des utilisateurs. Dans la solution proposée nous avons éliminé toute connexion ou lien direct entre le fournisseur d'identité et le fournisseur de service, toute communication se passe par l'intermédiaire de l'utilisateur.



La figure III.3 décrit l'architecture globale de la solution proposée:



**Figure III.3 :** Le déroulement des processus de façon globale du système proposé.

Selon la Figure III.3, le processus d'authentification dans le système proposée se déroule comme suit :

- 1-A partir du site du fournisseur de service, l'utilisateur choisit le fournisseur d'identité abonné avec ce fournisseur de service.
- 2- la confirmation de choix d'utilisateur à travers un jeton.
- 3- l'identification de l'utilisateur dans son propre fournisseur d'identité.
- 4- la création d'un jeton d'authentification d'utilisateur par le fournisseur d'identité.
- 5- la transmission du jeton d'authentification au fournisseur de service.
- 6- le fournisseur de services Analyse le jeton et offre le service à l'utilisateur si le jeton d'authentification est valide.

## III.2. Architecteur et fonctionnement

### III.2.1. Architecteur

Cette partie présente en détail l'architecture du système proposé. Nous introduisons le rôle très important que joue le navigateur dans le fonctionnement de notre système.

Le schéma détaillé est comme suit :

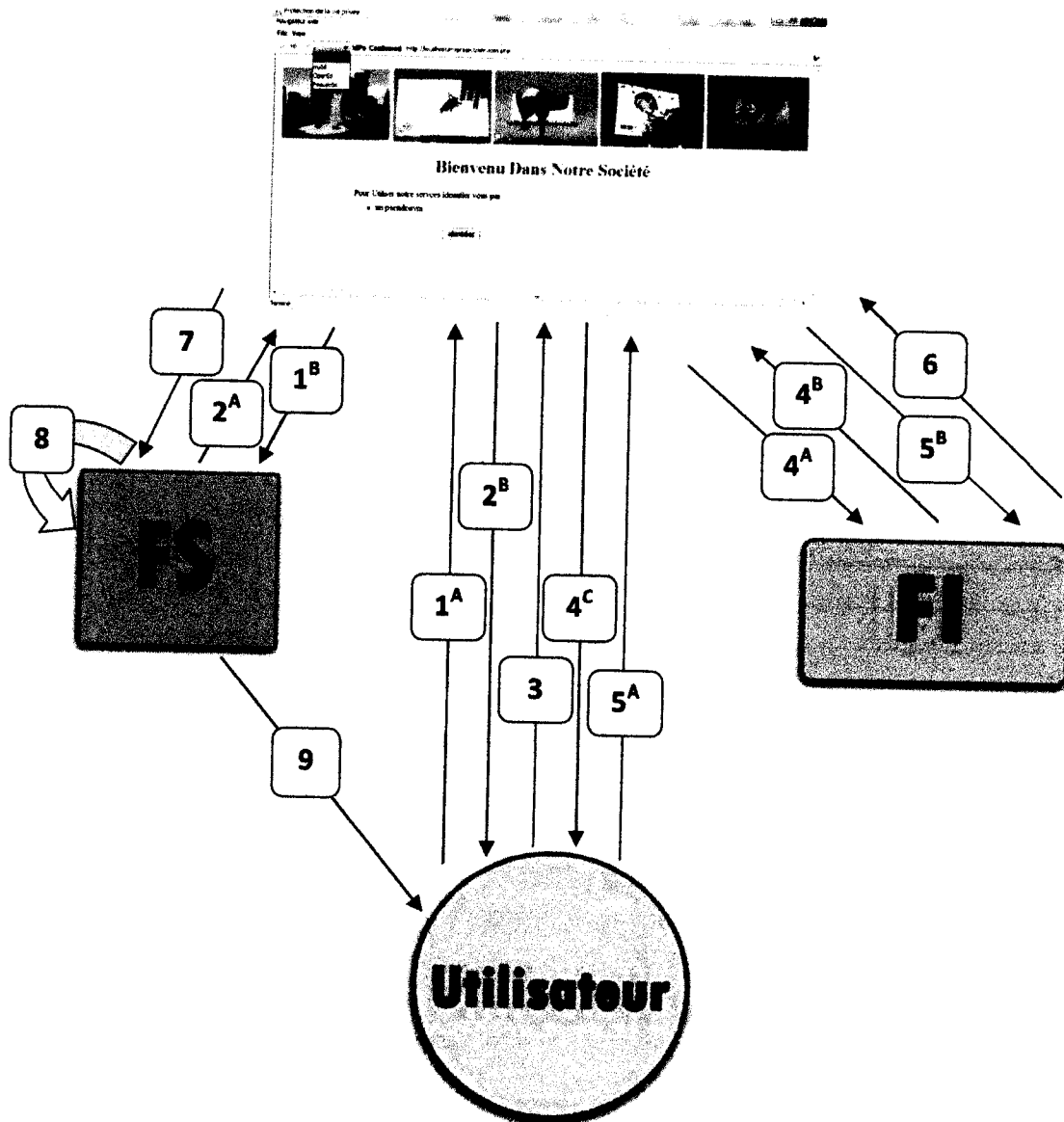


Figure III.4: Le déroulement des processus de façon détaillée du système proposé.

- 1- l'utilisateur fait le choix de son fournisseur de services à l'aide de son navigateur sans aucune identification (l'identification ce fait à l'étape avant dernière).
- 2- le fournisseur de services émit un cookie qui contient la liste de ces fournisseurs d'identité de confiance.
- 3- l'utilisateur fait le choix de son fournisseur d'identité à travers son navigateur.
- 4- le navigateur vérifie la correspondance du fournisseur d'identité (IdP) transmet par le SP avec celui choisi par l'utilisateur, dans l'affirmatif il redirige l'utilisateur vers la page d'authentification de fournisseur d'identité.
- 5- l'utilisateur s'identifie au niveau de son fournisseur d'identité en fournissant un pseudonyme et un mot de passe.

- 6- si l'utilisateur est authentifié, le fournisseur d'identité construit un jeton qui contient des informations sur l'utilisateur. ces informations sont signées et cryptées par une clé de session et stockées dans un cookie, qui sera analysé par le navigateur.
- 7- d'une façon automatique le navigateur décrypte les données émises par le FI et vérifie leur validité, si elles sont valides elles les envoient au fournisseur des services.
- 8- le fournisseur de services à son tour refait les étapes précédentes et extrait les données concernant l'utilisateur et vérifie leur validité, si elles sont valide il passe à l'étape final.
- 9- enfin le fournisseur de services offre ses services à l'utilisateur.

