République Algérienne Démocratique et Populaire
Université Abou Bakr Belkaid– Tlemcen
Faculté de technologie
Département d'Informatique

Mémoire de fin d'études
pour l'obtention du diplôme
Master en Informatique

Option : Système d'information et de connaissance (SIC)

Thème

Protection des données personnelles coté utilisateur avec les ontologies

Réalisé par :

- **M^lle^ BOUALI Hanane**
- **M^me^ SIRIARI Fadila**

Présenté devant le jury:

- ***M^r^ BENAMAR Abdelkim*** ***(président)***
- ***M^me^ DARI BEKARA Kheira*** ***(Encadreur)***
- ***M^r^ Mana Mohammed*** ***(Examinateur)***

Année universitaire :2016-2017

REMERCIEMENT

En tout premier lieu, nous remercions le bon Dieu, tout puissant, de nous avoir donné la force pour survivre, ainsi que l'audace pour dépasser toutes les difficultés.

Nous tenons à exprimer toute nos reconnaissance à notre encadreur de mémoire Madame BEKARA DARI KHEIRA. Nous la remercions de nous avoir encadré, orienté, aidé et conseillé.

Nos profonds remerciements pour les membres de jury qui ont accepté d'évaluer ce travail.

Nous remercions également M^{r} BENAMAR AEK chef département d'informatique.

Dédicace

Louange à Dieu, le seul et l'unique

Je dédie ce modeste travail

A mes parents qui sont investis corps et âme sur ma scolarité, qui sont sacrifies pour toujours m'apporter : soutien sans limite et sans faille, tolérance et encouragements qu'ils a bien voulu consentir pour moi .tous les mots restent faibles pour lui exprimer ma profonde reconnaissance et qu'ils veuillent bien accepter ces lignes en guise remerciement que dieu tout puissant les bénisse à jamais, lui accorde tout le bonheur qu'ils méritent et très langue vie pleine de santé

A mon frère et mes sœurs et ma belle sœur, mes nièces et mes neveux.

A toute ma famille.

A ma binôme SIRIALI FADILA,

A toute la promotion 2ème Années master SIC 2015-2016,

A tous mes collègues de travail,

A tous ceux qui ont participés de prés ou de loin dans la réalisation de ce travail,

HANANE

Dédicace

Je dédie ce mémoire,

A mes enfants mohamed et abdelilah

A mon marie djemmaa kamel

A ma copine bouali hanane

FADILA

Résumé :

Le développement de l'Internet ne peut masquer les réelles ambigüités de contrôle d'accès à l'information entre le client et le service provider.
Le présent travail examine d'abord le contexte légal de la vie privé et les divers moyens informatiques destinés à la protection des données personnelles.
Nous avons étudié les différents modèles de contrôle d'accès, montrant ainsi les avantages qui nous ont mener à choisir le modèle XP A CML (eXtensible Privency Acces Control Markcup Lunguge) conçu sur la base d'extensions apportées au modèle de contrôle d'accès XACML.
Notre travail s'inscrit dans la continuité d'extension du langage XPACML , en nous basons sur des outils du Web Sémantique, en particulier les ontologies.
Notre ontologie reflète les données de l'utilisateur à protéger, l'ensemble des fournisseurs de services qui sont susceptibles d'utiliser les données de l'utilisateur; et les politiques permettant à l'utilisateur de spécifier ses préférences en matière de protection de données.
Mot clés : ontologie, protection des données, web sémantique, XPACML.

ملخص:

تطوير شبكة الإنترنت لا يمكن أن يخفي في الواقع التحكم في الوصول إلى غموض المعلومات بين العميل ومقدم الخدمة.
في عملنا نحن ننظر أولا في السياق القانوني للخصوصية، و كذا وسائل الاعلام الالي المختلفة لحماية البيانات الشخصية.
لقد درسنا نماذج مختلفة لمراقبة الدخول واللغات المختلفة التي أوصلتنا إلى معرفة محسنات XPACML مما أدى بنا إلى اختياره لأنه مصمم على أساس التمديد الذى أدخل على نموذج التحكم في الوصول XACML.
يندرج عملنا ضمن الاستمرارية في تمديد لغة XPACML، معتمدين في ذلك على الويب الدلالي و بالأخص الانطولوجيا.
تعكس لنا الانطولوجيا بيانات المستخدم التي يجب حمايتها، وجميع مقدمي الخدمات الذين يمكنهم استخدام بيانات المستخدم.
والسياسات التي تسمح للمستخدم تحديد اولويات لحماية البيانات.
الكلمات المفتاحية: الانطولوجيا ،الويب الدلالي ،حماية البيانات،XPACML .

Abstract:

The development of the Internet can not really mask the ambiguities of access control to information between the client and the service provider.
In our work we first examine the legal context of privacy as well as the different means for the protection of personal data.
We have shown the different access control models as well as the different languages that led us to XPACML (eXtensible Privency Acces Control Markcup Lunguge) based on extension to the XACML access control model.
We develop a data protection tool based on the semantic web, more accurate ontology.
Our ontology reflects the user data model.
The set of providers are likely by the user, the service provider may use specific policy model policies to specify personal preferences.
Key word: ontology, data protection, semantic web, XPACML.

Table des matières:

Table des figures:

INTRODUCTION GENERALE:

L'apparition de l'informatique et de l'Internet a changé la nature de nos travaux et de nos échanges. Le traitement de nombreux types d'informations fut désormais automatisé. La duplication et le partage sont d'avantage facilités ; proliférant ainsi le nombre de services mises en ligne tel que le e-commerce, e-learning, etc...

Quoique l'informatisation du traitement des données ait généralement considérée comme un énorme progrès technique, elle s'est aussi accompagnée de risques et de menaces sur les données personnelles des utilisateurs.

En effet, avoir un accès à leurs informations personnelles avec des pratiques d'usages exprimées dans un jargon juridique textuel, incompréhensible, rebute la quasi-totalité des utilisateurs. Ainsi de plus en plus d'utilisateurs deviennent soucieux de leurs données personnelles et de leur vie privée, dont la nature des problématiques posées par les progrès informatiques ne cesse de se complexifier.

Aujourd'hui beaucoup de gens n'utilisent pas Internet pour effectuer des transactions financières et commerciales. Certaines personnes refusent tout simplement de faire confiance à l'authenticité des transactions commerciales sur Internet, comme dans le cas du commerce électronique, la plupart des fournisseurs de services recueillent des données personnelles, ils négligent d'informer les consommateurs de l'utilisation qui en sera faite, de la manière dont elles seront sécurisées,…

Ainsi, beaucoup d'entre eux (les consommateurs – utilisateurs) se montrent retissant au moment de divulguer leurs renseignements personnels et privés pour des questions de sécurité.

En réponse à ces problématiques, la protection de la vie privée a été premièrement affirmée en 1948 par la Déclaration universelle des droits de l'homme des Nations unies (art. 12). Et depuis le 17 juillet 1970 en France, via l'article 9 du Code civil. Cette protection était limité contre toute intervention arbitraire telle que la protection du domicile, de secret professionnel et médical, la protection de l'image, la protection de l'intimité**,** etc**…**

Concrètement, la protection de la vie privée repose dans sa mise en place en plus des moyens légaux (par exemple loi informatique et libertés), des moyens organisationnels (règles internes à l'organisation), et des moyens techniques (exemple cryptographie), etc.

Dans le cadre technique, l'objectif de ce mémoire est de découvrir quelques langages de protection des données personnelles qui permettent à l'utilisateur de mieux comprendre les pratiques d'usage appliquées sur ses données personnelles par le fournisseur de service. Ensuite, de remettre à l'utilisateur le contrôle de ses données personnelles en mettant en place un processus de négociation entre ses préférences personnelles et les pratiques d'usage appliquées par le SP en matière de protection de données personnelles.

Ce mémoire est organisé comme suit :

Le chapitre 1 porte sur la vie privé et la protection des données personnelles. Il détaille la définition de l'aspect vie privée en générale, et la sphère privée d'un individu en particulier. Ensuite, il présente la gestion de la protection des données personnelles collectées et stockées par les SP(s), en commençant par la présentation des aspects légaux du droit à la sphère privée (la loi française 78-17 « Informatique et Libertés », et la loi française 2004-801), ensuite en présentant les 7 principes clés de la protection des données personnelles et en concluant par les technologies de protection des données personnelles.

Le chapitre 2 est consacré à la présentation des langages de protection des données personnelles existants permettant la formalisation des aspects légaux liées à la protection des données ; et les modèles de contrôle d'accès (DAC, MAC, IBAC, RBAC et ABAC) mettant en pratique un contrôle affiné. Ce chapitre est conclu par l'étude du langage XPACML (ses principes, son modèle et son architecture de control d'accès) mettant en place les aspects légaux dans le modèle de contrôle d'accès XACML (basé sur le modèle ABAC) coté utilisateur.

La négociation étant ainsi permise avec tout agent du SP, nous concluant ce mémoire par le chapitre 3 qui adopte le concept d'ontologie et du web sémantique en vue d'apporter plus de souplesse au protocole de négociation. Notre travail d'implémentation met en lumière les différents aspects de négociation sur lesquels nous avons travaillé.

CHAPITRE I

1. INTODUCTION:

La notion de vie privée d'un individu (ainsi que la question de sa protection) est apparue et a évolué dans les sociétés occidentales parallèlement à l'émergence de l'ensemble des libertés individuelles. La *Magna Carta*, la Glorieuse Révolution britannique, les textes des Lumières, la Déclaration d'Indépendance des Etats-Unis d'Amérique ou la Déclaration des Droits de l'Homme et du Citoyen sont quelques une des étapes historiques témoignant de l'importance croissante donnée à l'individu dans la société. On y voir poindre les délicats équilibres et points de compromis qui existent entre les libertés individuelles et le bien collectif. La gestion de ces équilibres occupe toujours une place prépondérante dans l'organisation de la chose publique des nations policées.

Parmi ces libertés individuelles, le droit à la vie privée et à sa protection apparait historiquement assez tôt, mais uniquement dans son principe. L'apparition de la photographie, notamment, sera à la source d'études fondatrices comme l'article « *the right of privacy* » de Samuel D.Warren et Louis D. Brandeis, posant en 1890 les fondements du droit des individus à protéger leur image et, de manière plus générale, leur vie privée **[1]**.

La délimitation du droit à la vie privée (et à sa protection) n'est pas un problème trivial, car cette notion dépend de la culture, de l'histoire locale et des sensibilités individuelles. La conscience collective d'un besoin de protection de la vie privée peut notamment être influencée par la mise au grand jour de failles significatives dans cette dernière.

L'avènement de l'ère numérique, dans le dernier quart du vingtième siècle, a déclenché de nouvelles réflexions sur le sujet, les nouveaux outils introduisant de nouveaux risques.

De la même manière que la photographie a permis la propagation des images, Internet et les applications distribuées d'une manière générale permettent le partage, la duplication, le traitement automatisé de nombreux types d'informations. Le public prend assez rapidement conscience des possibilités offertes par les nouveaux développements techniques mais également des risques parallèlement encourus par leurs données personnelles.

Néanmoins, il reste un fossé entre le droit formellement exprimé, son application plus ou moins stricte par les acteurs du monde informatique et sa compréhension par les usagers, souvent rebutés par la forme des textes.

Les exemples du commerce électroniques (auxquels nous nous intéressons) et des réseaux sociaux, sont des applications emblématiques associées à deux phases d'expansion du

réseau internet, permettent d'illustrer ces difficultés. Dans le premier cas, l'utilisateur est forcé, pour acquérir un produit ou un service, de transmettre des informations personnelles et potentiellement sensibles, comme son adresse ou son numéro de carte de crédit au fournisseur de service (SP). Dans les réseaux sociaux, les utilisateurs sont incités, afin de profiter au mieux de l'environnement personnalisé, à dévoiler énormément d'informations sur eux-mêmes, sans toujours mesurer le risque associé.

1.1. Définitions :

Les notions liées à la vie privée peuvent difficilement être définies sans s'intéresser aux différentes significations du terme *privacy* en anglais. En effet, si l'on peut le faire correspondre en français à la notion assez générale de « caractère privé » d'une chose, ce mot semble être considéré comme recouvrant un certain nombre de concepts liés. Suivant leur culture et leur point de vue, les auteurs les plus consciencieux prennent soin de préciser ce que désigne pour eux le terme *privacy*, sans l'utiliser indifféremment pour le droit à la vie privée (*right of privacy)* ou la protection de la vie privée (*privacy protection).*

Nous considérerons ici le terme *privacy* comme une traduction imparfaite de l'expression

« vie privée », autorisant à définir des termes dérivés.