**III.2.2. Fonctionnement (Diagramme de séquence)**

Le diagramme de séquence suivant présente en détail les interactions effectuées par les entités de notre système :

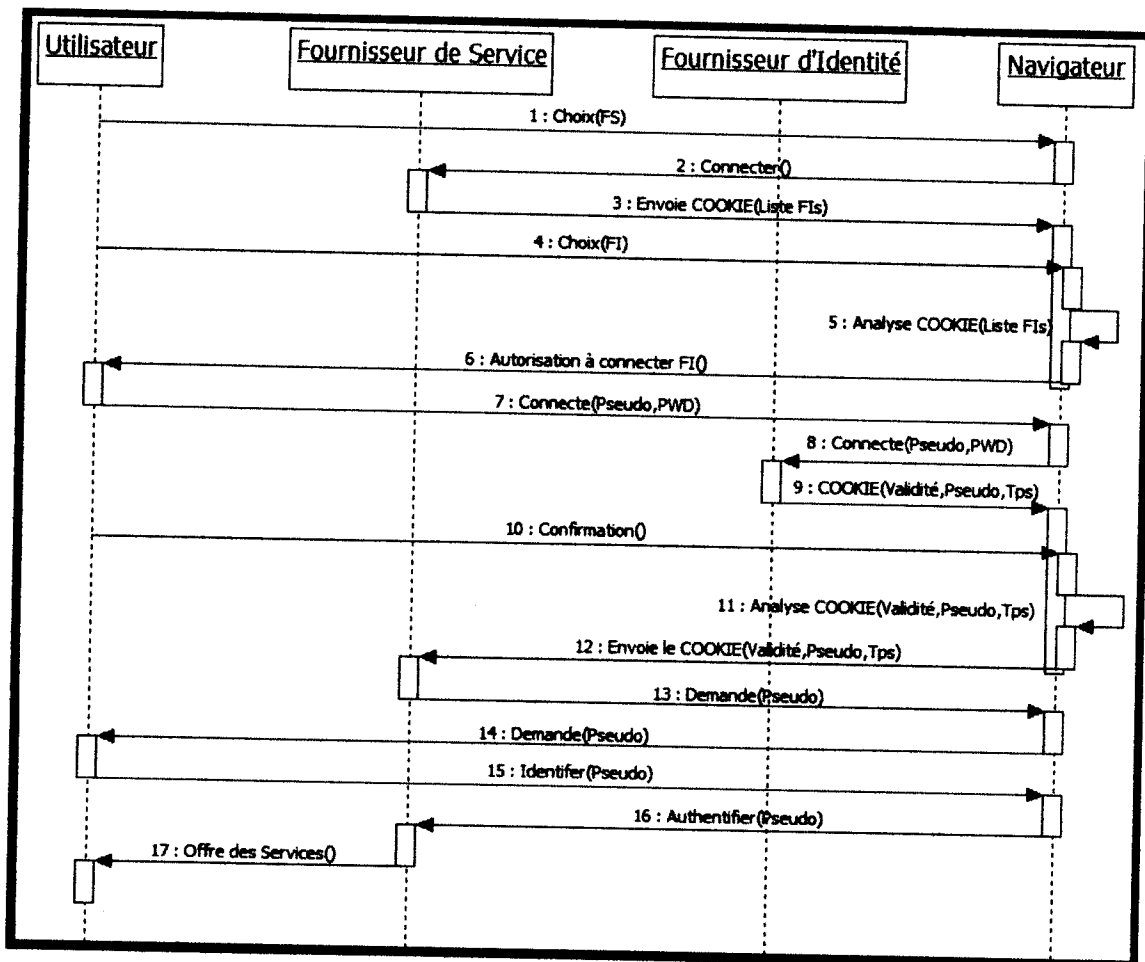


Figure III.5 : Diagramme de séquence des processus du système.

La suite de cette section introduit les détails de conception des entités impliquées dans le système proposé, à savoir le fournisseur d'identité, le fournisseur de services ainsi que l'extension du navigateur.

### III.2.2.1. Le fournisseur d'identité

Le rôle principal d'un fournisseur d'identité est de fournir une preuve d'authenticité de ses abonnés. Dans la solution proposée et après l'authentification d'un utilisateur, le FI génère un jeton qui contient trois informations principales :

- **la validité** : un champ qui indique si un utilisateur est correctement authentifié (OK), ou non (NO)
- **Le pseudonyme** : ce champ contient le pseudonyme choisi par l'utilisateur comme identifiant au niveau du fournisseur de service.
- **Le temps** : ce champ contient le temps de création du jeton. au moment de signature numérique des données, sera en GMT et convertit en secondes.

Après sa création, le jeton sera signé et crypté avant d'être transmis à l'utilisateur, le processus de signature et de cryptage des données se déroule comme suit :

#### 1. La signature électronique :

Dans notre système la signature électronique fait appel à deux familles d'algorithmes afin de pouvoir garantir l'authenticité et l'intégrité d'un document.

Les algorithmes asymétriques usuellement utilisés sont RSA et DSA<sup>1</sup>, les fonctions de hachages les plus courantes dont MD5<sup>2</sup> et SHA<sup>3</sup>.

Figure III.6 représente les différentes étapes du processus de signature et de vérification des données :

<sup>1</sup> DIGITAL SIGNATURE STANDARD (DSS), Federal Information Processing Standards, Publication 186, 1994 May 19, <http://www.itl.nist.gov/fipspubs/fip186.htm>

<sup>2</sup> R.L. Rivest, RFC 1321: *The MD5 Message-Digest Algorithm*, Internet Activities Board, 1992, <http://www.ietf.org/rfc/rfc1321.txt?number=1321>

<sup>3</sup> SECURE HASH STANDARD, Federal Information Processing Standards, Publication 180-1, 1995 April 17, <http://www.itl.nist.gov/fipspubs/fip180-1.htm>

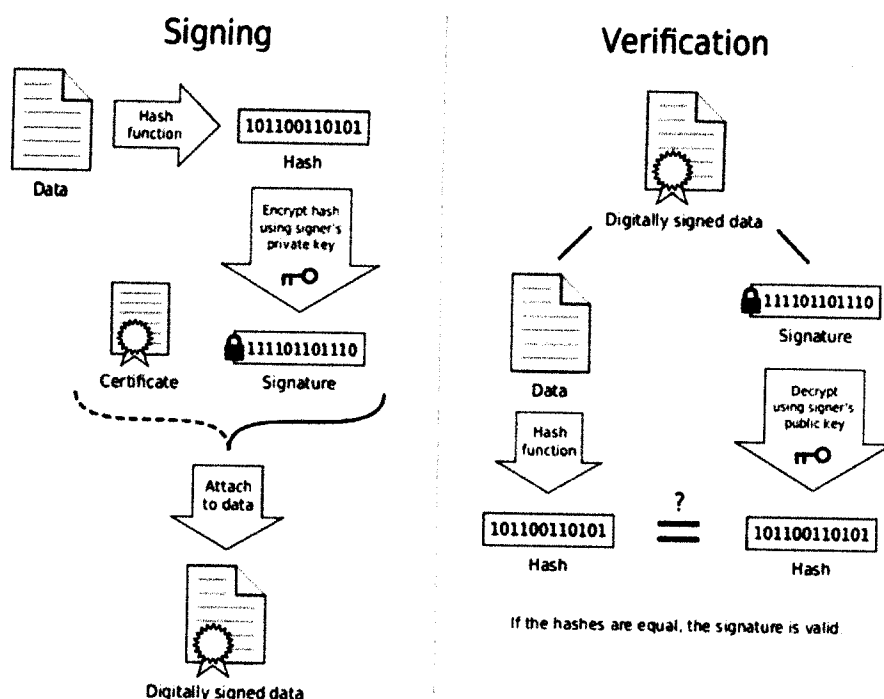


Figure III.6 : Principe de l'algorithme de signature électronique.

Pour l'implémentation des signatures nous avons utilisé la bibliothèque PHP « OpenSSL », la suite de ce paragraphe décrit quelques fonctions utilisées dans l'implémentation de notre système.

- **la fonction de signature:** « openssl\_sign » dont la signature et comme suit :

```
bool openssl_sign ( string $data , string &$signature , mixed $priv_key_id [, int $signature_alg = OPENSSSL_ALGO_MD5 ] )
```

**openssl\_sign()** calcule la signature des données *data* en utilisant l'algorithme MD5 (hashing) suivi du chiffage avec la clé privée *priv\_key\_id*. Cette fonction retourne **TRUE** en cas de succès, ou **FALSE** si une erreur survient.