Par la suite nous donnons les définitions de base introduites par **[2]** sur lesquelles nous nous somme basé dans ce travail.

La mention du concept de *vie privée* éveille chez tout un chacun un ensemble de problématiques liées à notre vie quotidienne ou à notre perception de procédés techniques ou liées à un certain contexte professionnel. Ainsi, la capacité à cacher un certain nombre de choses sur soi au public en général, à des collègues, à des connaissances, relève nécessairement de notre droit à la vie privée. La notion de surveillance des activités d'un individu, l'enregistrement ou le traitement d'informations le concernant, le fait d'entrer en communication avec lui sur la base des résultats d'un tel traitement sont autant d'actions en lien étroit avec la notion de vie privée ou de sphère privée. Le concept semble donc composite et par conséquent difficile à cerner. Néanmoins, certains auteurs ont proposé des définitions très restreintes dont on peut se demander si elles correspondent vraiment à cette vision naïve et intuitive de la vie privée.

Définition 1.1: La sphère privée d'un individu est l'ensemble des informations se rapportant à lui-même, qu'il considère comme sensibles et donc dignes d'être protégées. Cette sphère est personnelle (l'individu est le propriétaire des informations qu'elle contient), personnalisable (l'individu décide des informations qu'elle contient), dynamique (les informations peuvent y être ajoutées ou en être retirées) et dépendante du contexte (les

informations qu'elle contient peuvent, en nature et en nombre, dépendre du temps, des activités de l'individu ou d'autres paramètres).

Définition 1.2: Le droit à la vie privée d'un individu est sa prétention aux caractères personnel, personnalisable, dynamique et contextuel de sa sphère privée ainsi qu'au contrôle de la diffusion, de l'utilisation et de la conservation des informations contenues dans sa sphère privée, quelles que soient la représentation de ces informations et la localisation de cette représentation.

Définition 1.3: (Protection de la vie (ou de la sphère) privée). La protection de la vie privée est l'ensemble des mesures techniques visant à assurer le respect du droit à la vie privée.

Définition 1.4: (Protection des données personnelles). La protection des données personnelles consiste en l'ensemble des mesures techniques visant à assurer le respect du droit à la vie privée, limitées aux données de la sphère privée d'un utilisateur explicitement représentées sous forme numérique et mises en jeu dans le cadre d'une application informatique.

2. LA GESTION DE LA PROTECTION DES DONNEES PERSONNELLES COLLECTE ET STOCKEES PAR LES SPS :

Une fois en transaction avec un fournisseur de services (SP), et soumis à un contrat d'usage assez lent et incompréhensible; l'utilisateur est contraint à dévoiler un certain nombre de données personnelles requises afin d'avoir accès au service. Sans pour autant savoir l'essor de ces données, une fois migrées de son poste.

Le fournisseur de service selon un certain nombre de paramètres (son type et son secteur, ... etc.) prétend un usage définit, une durée de rétention délimitée et un certain nombre de moyens de protection des données personnelles acquises comme suit:

2.1. Les données personnelles collectées:

Deux catégories de données sont stockées et gérées par les fournisseurs de service:

- ✓ Les données fournies par l'utilisateur, en général au moment de la souscription au service:
- ✓ Données nécessaires à l'identification de l'usager ou du client (nom, prénom, date de naissance, sexe, etc).
- ✓ Données nécessaires à la facturation du service (adresse de facturation et moyen de paiement)

- ✓ Données nécessaires à la fourniture du service
- ✓ Données complémentaires, optionnelles, que le fournisseur de service peut collecter pour un usage marketing à condition d'avoir obtenu l'accord explicite du client.
- ✓ Données complémentaires demandées en échange de la fourniture d'un service gratuit sur Internet
- ✓ Les données collectées pendant la fourniture du service. Il s'agit des données les plus sensibles en matière de vie privée. Ces données dépendent du type de service fourni :
- ✓ Opérateurs de télécommunication - tickets de facturation et d'usage permettant de retrouver
- ✓ l'ensemble des correspondants d'une personne (émission et réception)
- ✓ l'ensemble des sites web visités et des pages vues
- ✓ l'ensemble des lieux physiques (cellules) où un téléphone mobile allumé a été vu, même s'il n'a pas appelé
- ✓ Services bancaires
- ✓ ensemble des transactions de paiement horodatées avec les coordonnées de la contrepartie (commerçant)
- ✓ incidents de paiement
- ✓ crédits et incidents de remboursement
- ✓ revenus récurrents
- ✓ Services de santé (incluant Sécurité sociale et assurance) - information sur les maladies d'une personne et les médicaments achetés
- ✓ Services d'éducation
- ✓ Fournisseurs de service gratuit sur Internet : Données d'usage du service collectées par le fournisseur de service (tel que l'historique de navigation, les choix utilisateurs, les sites et documents consultés, contenu des messages mails échangés) que celui-ci utilise pour mieux connaitre ses usagers ou financer ses activités. Il est à noter que de telles informations font l'objet de commerce lucratif entre les fournisseurs de services et d'autres partenaires de la sphère Internet (Data Brokers, Marketing, Publicitaires). Ces utilisations étant de plus en plus contestée aujourd'hui car faites en totale opacité, sans consentement de l'utilisateur et générant des bénéfices substantiels souvent disproportionnés avec le service fourni.
- ✓ Fournisseurs de service payant sur Internet: Les mêmes problématiques que le cas précédent peuvent se poser lorsque le fournisseur d’accès payant se permet d'utiliser les donnés collectées dans un but significativement différent de celui accepté par l'utilisateur.

2.2. L’usage des données:

Les données collectées sont utilisées pour fournir le service, le facturer et gérer les réclamations mais aussi de multiples autres utilisations. (en se référant à la **CNIL (**Commission nationale de l'informatique et des libertés**)**).

Mais les préoccupations en matière de protection de la vie privée concernent :

- ✓ La capacité du fournisseur de service à protéger les données stockées, d'un accès extérieur malveillant.
- ✓ L'usage qui est fait des données collectées pour une usage non prévu par le contrat (action marketing, location de fichiers, etc.).
- ✓ Le partage de ces informations avec des tiers non autorisés.
- ✓ L'usage des données pour des buts non attendus ou non désirés par le Citoyen (communément appelé le "Data Subject").

Les établissements financiers et les opérateurs de télécommunication possèdent des informations extrêmement sensibles sur la vie privée de leurs clients. La CNIL veille particulièrement à ce que ces données ne soient pas traitées pour des actions marketing sans accord préalable.

2.3. Durée de conservation des données:

La loi sur la protection des données personnelles exige que les données ne soient pas conservées plus longtemps que nécessaire pour fournir le service demandé, néanmoins d'autres nécessités légales peuvent influencer cette durée.

- ✓ La conservation des données administratives peut-être extrêmement longue, et durer toute la vie (Cotisations retraites, transactions immobilières par exemple).
- ✓ La durée de conservation par les fournisseurs de services non administratifs est soumise en France à l'accord préalable de la CNIL.

Les données ne doivent être en principe conservées que pendant la durée nécessaire à fourniture du service, ainsi que pendant une période complémentaire raisonnable. L'article L110-4 du Code de commerce français prévoit que « Les obligations nées à l'occasion de leur commerce entre commerçants ou entre commerçants et non-commerçants se prescrivent par cinq ans si elles ne sont pas soumises à des prescriptions spéciales plus courtes ».

- ✓ En cas de résiliation du service, les données nécessaires à la collecte du dernier paiement peuvent être conservées tant que la dernière facture est considérée comme exigible.
- ✓ En cas de litige, la durée peut être de plusieurs années.

Néanmoins la loi peut exiger une conservation beaucoup plus longue des données; il en est ainsi :

- ✓ De la conservation des données bancaires qui peuvent être utilisées comme preuve dans des conflits commerciaux.
- ✓ Des données de télécommunication qui peuvent servir de preuve lors des procès au pénal ou pendant les enquêtes judiciaires qui peuvent avoir lieu longtemps après les faits.

2.4. Protection des données:

C'est un point extrêmement sensible, puisque l'on sait que malgré les moyens de protection mis en œuvre, aucun système ne peut être considéré comme inviolable.

Ainsi le fournisseur de service doit :

- ✓ S'assurer que l'accès à son système d'information est protégé des accès extérieurs. Il doit également s'assurer de l'impossibilité d'altérer les données.
- ✓ Respecter les règles en matière de stockage des données bancaires (interdiction de stocker en clair les numéros de carte par exemple)
- ✓ Éviter de stocker les mots de passe des clients, mais se contenter de stocker l'information permettant de contrôler les mots de passe.

3. ASPECTS LEGAUX DU DROIT A LA SPHERE PRIVEE:

Dans le cadre de nos travaux, nous nous intéresserons, au cadre légal existant en France. Depuis 2004, il est globalement accepté que le droit français en matière de protection de la vie privée est compatible avec la législation de la plupart des pays de l'Union Européenne (se situant parmi les plus protectrices, derrière l'Espagne notamment), ou avec celle du Canada, par exemple. La législation fédérale américaine, par contre, est très différente.

En France, depuis la création du code civil en 1803, le droit des individus à la vie privée est affirmé par son article 9 **[3],** mais de manière générique et appelant une interprétation appuyée au regard des moyens de traitements mis à disposition par le développement de l'informatique. En effet, les moyens techniques actuels permettent la mise en œuvre de collectes et de traitements automatisés des données personnelles, contexte qui n'était pas envisageable en 1803. Cette adaptation débute en 1978 par une loi nationale **[4]** et va se poursuivre dans le cadre de l'Union Européenne. L'essentiel du cadre législatif en

matière de protection de la vie privée réside dans deux lois nationales **[4], [5]**, reprenant deux directives européennes **[6], [7].**

3.1. La loi française 78-17 « Informatique et Libertés »:

La spécificité des risques induits par les traitements automatisés des informations (et notamment ceux mis en œuvre par les administrations publiques), a motivé la création d'une nouvelle loi en 1978. La France est alors le premier pays européen à inclure dans son droit national des dispositions spécifiques concernant les traitements informatiques de données personnelles. La loi 78-17 du 6 janvier 1978 **[4]**, « relative à l'informatique, aux fichiers et aux libertés » (dite loi « Informatique et Libertés ») parle de traitements automatisés sur des données nominatives. Ce texte pose des principes qui sont depuis rentrés dans la culture française, de par les mentions légales apparaissant systématiquement dans les formulaires de collecte de données.

Les données nominatives sont définies comme celles pouvant être rattachées à une personne physique, de manière directe ou indirecte. Cette précision permet d'englober les données pouvant être liées à une personne après recoupement avec d'autres informations.

Le responsable d'un traitement automatisé ou d'une base de données contenant ce type d'informations ne les possède pas, mais est responsable de leur sécurité. Le texte introduit l'obligation fondamentale d'informer l'intéressé sur divers points concernant le traitement et de recueillir son consentement (sauf dans certains cas particuliers prévus par la loi, comme par exemple l'accomplissement d'une mission de service public). L'intéressé jouit également d'un droit d'accès et de rectification des données collectées, exerçable auprès du responsable du traitement.

La loi impose une limitation sur la durée de conservation des données. Celles-ci devront en effet être détruites une fois qu'elles ne sont plus nécessaires au traitement.

3.2. La loi française 2004-801:

La loi française numéro 2004-801 « relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel » **[5]** a pour principal objectif de modifier la loi 78-17 déjà existante, pour la mettre en conformité avec les directives européennes de 1995 et 2002.

Cette loi entérine le terme d'informations personnelles en remplacement des informations nominatives. La notion de justification de la collecte et du traitement (principe de collecte minimale) est transposée dans le texte français en utilisant les termes de la

directive de 1995. Le texte de loi consolidé modifie également les pouvoirs de la CNIL. En particulier, l'étendue de la déclaration préalable est largement diminuée, et la commission dispose d'un rôle de contrôle plus important, assorti d'un pouvoir de sanction.

La principale contribution de cette loi aux principes généraux de la protection des données à caractère personnel dans le cadre des traitements automatisés consiste en l'introduction dans la loi française de la notion de justification. Pour le reste, elle reformule la loi existante pour l'harmoniser avec la directive européenne.

Enfin, en accord avec les directives européennes de 1995 et 2002, sept principes de protection des données personnelles ont été introduites par différents travaux.

4. LES 7 PRINCIPES CLES DE LA PROTECTION DES DONNEES PERSONNELLES:

4.1. Le principe de finalité :

Les données à caractère personnel ne peuvent être recueillies et traitées que pour un usage déterminé et légitime, correspondant aux missions de l'établissement, responsable du traitement. Tout détournement de finalité est passible de sanctions pénales.