- **la fonction de vérification des signatures électronique :** nous avons utilisé la fonction

« openssl\_verify » , cette fonction est utilisée au niveau de fournisseur des services , les paramètres de cette fonction sont comme suit

```
bool openssl_verify('tampered'.$data, $binary_signature, $public_key, OPENSSSL_ALGO_MD5);
```

Cette fonction retourne **TRUE** si les deux signatures (celle de l'émetteur et celle de l'information de récepteur) sont identiques, et **FALSE** en cas d'un changement dans la signature.

La variable *\$data* signifie l'information qu'on veut signer,

\$binary\_signateur : est la variable qui contient la signature numérique des informations,  
 \$private\_key : c'est la variable qui contient la clé privée du fournisseur d'identité,  
 et OPENSSL\_ALGO\_MD5 : c'est l'algorithme d'hachage.

**2. Le cryptage de jeton signé :**

La signature des données n'est pas suffisante, en cas où les données sont volées par un espion, même s'il ne peut pas les utiliser pour accéder aux services web, parce qu'elles sont signées, il peut lire facilement le contenu du jeton comme le pseudonyme, la validation des données et le temps de création du jeton, cela peut introduire un grand risque sur la vie personnelle d'autrui.

Pour couvrir cette défaillance de sécurité, nous avons ajouté une phase de cryptage des données après la procédure de signature, en utilisant une clé de session. Le cryptage se fait au niveau du fournisseur d'identité, le décryptage se fait par le fournisseur de service avant la vérification de la validité de la signature.

Avant d'expliquer la procédure de cryptage symétrique à base des clés de session, on mentionne qu'il y a des moyens plus sécurisés que cette méthode, par exemple l'utilisation de **Blind Signature**<sup>[21]</sup>, qui offre des moyens de sécurité de haut niveau, le service permet aux utilisateurs de générer un pseudonyme utilisé pour se connecter au fournisseur de services sans être chainable avec l'identité réelle de l'utilisateur. Il existe d'autres techniques telles que les protocoles **zero-knowledge proof**<sup>[21]</sup>, qui supportent la divulgation sélective des propriétés des utilisateurs.

Nous avons adopté une méthode de cryptographie symétrique à base des clés de session pour chiffrer les jetons. Par un algorithme de cryptage et de chiffrement symétrique moderne comme DES, IDES, AES... qui protège le contenu du jeton en cas si il est capturé par un espion.

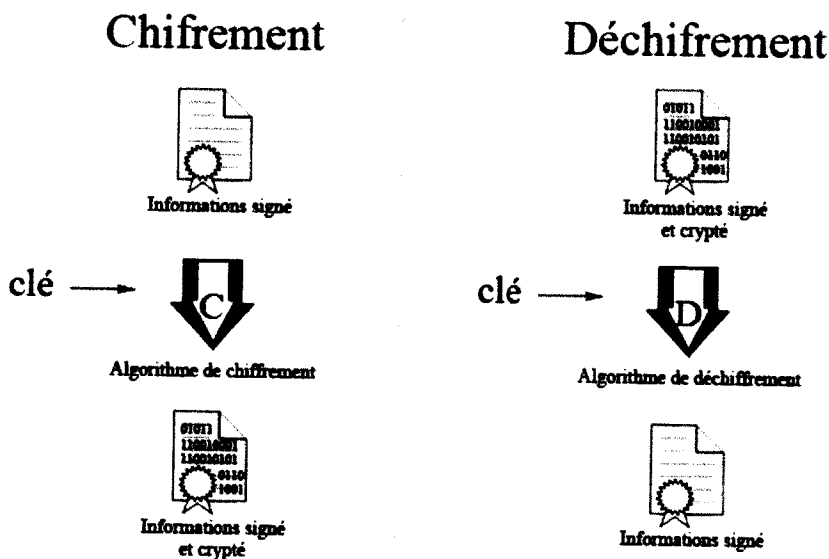


Figure III.7 : Opération de chiffrement de données signées.

La Figure III.7 présente les différentes étapes de chiffrement à clé symétrique, le chiffrement ce fait au niveau de fournisseur d'identité, et le déchiffrement au niveau de fournisseur de services. Cette étape a pour but d'augmenter et assurer l'intégrité des données d'une façon discrète et confidentielle.

### III.2.2.2. Le fournisseur des services

Le même principe que le fournisseur d'identité, sauf que le FS s'adresse à l'étape de déchiffrement des données, et de vérification de signature.

En premier lieu, le fournisseur de services, met un cookie temporaire sur le poste utilisateur. Le cookie contient la liste des fournisseurs d'identité dont le fournisseur de services fait confiance.

Après que l'utilisateur s'authentifie au FI, le FS analyse le cookie généré par le FI.

Les étapes d'analyse de cookie sont :

- le déchiffrement des données à l'aide de la clé de session,
- la vérification de signature électronique,
- si les deux étapes précédentes sont bien passées, il passe à l'étape de vérification de jeton. :
- La procédure de vérification du jeton prend en compte les éléments suivants :
  - La validité :** si le variable de validité est **NO**, le fournisseur de service va afficher une page d'erreur d'authentification, si cette variable est à **OK** il passe à la vérification du temps.
  - Le temps :** est une variable qui contient le temps de création de jeton par le fournisseur d'identité. la différence entre le temps de réception de jeton et le temps de sa création doit être inférieur à une certaine valeur (30 seconde par exemple) sinon l'authentification échoue.

### III.2.2.3. Le navigateur

Le navigateur joue le rôle d'un processus intermédiaire qui relie les fournisseurs des services et d'identités, par certaines règles et opérations :

La première fonctionnalité de navigateur concerne le choix de fournisseur d'identité, et la confirmation de son abonnement avec le fournisseur des services.

La deuxième fonctionnalité est le transfert des cookies d'authentification depuis le FI vers le FS, et la redirection des utilisateurs vers leurs site de FS.

## IV. Implémentation

### Présentation

Le but de notre application est d'éliminer les liens de confiance entre les FSs et FIs, et de le remplacer par les processus suivants qui sont intégrés au navigateur JAVA :

- 1- une liste de choix qui contient les fournisseurs d'identité existants.
- 2- bouton de confirmation après le choix de fournisseur d'identité, qui contient la fonction d'assertion si le fournisseur des services est abonné avec cet IdP ou non.
- 3- un autre bouton de confirmation qui vérifie l'intégrité des données de clients.

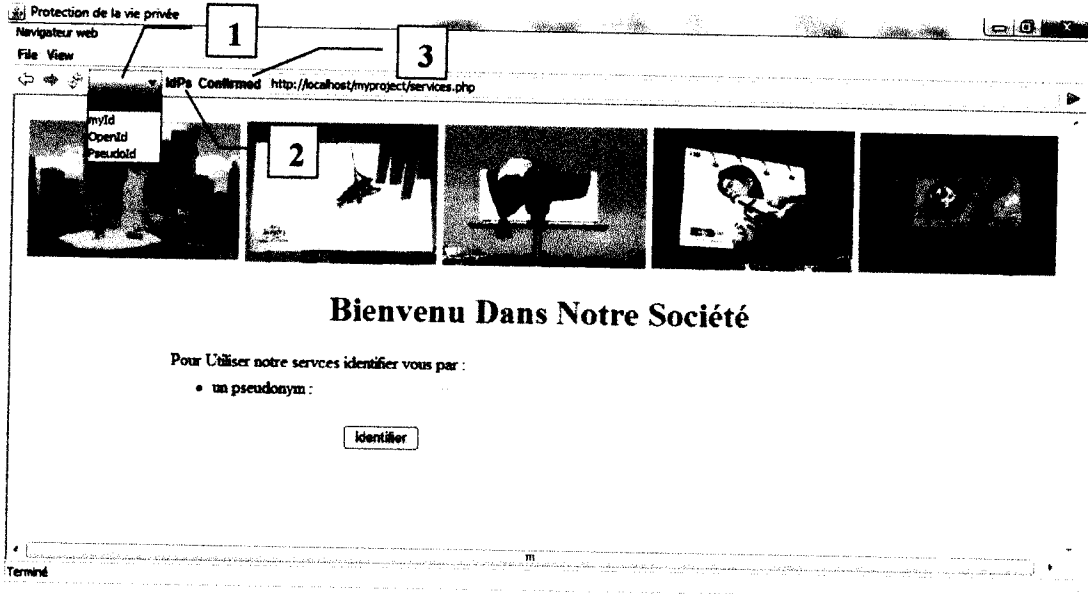


Figure III.8: Navigateur de java avec intégrité des services web.