4.2. Le principe de proportionnalité:

Seules doivent être enregistrées les informations pertinentes et nécessaires pour leur finalité.

4.3. Le principe de pertinence des données :

Les données personnelles doivent être adéquates, pertinentes et non excessives au regard des objectifs poursuivis.

4.4. Le principe de durée limitée de conservation des données:

C'est ce que l'on appelle le droit à l'oubli. Les informations ne peuvent être conservées de façon indéfinie dans les fichiers informatiques. Une durée de conservation doit être établie en fonction de la finalité de chaque fichier.

Au-delà, les données peuvent être archivées, sur un support distinct.

4.5 .Le principe de sécurité et de confidentialité:

Le responsable du traitement, est astreint à une obligation de sécurité. Il doit faire prendre les mesures nécessaires pour garantir la confidentialité des données et éviter leur divulgation :

- ✓ Les données contenues dans les fichiers ne peuvent être consultées que par les services habilités à y accéder en raison de leurs fonctions.
- ✓ Le responsable du traitement doit prendre toutes mesures pour empêcher que les données soient déformées, endommagées ou que des tiers non autorisés y aient accès. S'il est fait appel à un prestataire externe, des garanties contractuelles doivent être envisagées.
- ✓ Les mesures de sécurité, tant physique que logique, doivent être prises. (par ex : Protection anti-incendie, copies de sauvegarde, installation de logiciel antivirus, changement fréquent des mots de passe alphanumériques d'un minimum de 8 caractères.)
- ✓ Les mesures de sécurité doivent être adaptées à la nature des données et aux risques présentés par le traitement.

4.6. Le principe de transparence :

La loi garantit aux personnes l'information nécessaire relative aux traitements auxquels sont soumises des données les concernant et les assure de la possibilité d'un contrôle personnel. Le responsable du traitement de données personnelles doit avertir ces personnes dès la collecte des données et en cas de transmission de ces données à des tiers.

4.7. Le principe du respect du droit des personnes :

4.7.1. Informer les intéressés*:*

Lors de l'informatisation de tel ou tel service, ou lorsque des données sont recueillies par exemple par voie de questionnaires, les usagers concernés et le personnel de l'organisme doivent être informés de la finalité du traitement, du caractère obligatoire ou facultatif du recueil, des destinataires des données et des modalités d'exercice des droits qui leur sont ouverts au titre de la loi "Informatique et Libertés" : droit d'accès et de rectification mais aussi, droit de s'opposer, sous certaines conditions, à l'utilisation de leurs données.

Cette information doit être diffusée, par exemple, au moyen d'affiches apposées dans les services recevant du public, de mentions portées sur les formulaires de collecte papier et électroniques, ainsi que sur les courriers et courriels adressés aux personnes dont les données sont collectées. Des modèles de mentions d'information sont disponibles sur le site

4.7.2 .Les droits d'accès et de rectification**:**

Toute personne peut demander communication de toutes les informations la concernant contenues dans un fichier détenu par l'établissement et a le droit de faire rectifier ou supprimer les informations erronées.

4.7.3. Le droit d'opposition**:**

Toute personne a le droit de s'opposer, pour des motifs légitimes, à ce que des données la concernant soient enregistrées dans un fichier informatique, sauf si celui-ci présente un caractère obligatoire.

5. TECHNOLOGIES DE PROTECTION DES DONNEES PERSONNELLES:

Des propositions de plus en plus nombreuses sont présentées comme améliorant la protection de la vie privée. On trouve dans cette approche des travaux « atomiques », traitant un aspect du problème en fournissant un outil informatique ou méthodologique (à l'image du standard Platform for Privacy Preferences du W3C **[8]**), ou encore des propositions composites reposant sur de telles briques fonctionnelles. Cette dernière catégorie recouvre des travaux plus ou moins ambitieux, allant de la spécification de profil utilisateur sécurisé (comme dans la proposition de Stéphanie Riché et Gavin Brebner **[9]**) aux infrastructures intégrées portées par les projets européens PRIME **[10]** ou PISA **[11**], par exemple.

Nous nous proposons d'analyser les quelques outils et principes que l'on retrouve couramment dans les propositions existantes, et qui diffèrent du simple contrôle d'accès non spécifique à la protection des données personnelles.

5.1. Platform for Privacy Preferences:

Le standard P3P du World Wide Web Consortium **[8]** est un outil désormais incontournable de la communication des sites web sur leur politique de protection des données personnelles.P3P est une spécification de documents XML décrivant les politiques de traitement des données personnelles déclarées par un site web. Ces documents sont conçus pour être accessibles par un navigateur à partir de la page d'accueil du site. L'objectif de ce projet est de rationnaliser la manière dont les sites web communiquent sur leurs traitements. Les données présentes dans un document P3P couvrent les aspects suivants :

- L'identité de l'entité collectant les données ;
- La nature des données collectées ;
- La destination (ou justification) de la collecte de données ;
- L'identification des données pouvant être partagées avec des tiers ;
- L'identification de ces tiers ;
- La possibilité offerte ou non aux utilisateurs de modifier la manière dont leurs données sont traitées ;
- Les méthodes de résolution des conflits éventuels (et le ressort juridique compétent) ;

- La durée de rétention de chacune des informations collectées ;
- Un lien vers une version de la politique lisible par un humain.

Il faut bien comprendre ici, et les documents du W3C le soulignent, que P3P n'impose aucune politique minimale, il ne fait que fournir le moyen de l'exprimer. De plus, P3P ne permet pas de vérifier que la politique est effectivement appliquée par le site web en question. P3P a pour seul objectif (comme précisé dans les spécifications) de résoudre le problème de l'information de l'utilisateur, à l'exclusion des cinq autres axes de la protection des données personnelles.

On peut toutefois noter que les diverses extensions à P3P permettent également de traiter partiellement le problème du consentement de l'utilisateur. En effet, le langage APPEL

[12] permet de spécifier des préférences du côté de l'utilisateur. Ainsi, le navigateur est capable de détecter automatiquement (via des moteurs fournis par le W3C) si une politique

P3P est conforme aux préférences APPEL, le fait étant alors considéré comme un consentement a priori de l'utilisateur. Ce système est par exemple utilisé dans le cas simple de la décision d'acceptation d'un cookie par un navigateur.