Dans cette partie nous allons présenter les différentes étapes d'utilisation des éléments du système proposé :

Premièrement : l'utilisateur se connecte à un FS (Figure III.9).



Figure III.9 : La page d'accueil et d'identification de fournisseur des services.

Fait son choix depuis la liste 1 (liste des Idps) et confirme son choix par le bouton 2 (Idps).



Le navigateur ouvre la page d'identification du fournisseur d'identité choisi dans l'étape précédente (Figure III.10) :

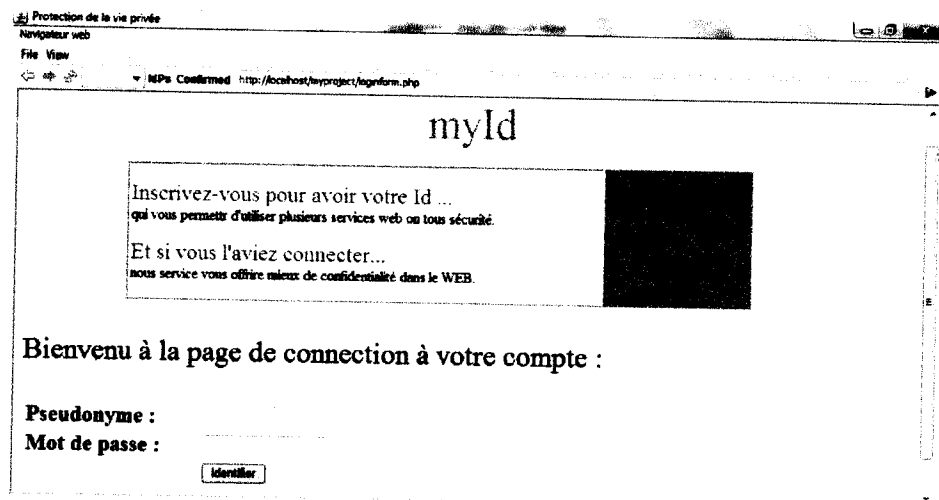


Figure III.10 : La page d'identification de fournisseur d'identité.

Si l'utilisateur n'est pas encore inscrit il clique sur le bouton d'inscription de la page login du FI, la page suivante s'affiche :

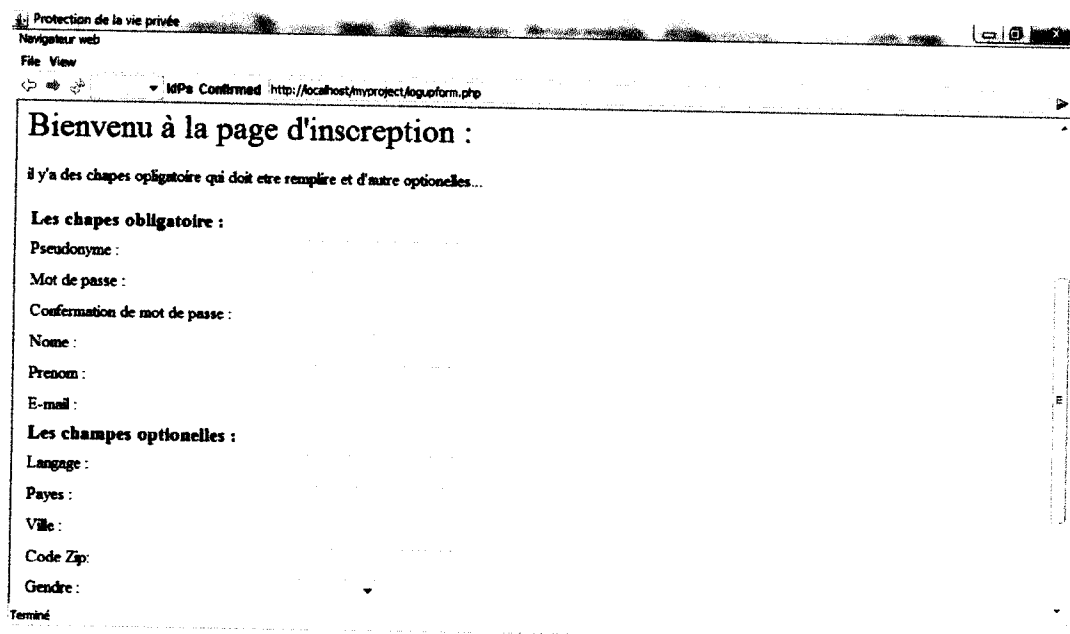


Figure III.11 : La page d'inscription de fournisseur d'identité.

Un utilisateur déjà inscrit doit s'identifier par son pseudo et son mot de passe (Figure III.10). Il est bien de noter que les mots de passe sont enregistrés de façon crypté dans la base de FI ce qui signifie que même ce dernier ne peut pas connaitre le mot de passe de l'utilisateur.

Si l'authentification aboutit à un succès, la page suivante sera affichée :

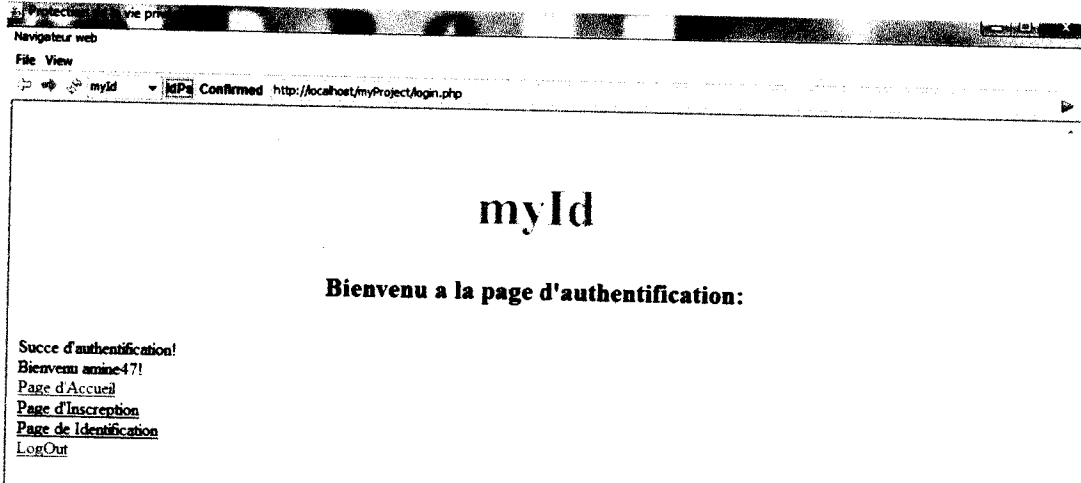


Figure III.12 : La page d'authentification du fournisseur d'identité.

L'utilisateur appuie sur le bouton 3 du navigateur (confirmed) qui le redirige vers son FS.

L'utilisateur s'identifie sur la page de FS par son pseudonyme Le FS affiche la page de succès et autorise l'utilisateur à utiliser les services fournis.

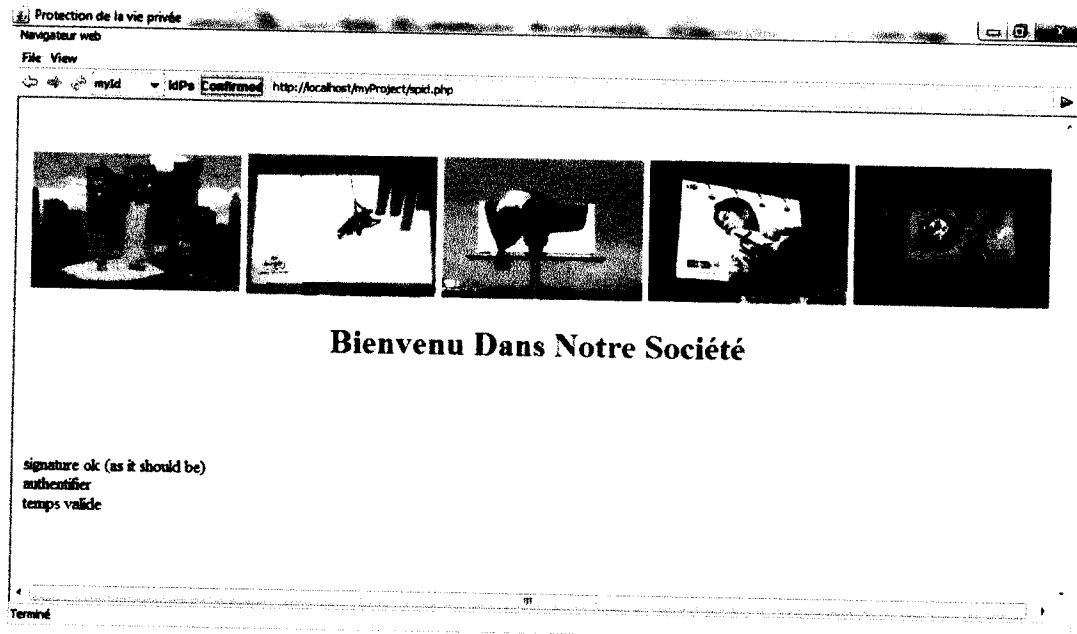


Figure 13 : La page d'authentification de fournisseur des services.

**V. Les outils de développement****V.1. PHPmySQL (WampServer Version 2.2)**

Le couple PHP / MySql est de plus en plus utilisé pour proposer du contenu dynamique sur le web. Cela signifie que les pages ne sont pas stockées telles quelles sur le serveur, mais générées en direct par ce dernier, le résultat étant envoyé directement au navigateur. La page visitée n'a pas d'existence en tant que telle, elle est créée au moment où on la demande.

Dans notre travail nous avons utilisé : Apache: 2.2.21 comme serveur Web, . Le PHP version 5.3.10 pour la construction des pages web dynamique, et la version 5.5.20 de MySQL pour la construction des bases des données web.

**V.2. L'IDE NetBeans 6.8**

Le navigateur java utilisé dans notre système est implémenté sous l'IDE NetBeans 6.8, le choix de cet environnement de développement est due à sa facilité d'utilisation et la richesse de la documentation relative.

**VI. Conclusion**

Dans ce chapitre nous avons présenté la partie pratique de notre travail. Nous avons présenté un système de gestion d'identité qui augmente la sécurité de transfert des informations et protège la vie privée des utilisateurs sur l'internet (web services), cette approche est basée sur l'élimination des liens d'échange d'informations entre les fournisseurs des services et les fournisseurs d'identité.

L'un des avantages de notre système est l'incapacité des fournisseurs d'identité de connaitre sur quel fournisseur de services l'utilisateur connecte et utilise son identité.



CONCLUSION GÉNÉRALE

# *Conclusion générale*

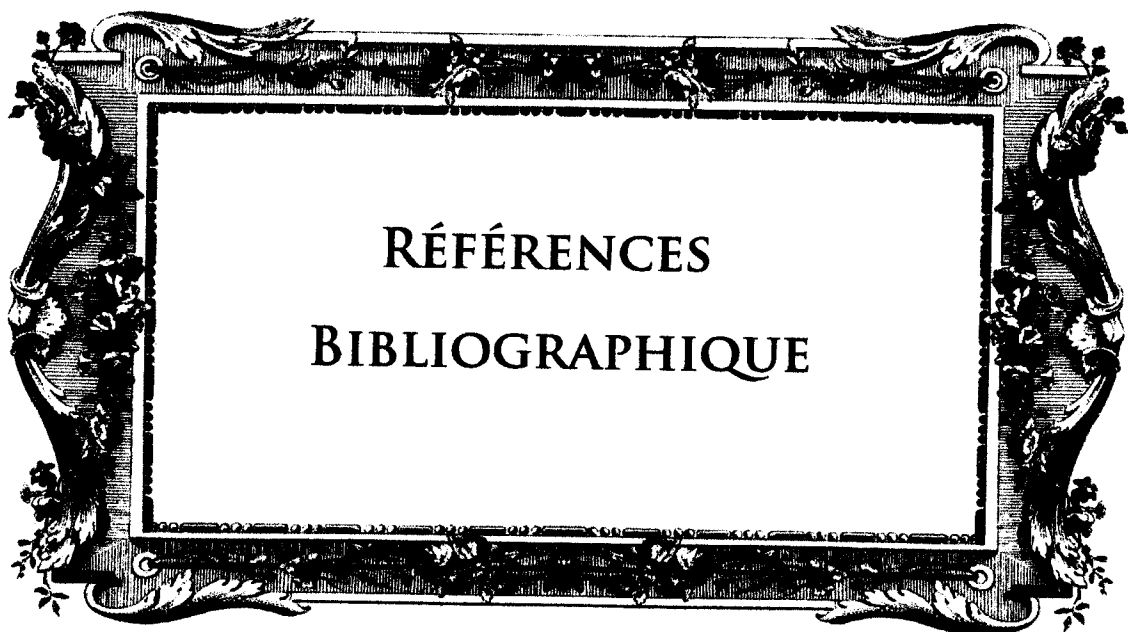
L'étude présentée dans ce mémoire a montré comment l'architecture d'identité fédérée peut simplifier la gestion des identités numérique surtout dans les environnements de collaboration telle que les services d'e-Gouvernement et e-Santé. De plus, le modèle de protection de vie privée proposé au 3<sup>ème</sup> chapitre introduit des améliorations sur certaines solutions existantes comme l'OpenID, et assure la sécurité des transactions entre les FSs et les FIs.

Après que nous avons présenté le prototype de notre système, les méthodes de son fonctionnement, et la manière dont il élimine le problème de l'OpenID, nous avons démontré comment ce prototype offre plus de sécurité et améliore la protection de la vie privée des personnes dans les web services.

Malgré les avantages de notre système, il reste toujours des améliorations à faire, surtout au niveau des algorithmes de cryptage et d'augmentation de sécurité. Nous envisageons d'introduire d'autres algorithmes plus sécurisés et plus adaptés pour ce genre de problématique comme le Blind signature et le zero-knowledge proof ...