Les concepteurs de systèmes de protection des données personnelles ont tout intérêt à s'appuyer sur le standard P3P, ou en tout cas à demeurer interopérable avec lui. En effet, il permet de résoudre de manière simple le problème de l'information de l'utilisateur, en étant capable de décrire les divers aspects relatifs au traitement des données. Si les listes de choix prédéfinies pour la spécification du type de traitement, de leur justification ou du type de données restent limitées, elles sont extensibles par le biais de schémas XML. De plus, P3P est déjà largement utilisé par les sites web pour leur communication, et de nombreux outils sont capables de manipuler le formalisme d'une manière ou d'une autre.

Toutes ces raisons poussent à favoriser au maximum l'interopérabilité avec P3P, préférentiellement à d'autres langages de politiques comme SPARCLE **[13]**, moins génériques et moins répandus.

Il faut toutefois rester conscient des limitations de P3P. Tout d'abord, la restriction à un rôle d'information (et éventuellement de consentement). Enfin et surtout, P3P exprime la politique d'un site web indépendamment de tous les types de réglementations que nous avons pu identifier. Un utilisateur n'a alors aucun moyen de savoir si ces politiques respectent telle loi ou telle directive. Il reste donc ici un travail d'information et de raisonnement à effectuer.

5.2. Sticky policies (politiques collantes):

Comme nous venons de le voir, des outils comme P3P n'assurent pas réellement La protection des données. Une fois qu'une politique de traitement est déterminée, il faut donc que les processus de traitement, de transmission et de stockage des données se chargent de l'appliquer. Une approche courante (et commune à de nombreuses propositions) consiste à

attacher aux informations sensibles les méta-données de description de la politique de sécurité associée, les applications s'engageant à respecter cette « politique collante ». Ces sticky policies ont été introduites par Günter Karjoth et Matthias Schunter en 2002 **[13].**

La proposition, dans son principe, attache des règles aux données personnelles, qui ne peuvent être manipulées par l'application que si ces règles sont respectées. On retrouve par exemple cette idée dans l'architecture intégrée du projet PRIME **[14], [15]** ou dans l'architecture proposée parallèlement par Marco Casassa Mont et al. **[16].** Si le concept est intuitif, pratique et adapté à la distribution des applications et des données (permettant un premier pas vers une réelle protection étendue), il ne donne cependant (dans sa version de base) aucune garantie à l'utilisateur quant au respect de la politique par une application distante. Il nous faudra donc détailler comment les sticky policies peuvent être utilisées de manière à réellement contraindre l'utilisation des données à distance. Les différents types d'utilisation de ce concept devront être analysés en fonction des garanties qu'ils fournissent à un observateur distant et au propriétaire des données en particulier.

Il convient également d'observer, en ce qui concerne cette famille de propositions, que la source desdites politiques n'est pas toujours spécifiée. Bien souvent, elle est issue de négociations entre les parties ou directement déduite de politiques déclaratives de type P3P.

Ce mode de fonctionnement ne permet donc pas nativement de prendre en compte les contraintes réglementaires ou légales applicables aux traitements, que ce soit à la création de la politique ou bien au moment du traitement de l'information. Nous sommes donc toujours ici en besoin d'un outil adapté pour manipuler et prendre en compte les contraintes issues des réglementations.

5.3. Gestion déportée des données sensibles:

Des propositions ont également été faites pour permettre aux utilisateurs de profiter de services en ligne tout en évitant à ces derniers de pouvoir tracer leurs activités. C'est un type d'application qui relève donc davantage de l'accès anonyme aux services et des autorisations préservant la vie privée. Néanmoins, certaines de ces propositions se rapportent plus particulièrement à la gestion des données personnelles de l'utilisateur dans ces scénarios. C'est le cas notamment du protocole SAML2.0 **[17]** établi par Liberty Alliance ou du protocole IDsec **[18],** établi par l'IETF. Il consiste en la déportation de la gestion des données utilisateur sur un serveur spécialisé, mettant en œuvre des mécanismes de contrôle d'accès sophistiqués, visant à s'assurer du bien-fondé des différentes demandes d'accès au profil qui lui sont faites.

En préalable au déroulement d'une transaction entre l'utilisateur et le fournisseur de service , l'utilisateur s'identifie auprès du serveur gestionnaire de son profil, qui en retour lui procure un certificat de session (qui servira de jeton d'accès temporaire). Ce certificat est ensuite transmis au service, qui le présente au gestionnaire de profil , accompagné d'une requête concernant le profil de l'utilisateur et d'un certificat de créance sur ses propres propriétés techniques. Le gestionnaire de profil valide le certificat de session et examine le certificat de créance. Si le gestionnaire est satisfait par ce certificat, c'est-à-dire si la correspondance entre la requête et les propriétés, quelles qu'elles soient, du fournisseur de service correspondent aux exigences connues de l'utilisateur, alors les informations de profil sont transmises au service.

Les approches de ce type ont apporté des idées intéressantes, notamment dans le cadre de la gestion des identités virtuelles telle proposée par le projet FC2 (Fédération des Cercles de

Confiances) par exemple ; mais souffrent de limitations discriminantes. Tout d'abord, la localisation des données personnelles de l'utilisateur sur un serveur délocalisé et clairement identifié pose un problème de sécurité mis en avant par les concepteurs mêmes d'IDsec : le serveur gestionnaire, dépositaire de nombreuses données potentiellement sensibles, devient en effet une cible privilégiée pour des attaques informatiques. Cet aspect du problème limilite fortement en faveur d'une gestion des données personnelles directement par leurs propriétaires, de manière distribuée. De plus, ce protocole ne s'inquiète que de l'accès initial aux données et ne fournit aucun moyen pour assurer leur protection étendue.

Enfin, les possibilités offertes à l'utilisateur de spécifier des propriétés techniques à vérifier pour qu'un fournisseur de service puisse accéder à telle ou telle partie de son profil personnel sont très limitées en termes d'expressivité. En effet, pour traiter réellement de protection des données personnelles, il faudrait pouvoir définir des politiques capables de se référer aux 7 principes clés de la protection des données personnelles Définis précédemment

5.4. Agents utilisateurs:

Certaines propositions émanant du domaine des systèmes multi-agents impliquent des agents artificiels **[19]** dans la protection des données personnelles des utilisateurs. Dans ce contexte, les agents désignent des entités logicielles capables d'interagir de manière autonome avec d'autres entités ainsi qu'avec l'environnement dans lequel elles sont situées.