D'autres perspectives de notre travail sont :

- Intégrer les fonctionnalités notre système comme extensions de navigateurs, comme Internet Explorer, Firefox, Safari et Opera, à l'aide des langages de développement et de construction des plug-ins comme XUL.
- Améliorer notre système par l'ajout de nouvelles fonctionnalités offrant la possibilité de choix sélectif des informations à révélées, et la possibilité d'utilisation de plusieurs pseudonymes.
- Sécuriser les sites des FIs et FSs, à l'aide des protocoles plus sécurisés comme : HTTPS et SSL ...



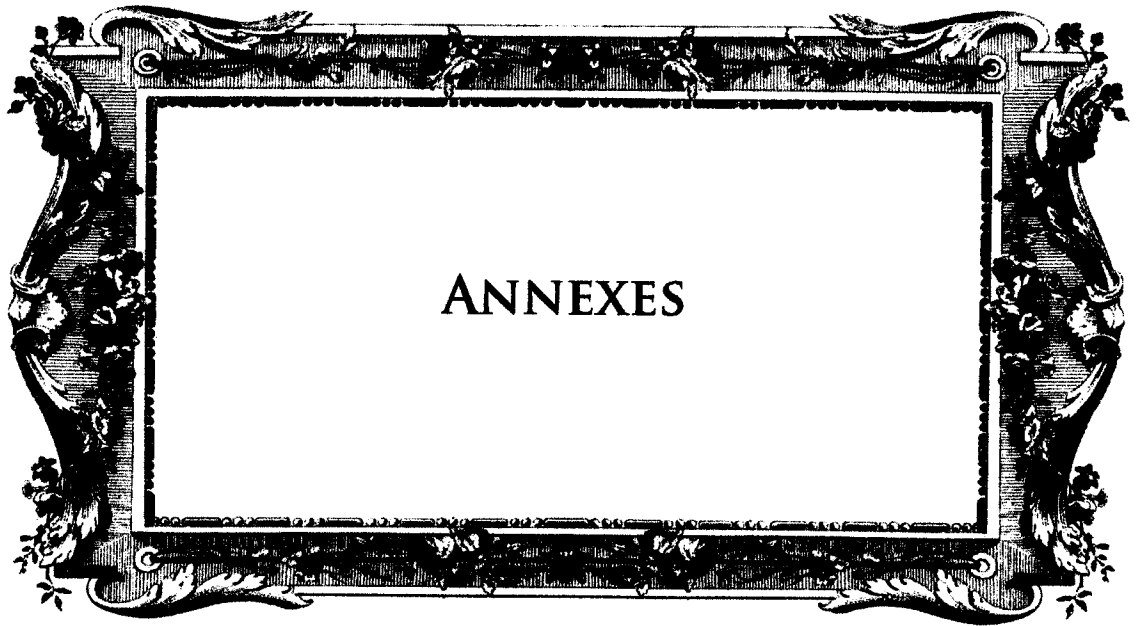
RÉFÉRENCES  
BIBLIOGRAPHIQUE

## *Références bibliographiques*

- [1] Nathalie MALLET-POUJOL, « *Protection de la vie privée et des données personnelles* », Chargée de Recherche au CNRS , Université Montpellier I- UMR 5815, Février 2004.
- [2] Loi n° 2004-801 , Article 1, « *La définition des données personnelles* », relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel, 06 août 2004.
- [3] Centre du cyberfutur de l'Alberta , « *La vie privée sur Internet* », guide, Équipe d'affaires électroniques de l'Ouest , 2005.
- [4] OCDE Document exploratoire sur le vol d'identité en ligne  
Document de référence DSTI/CP(2007)3/FINAL  
Réunion Ministérielle de l'OCDE le futur de l'économie Internet.
- [5] Andreas PFITZMANN, « *A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management* », Version v0.34, 10.Aug.2010.
- [6] Vincent Regnault, « *Protection de la vie privée des patients par la traçabilité des accès aux applications médicales* », Fécamp : Centre Hospitalier Intercommunal, 2012.
- [7] Pierre Truche « *Ministère de la fonction publique et de la réforme de l'État* », République française , Le Ministre N/REF/CAB/2001 -80/GB, 13 décembre 2001.
- [8] Daniel Le Métayer et Guillaume Piolle, « *Droits et obligations à l'ère numérique : protection de la vie privée* », 30 septembre 2010.
- [9] Pierre-Jean Benghozi .al, « *L'Internet des objets* », Ecole polytechnique TELECOM parisTech, Octobre 2008.
- [10] Yves Deswarte, « *Des Technologies pour protéger la vie privée sur Internet* », Université de Toulouse, avenue du Colonel Roche 31077 Toulouse cedex 4 France.

- [11] Carlos AGUILAR MELCHOR, « *Les communications anonymes à faible latence* », thèse de doctorat, Laboratoire d'Analyse et d'Architecture des Systèmes du CNRS, 27 Oct 2006.
- [12] Sébastien Gambs, « *Réseaux de communication anonyme* », Cours, 22 novembre 2011.
- [13] « Spécification de la plateforme pour les préférences de confidentialité 1.0 (P3P 1.0) », consulté le 05/03/2011, disponible sur :  
<http://www.yoyodesign.org/doc/w3c/p3p1/index.html>
- [14] Institute for prospective Technological Studies, « *Sécurité et Respect de la vie Privée du citoyen A l'ère du Numérique après Le 11 septembre: Vision prospective* », Document de Synthèse, Juillet 2003.
- [15] Club de la sécurité de l'information France, « *GESTION DES IDENTITES* », dossier technique, Juillet 2007.
- [16] LRDP Kantor Ltd et Centre for Public Reform, « *DIFFÉRENTES APPROCHES DES NOUVEAUX DÉFIS EN MATIÈRE DE PROTECTION DE LA VIE PRIVÉE, EN PARTICULIER À LA LUMIÈRE DES ÉVOLUTIONS TECHNOLOGIQUES* », étude comparative, Janvier 2010.
- [17] M. Uciel FRAGOSO-RODRIGUEZ, « *Modèle de Respect de la Vie Privée dans une Architecture d'Identité Fédérée* », thèse de doctorat, version 1-1 Décembre 2010.
- [18] Housseem Jarraya et Maryline Laurent-Maknavicius, « *Liberty Alliance et le respect de la vie privée* », étude d'un stage, 2007.
- [19] « Wikipedia, the free encyclopedia », disponible sur :  
<http://fr.wikipedia.org/wiki/OpenID>.
- [20] Ministère de l'éducation nationale, de la jeunesse , « *Schéma directeur des espaces numériques de travail Recommandations pour l'Authentification-Autorisation-SSO (AAS)* », République Française, 13/07/2011.
- [21] Arkajit Dey<sup>1</sup> and Stephen Weis Massachusetts Institute of Technology,  
*PseudoID: Enhancing Privacy for Federated Login*  
Cambridge, MA, USA 02139 Google Inc., Mountain View, CA, USA 94043
- [22] Sébastien Gambs « *Respect de la vie privée dans la société de l'information* », Université de rennes 1 – INRIA/IRISA, 29 mars 2011





ANNEXES

# Annexe A

---

**La présente annexe fournit quelques conseils généraux permettant aux utilisateurs de mieux protéger leur vie privée en naviguant sur le web**

## 1. La navigation privée

Tous les navigateurs récents offrent une mode de navigation privée. Lorsque cette fonction est activée, aucun historique (formulaire, téléchargement, recherche, fichiers témoins, etc.) de la session de sera conservé par le navigateur. C'est une fonction pratique en allant sur le web à partir d'un ordinateur public ou partagé.

Un raccourci-clavier permet de basculer en mode de navigation privée : [Ctrl]+[Maj]+[p] pour IE, FF et [Ctrl]+[Maj]+[n] pour Chrome et Opera.