Ces agents peuvent être des briques logicielles destinées à mettre en oeuvre le traitement en lui-même, comme dans l'architecture proposée par John J. Borking **[20].**

Ils peuvent également se présenter sous la forme d'agents mobiles et se déplacer avec les données. Dans des travaux comme ceux de Stéphanie Riché, Gavin Brebner et Mickey

Glitter **[21], [22],** ces agents sont des assistants logiciels au service d'un utilisateur placé au centre de l'application. Ces agents personnels ou agents utilisateurs sont alors chargés de surveiller et de contrôler l'utilisation qui est faite des donné es, de manière que cette utilisation reste conforme avec la politique établie.

Cette approche permet de répondre au problème majeur posé par la gestion déportée des données personnelles par une reprise de contrôle de l'utilisateur sur ses informations, tout en décentralisant les fonctionnalités de raisonnement dans des entités autonomes. L'agent ou les agents utilisateurs permettent d'interfacer les relations entre l'utilisateur humain et les différents services et applications, en lui confiant la responsabilité de certaines décisions (comme par exemple dans le cas du consentement a priori). Dans cette perspective, c'est l'agent utilisateur qui met en œuvre les différents mécanismes (principalement de contrôle d'accès) assurant la sécurité des données personnelles de l'utilisateur.

Le paradigme des agents utilisateurs est donc séduisant à plusieurs titres. Si le principe de l'agent n'est pas en soi un outil de protection des données personnelles, il peut être chargé de leur mise en œuvre.

6. CONCLUSION :

Dans le chapitre 1nous avons vu une généralité de la vie privée et la protection des données personnelles, que la notion de vie privée d'un individu est basée sur des libertés individuelles, parmi ces libertés le droit de la vie privée et sa protection.

Plusieurs définitions de la vie privée droit à la vie privée, protection de la vie privée, protection des données personnelles.

La gestion de la protection des données personnelles par le fournisseur de service.

L'aspect légal de droit à la sphère privée nous nous intéresserons au cadre légale existant en France. Depuis 2004.

Les 7 principes clés de la protection des données personnelles, puis la technologie de protection des données personnelles.

Dans le chapitre 2 on va étudier les différents modèles de contrôle d'accès et le langage XACML


CHAPITRE II

1. INTRODUCTION :

Malgré les avancés en matière législative afin de protéger la vie privée des utilisateurs, ces efforts restent insuffisants pour établir une relation de confiance entre l'utilisateur et le SP. La problématique est encore plus accentuée quand les utilisateurs et le SP proviennent de pays avec différentes législations. Depuis que la protection des données personnelles a été caractérisée comme une des problématiques principales à gérer dans le nouvel air numérique, plusieurs spécifications techniques ont été définies afin de protéger les informations personnelles des menaces de web. Premièrement, le W3C a définit la plateforme P3P **[23]** afin de définir un vocabulaire de protection des données personnelles dans un format compréhensible par la machine. Ensuite différents travaux comme APPEL **[24]**, XPref **[25]** ont été définit afin de laisser l'utilisateur exprimer ses préférences. Ces travaux ont permis de mettre l'utilisateur en action en lui donnant un accès rapide aux politiques du SP qui peut utiliser les données pour des buts non acceptables pour l'utilisateur.

Des travaux comme XPACML **[2]**, sur lequel nous nous basons, a permis de donner plus de contrôle à l'utilisateur, en lui permettant de comparer la politique proposée par le SP à ses préférences. Ceci a été possible en mariant les travaux précédents à un modèle de contrôle d'accès. L'agent utilisateur lit automatiquement les politiques P3P, les affiches à l'utilisateur dans un format compréhensible. Enfin il effectue la comparaison entre les politiques du SP et préférences de l'utilisateur.

Par conséquent des protocoles de négociation de politiques ont été introduits afin de préserver la vie privée de l'utilisateur. Et afin de permettre une telle automatisation de ces protocoles, le concept d'ontologie a été adopté pour la protection de la vie privée.

2. LANGAGES DE PROTECTION DES DONNEES PERSONNELLES EXISTANTS :

La sécurité est requise mais insuffisante pour la protection de la vie privée. Dans un environnement ICT (*information and communication technologies)*, il existe plusieurs outils de sécurité, tel que le protocole SSL (*Secure Sockets Layer*) avec ses différentes versions, développé par Netscape, PGP (Pretty Good Privacy) est un autre outil, ainsi que la cryptographie avec clé publique RSA (algorithme cryptographique de cryptographie asymétrique). De tels outils de sécurité peuvent aider à la protection de la vie privée en limitant l'accès aux informations personnelles aux parties autorisées, mais la protection de la

vie privée requiert plus que cela. Raison pour laquelle les chercheurs innovent de nouveaux outils permettant d'assurer une garantie de la protection de la vie privée sur Internet.

2.1. Platform for Preferences Privacy P3P:

Le W3C a développé les spécifications de P3P afin de permettre aux SPs d'exprimer de façon transparente leurs politiques de vie privée dans un format standard interprétable par machine **[22].** De telles politiques sont traitées automatiquement avec les navigateurs permettant P3P durant les transactions en ligne. Les navigateurs retrouvent la politique P3P, ensuite un agent utilisateur interprète la politique et informe l'utilisateur sur quelles données sont collectées par le site, et comment les données vont être utilisées. Tout conflit entre les pratiques du site et les préférences de l'utilisateur est affiché à l'utilisateur. Ainsi, l'utilisateur est informé comment ses données personnelles sont traitées. Le W3C a présenté la première recommandation de P3P en 2002. La deuxième version de P3P 1.1 a été publiée en 2006. Cette deuxième version a non seulement révisé les erreurs de la première version mais elle a ramené plusieurs extensions. P3P dans ses deux spécifications définit **[22]** :

- Un schéma standard pour les données qu'un site veut collecter, connu sous schéma de données basique de P3P.
- Un ensemble standardisé des usages, récipients, catégories de données, et autres pratiques de révélations.
- Un format XML d'expression de politiques de révélation de protection.
- Des outils pour associer les politiques de protection avec les sites et les cookies.
- Un mécanisme pour transporter les politiques P3P dans le http.

La contribution principale de P3P peut se résumé dans les deux points suivants :

- P3P exprime les politiques P3P traditionnelles dans un format interprétable par la machine car il utilise le standard XML.
- P3P définit un ensemble de vocabulaire pour chaque élément de données utilisé.