## 2. Le web of trust



Un complément finnois qui s'installe sur les navigateurs principaux (Chrome, IE, FF, Opera et Safari) et qui affiche le niveau de confiance lié au site web vous visitez. La cote est affichée sous forme de lumière verte, jaune, rouge (il existe aussi une version pour daltonien). WOT (Web Of Trust) protège ainsi les internautes des marchands douteux, des sites web présentant des contenus inappropriés et des sites web qui collectent des informations personnelles sans permission préalable. Lien : <http://www.mywot.com/>

## 3. The Onion Router

Vous désirez être complètement anonyme sur le web, c'est Tor (The Onion Router) qu'il vous faut. Utilisé par les activistes, journalistes et militaires depuis belle lurette (1996), Tor empêche les intrus de connaître votre emplacement et vos habitudes de navigation. Il s'agit d'un logiciel libre disponible en français pour Windows, Mac, Linux et Android! Lien : <https://www.torproject.org/>



#### **4. Collusion**

Sympathique petite extension pour Firefox qui affiche visuellement les différents fichiers témoins actifs qui enregistrent votre navigation. La prochaine version devrait nous permettre de désactiver sélectivement ces fichiers témoins. Pour savoir qui suit vos mouvements en ligne! Lien : <http://www.mozilla.org/en-US/collusion/>

#### **5. Priv.ly**


Une nouvelle extension pour FF et Chrome qui propose de garder votre contenu à vous! En acceptant les licences d'utilisation des services comme Facebook, Google, Twitter et compagnie, vous leur donnez le droit d'utiliser VOS données à LEUR guise. Avec Privly, le contrôle des données demeure vôtre. C'est une extension qui en est encore à l'étape de preuve de concept, en phase bêta sur invitation et en collecte de fonds. Mais le principe de base est fort probant. Lien : <https://priv.ly/>

# Annexe B

---

## OpenID : comment ça marche ?

### OpenID, c'est quoi exactement ?

 Avec OpenID vous n'avez plus qu'un seul identifiant qui vous permet de vous connecter en quelques secondes à vos sites favoris et à tous les nouveaux sans avoir à créer à chaque fois un nouveau compte avec un nouveau mot de passe.

AOL, Facebook, France Telecom, Google, Microsoft, MySpace, Yahoo! utilisent la technologie OpenID et le nombre de sites mettant en place un système de connexion via cette technologie ne cesse de grimper.

### Comment obtenir une OpenID ?

Surprise! Vous en avez peut être déjà une. Si vous utilisez l'un des services suivants, vous avez déjà votre propre OpenID. (quand vous voyez du texte en italique, vous devrez le remplacer par votre nom d'utilisateur, pseudo ou nom de membre sur le site pour obtenir votre login OpenID.)

AOL : [openid.aol.com/pseudo](http://openid.aol.com/pseudo)

Google : Cherchez le bouton "Sign in with a Google Account"

MySpace : Cherchez le bouton "Login with MySpaceID" ou entrez [myspace.com/username](http://myspace.com/username)

Yahoo! : Cherchez le bouton "Sign in with Yahoo! ID"

Blogger : [votreblog.blogspot.com](http://votreblog.blogspot.com)

Flickr : Cherchez le bouton "Sign in with Yahoo! ID" ou entrez [www.flickr.com/photos/username](http://www.flickr.com/photos/username)

LiveDoor : [profile.livedoor.com/username](http://profile.livedoor.com/username)

LiveJournal : [username.livejournal.com](http://username.livejournal.com)

Orange (France Telecom) : [orange.fr](http://orange.fr)

SmugMug : [username.smugmug.com](http://username.smugmug.com)

Technorati : [technorati.com/people/technorati/username](http://technorati.com/people/technorati/username)

Vox : member.vox.com

WordPress.com : username.wordpress.com

Et si chose hautement improbable, vous ne possédiez aucun compte dans les services précédemment cités, vous pouvez toujours vous créer une OpenID : <http://www.openidfrance.fr/>.

A noter que votre compte Hotmail pourra bientôt lui aussi faire office d'OpenID.

# Annexe C

Cette annexe présente quelques détails d'implémentation sur les fonctionnalités utilisées dans le système proposé à savoir : l'utilisation des bases de données, les fonctions cryptographiques, la manipulation des cookies ainsi que la bibliothèque « **JWebBrowser** » utilisé pour implémenter le navigateur java.

## 1. Mode d'utilisation MySQL :

### Création d'une table :

```
CREATE TABLE FilmSimple  
(titre VARCHAR (30),annee INTEGER,nom_realisateur VARCHAR  
(30),prenom_realisateur VARCHAR (30),annee_naissance INTEGER);
```

### Insertion des données :

```
INSERT INTO FilmSimple (titre,annee,prenom_realisateur,nom_r  
ealisateur) VALUES ( ' Pulp Fiction ' , 1995 , ' Quentin ' , ' Tarantino ' );
```

### Interrogation et modification :

```
SELECT titre, annee FROM FilmSimple  
WHERE annee > 1980 DELETE FROM FilmSimple WHERE annee <= 1960  
UPDATE FilmSimple SET nom_realisateur= 'Wu',prenom_realisat  
eur= 'Yusen' WHERE nom_realisateur= 'Woo'
```

### Quelques commandes utiles :

**SELECT DATABASE();** C'est une pseudo-requête SQL (sans FROM) qui affiche le nom de la base courante.

**SELECT USER();** Idem, cette pseudo-requête affiche le nom de l'utilisateur courant.

**SHOW DATABASES;** Affiche la liste des bases de données.

**SHOW TABLES;** Affiche la liste des tables de la base courante.

**SHOW COLUMNS FROM** *NomTable* ; Affiche la description de la table *NomTable*.

## Les fonctions principales PHP/MySQL :

Fonction	Description
<code>mysql_connect()</code>	Pour établir une connexion avec MySQL, pour un compte utilisateur et un serveur donnés. Renvoie une valeur utilisée ensuite pour dialoguer avec le serveur.
<code>mysql_pconnect()</code>	Idem, mais avec une connexion <i>persistante</i> (voir annexe C). Cette deuxième version est plus performante quand l'interpréteur PHP est inclus dans Apache.
<code>mysql_select_db()</code>	Permet de se placer dans le contexte d'une base de données. C'est l'équivalent de la commande <code>USE base</code> sous <i>mysql</i> .
<code>mysql_query()</code>	Pour exécuter une requête SQL ou n'importe quelle commande MySQL. Renvoie une variable représentant le résultat de la requête.
<code>mysql_fetch_object()</code>	Récupère une des lignes du résultat et positionne le curseur sur la ligne suivante. La ligne est représentée sous forme d'un <i>objet</i> (un groupe de valeurs).
<code>mysql_fetch_row()</code>	Récupère une des lignes du résultat, et positionne le curseur sur la ligne suivante. La ligne est représentée sous forme d'un <i>tableau</i> (une liste de valeurs).
<code>mysql_error()</code>	Renvoie le message de la dernière erreur rencontrée.

### 2. Mode d'utilisation PHP :

#### Les sessions :

`session_start()`; // Initialise les informations de session. Si aucune session n'existe, un identifiant est engendré et transmis dans un *cookie*. Si la session (connue par son identifiant) existe déjà, alors la fonction instancie toutes les variables qui lui sont liées. Cette fonction doit être appelée au début de tout script utilisant les sessions (il faut que l'instruction Set-Cookie puisse être placée dans l'en-tête HTTP).

```
$_SESSION['user']=$_POST['pseudonyme']; //nom de session user.
```

```
Session_id(); // Renvoie l'identifiant de la session.
```

```
session_unset(); // supprimer tout les variable de la session.
```

```
session_destroy(); // Détruit toutes les informations associées à une session..
```

#### la fonction de cryptage :

```
function encrypt($data) {
```

```
    $key = $_POST['pseudonyme']; // Clé de session à 8 caractères au maximum
```

```
    $data = serialize($data);
```

```
    $td = mcrypt_module_open(MCRYPT_DES, "", MCRYPT_MODE_ECB, "");
```

```
    $iv = mcrypt_create_iv(mcrypt_enc_get_iv_size($td), MCRYPT_RAND);
```

```
    mcrypt_generic_init($td,$key,$iv);
```

```

    $data = base64_encode(mcrypt_generic($td, '!.$data));
    mcrypt_generic_deinit($td);
    return $data;
}

```

#### **Fonction de décryptage :**

```

function decrypt($data) {
    $key = $_POST['pseudo'];
    $td = mcrypt_module_open(MCRYPT_DES, "", MCRYPT_MODE_ECB, "");
    $iv = mcrypt_create_iv(mcrypt_enc_get_iv_size($td), MCRYPT_RAND);
    mcrypt_generic_init($td, $key, $iv);
    $data = mdecrypt_generic($td, base64_decode($data));
    mcrypt_generic_deinit($td);
    if (substr($data, 0, 1) != '!')
        return false;
    $data = substr($data, 1, strlen($data)-1);
    return unserialize($data);
}

```

#### **Les cookies :**

##### **Créer un cookie avec PHP**

```

int setcookie ( string name [, string value [, int expire [, string path [, string domain [, int
secure]]]]);

```

```

<?php setcookie('cookie_name', 'blablabla', (time() + 3600)); ?>

```

##### **Lecture d'un cookie**

```

<?php /* Les trois exemples suivants afficheront tous "blablabla" */
echo $cookie_name; // exemple 1 (si registrar_globals est à on dans php.ini)
echo $HTTP_COOKIE_VARS['cookie_name']; // exemple 2
echo $_COOKIE['cookie_name']; // exemple 3 (si on est sur PHP 4.1.0 ou plus)
?>

```

##### **Détruire un cookie**

```

<?php /* Les deux exemples suivants sont équivalents*/
setcookie('cookie_name'); // exemple 1
setcookie('cookie_name', "", 1); // exemple 2
?>

```



### Connection au base des données

```
<?php
    $conn = mysql_connect("localhost","root","");
    $db = mysql_select_db("myid",$conn);
?>
```

### Extraction des données de page d'inscriptions

```
$pseudonyme = $_POST['pseudonyme'];
    $pw1 = MD5($_POST['pw1']); //mot de passe crypté (inconnu même par le FI)
    $pw2 = MD5($_POST['pw2']);
    $nome = $_POST['nome'];
    $prenome = $_POST['prenome'];
    $email = $_POST['email'];
    $langage = $_POST['langage'];
    $payes = $_POST['payes'];
    $ville = $_POST['ville'];
    $codezip = $_POST['codezip'];
    $genre = $_POST['genre'];
    $age = $_POST['age'];
```

### Insertion des données dans la base MySQL

```
$sql = "INSERT into users values('".$pseudonyme."','".$pw1."','".$nome."','".$prenome."','".$email."','".$langage."','".$payes."','".$ville."','".$codezip."','".$genre."','".$age."')";
    $query = mysql_query($sql);
```

### Vérification des données à la page d'identifications

```
$pseudonyme=$_POST['pseudonyme'];
    $pw=MD5($_POST['pw']);
    $sql="select count(*) from users where(
    pseudonyme='".$pseudonyme.'" and pw='".$pw."')";
    $query=mysql_query($sql);
    $result=mysql_fetch_array($query);
```

### 3. Mode d'utilisations JWebBrowser au NetBeans

On a utilisé les bibliothèques suivantes :

DJNativeSwing-SWT.jar

swt-debug.jar

Le codes qui manipule le browser ce sont intégré à les bibliothèques précédents ce que nous intéressé sont les actions additionnelles qui offrir la sécurité entre les FSs et les FIs.

#### Création d'une liste des fournisseurs d'identité :

```
final JComboBox combox=new JComboBox();
combox.addItem("");
combox.addItem("myId");
combox.addItem("OpenId");
combox.addItem("PseudoId");
```

#### Ajouté la liste des fournisseurs d'identité à la barre de JWebBrowser

```
buttonBar.add(combox);
```

#### Création d'un bouton de choix de fournisseur d'identité

```
final JButton button = new JButton("IdPs");
button.setForeground(Color.BLACK);
button.setFont(new Font("sansserif",Font.BOLD,12));
button.addActionListener(new ActionListener() {
    public void actionPerformed(ActionEvent e) {
// Les codes qui concerne ce bouton
    }
}
```

La même chose pour le bouton de confirmation des données authentifié par le FI.

## Résumé

L'augmentation des transactions en ligne dans les services web entre les fournisseurs des services FSs et d'identités FIs, conduit ces deux derniers à créer des systèmes de gestion d'identité SGI pour faciliter l'utilisation des services web et l'accès des utilisateurs à leurs informations personnelles.

Ce système est basé sur le concept de Single Sign On (SSO) comme l'OpenID, Microsoft CardSpace, ... Le problème de ce système (OpenID) est les liens de transactions entre les FSs et les FIs, qui permettent au FIs de traquer les services utilisés par l'utilisateur, à cause de cette permission, la protection de la vie privée est diminuée.

Le but de notre projet est d'éliminer ces liens de transaction et de les remplacer par un processus automatique géré par le navigateur à l'aide des signatures numériques et des Algorithmes de cryptographie.

**Les mots clé :** Web-Service, Single Sign On, OpenID, protection de vie privée, identité numérique, système de gestion d'identité (centralisé, fédéré).

## Abstract

The increase of online transactions in web services between services providers SPs and identities providers IDPs, conduct to creates identity management systems IMS which make more easier the use of web services and users access to their personal information.

These systems are based on the concept of Single Sign On (SSO) eg: OpenID IDPs, Microsoft CardSpace, ... The problem of OpenID system is the trust links between the SPs and IDPs which enable the IDPs to track user services accesses, thus, a privacy protection lacks.

The aim of our project is eliminate these bonds of trust and replaced them by an automatic process done by the browser using signature and cryptographic algorithms.

**Key words:** Web Service, Single Sign Sn, OpenID, protection of privacy life, digital identity, identity management system (centralized, federated).

## ملخص

لقد أدت الزيادة في المعاملات الفورية للصفقات و الخدمات في الأنترنت بين مزودي الخدمات FSs و مزودي الهوية FIs إلى إنشاء نظام تسيير الهوية SGI، لتسهيل استخدام خدمات شبكة الإنترنت، و وصول المستخدمين إلى المعلومات الشخصية الخاصة بهم. وهذا النظام الذي يعتمد على مفهوم التسجيل بالهوية الموحدة، مثل: مزودي هوية OpenID و Microsoft CardSpace ... مشكلة نظام OpenID تكمن في إنشاء روابط ثقة بين مزودي الخدمات و مزودي الهوية، و الذي يتيح لمزودي الهوية تتبع خدمات المستخدمين مما يؤدي إلى نقصان الحماية لحياة الأشخاص في الأنترنت.

و الهدف من مشروعنا، هو إزالة روابط الثقة هذه وتعويضها بعمليات تلقائية تدمج في المتصفح و ذلك باستعمال التوقيع الإلكتروني و خوارزميات التشفير.

الكلمات المفتاح: خدمات الويب، حماية الخصوصية والهوية الرقمية، نظام تسيير الهوية.