En ce qui suit, les principaux éléments XML utilisés par P3P sont listés sous différents tags XML :

- ENTITY : Cet élément donne une description précise de l'entité légale faisant la présentation des pratiques de vie privée.
- DATA : Ce tag exprime la donnée collectée elle-même par le SP. Les items de données sont regroupés par catégories. Dans une politique P3P, les données sont groupées dans un élément nommé Data-Group.

- PURPOSE : Cet élément exprime pourquoi le SP demande les données. En ce qui suit une partie de ses valeurs prédéfinies :
  - current: Les données (informations) collectées peuvent être utilisées par le SP pour compléter l'activité pour laquelle il a été fourni.
  - admin: L'information peut être utilisée pour un support technique du site du SP et son système d'information.
  - develop: L'information collectée peut être utilisée pour personnaliser ou modifier le contenu du site du SP.
  - pseudo-analysis: Le SP collecte les informations non identifiables de l'utilisateur afin de déterminer ses habitudes, ses intérêts ou autres caractéristiques pour des buts de recherche, analyse et éditions de rapports internes.
  - individual-analysis: Les informations collectées peuvent être utilisées pour déterminer des habitudes, des intérêts, ou autres caractéristiques des individus. Ils peuvent également les combiner avec des données identifiantes pour des buts de recherche ou d'analyse de consommateur.
  - telemarketing: Le SP collecte les informations personnelles pour contacter les utilisateurs pour des buts de marketing comme les annonces de nouveaux services via téléphone par exemple.

- RECIPIENT : Ce tag contient la liste des autres SPs avec lesquels le SP actuel va partager les données collectées. La liste suivante est un sous ensemble de ses valeurs prédéfinies :
  - ours: Seulement le SP ayant fait la demande recevra les données collectées.
  - delivery: Le SP peut partager les informations avec les entités légales effectuant la livraison du service.
  - same: Les informations collectées peuvent être envoyées à des entités légales ayant les même pratiques de protection (politiques d'usage) que le SP en question.
  - other-recpient: Le SP peut envoyer les informations collectées à d'autres entités légales ayant des pratiques de protection différentes que lui.
  - public: Le SP mettra certaines informations, dans les répertoires publiques, ou les répertoires de CD-ROM commerciaux.

- RETENTION : Ce tag indique la période de temps durant laquelle les données collectées vont être stockées dans les BDDs du SP. Quelques exemples de ces valeurs peuvent être :

- no-retention: L'information collectée sera gardée par le SP pour une durée très brève.
- stated-purpose: L'information collectée est retenue pour remplir les buts précités.
- legal-requirement: La durée de rétention requise légalement par la loi.
- applicable law: L'information collectée est retenue pour remplir un but cité à prime à abord mais les exigences légales peuvent conduire à la retenir un peu plus longtemps que le but cité.
- indifinitely: L'information collectée est retenue pour une durée de temps indéfinie.

❖ ACCESS : Cet élément exprime le type d'accès, autorisé par le SP pour l'utilisateur possesseur des données collectées.
❖ STATEMENT : Cet élément regroupe les éléments Purpose, Recipient, Retention appliqués pour un élément data-group spécifique.
❖ DATA-GROUP : Cet élément contient les éléments DATA qui sont collectés pour une politique citée dans l'élément STATEMENT.

En général une politique P3P est un ensemble de déclarations qui définissent les pratiques sur un groupe d'éléments de données (data elements). Optionnellement une politique P3P peut définir des extensions. A titre d'exemple, la description humaine d'une déclaration peut être écrite dans l'élément CONSEQUENCE. Aussi, une déclaration de politique peut inclure des CATEGORIES d'éléments qui sont obligatoire d'utiliser pour faciliter l'implémentation et l'utilisation des agents utilisateurs **[23]**.

Pour comprendre une implémentation pratique des spécifications P3P, une politique P3P basique est montrée ci-dessous dans la **Figure 2.1** :

```
<POLICIES xmlns="http://www.w3.org/2002/01/P3Pv1">
 <POLICY ... name="policy">
    <ENTITY/>
       <STATEMENT>
          <CONSEQUENCE>Our Web server collects access logs containing
  this information.
  </CONSEQUENCE>
  <PURPOSE>
     <admin/>
     <current/>
     <develop/>
  </PURPOSE>
  <RECIPIENT>
     <ours/>
  </RECIPIENT>
  <RETENTION>
     <indefinitely/>
  </RETENTION>
  <DATA-GROUP>
     <DATA ref="#dynamic.clickstream"/>
     <DATA ref="#dynamic.http"/>
  </DATA-GROUP>
    </STATEMENT>
 </POLICY>
</POLICIES>
```

Figure 2. 1 : Exemple de politique P3P basique.

Cette politique est définie pour un SP qui collecte seulement les données stockées dans un serveur standards de log. La politique contient un seul élément STATEMENT, qui indique que les données sont collectées pour compléter la transaction actuelle, pour le site web du SP, l'administration du système, pour la recherche et pour développement.

Les données sont utilisées par le SP et ses agents, et sont retenues pour une durée indéfinie.

2.2. Langages APPEL et XPref :

Dans un but de comparaison entre la politique du SP et les préférences de l'utilisateur, W3C a complété la plateforme P3P avec le langage APPEL (Privacy Policy Exchange Language) **[24].** En effet, APPEL est utilisé pour exprimer les préférences de l'utilisateur dans un format interprétable par la machine, vu qu'il est basé sur le langage XML.

Une politique APPEL contient une partie des éléments P3P. Elle réutilise spécifiquement les tags POLICY, STATEMENT et tous leur sous éléments. L'utilisateur spécifie dans un ensemble de règles ses préférences concernant comment il souhaite que ses données soit traitées.

Ces préférences sont ensuite incorporées dans la politique APPEL POLICY sous l'élément RULESET. Chaque élément RULE contient un attribut « behaviour » et optionnellement un attribut « description » qui fournit une explication lisible de la règle. Cet élément « RULE » est utilisé pour exprimer la volonté de l'utilisateur à révéler un groupe de données personnelles sous certaines conditions placées dans l'élément STATEMENT.

L'attribut « behaviour » peut avoir trois valeurs :

- Block : cette valeur indique que l'échange des données dans une règle est refusé.
- Request : cette valeur de comportement permet le traitement de la requête.
- Limited : L'accès aux données requises sera limité. Seulement les données obligatoires sont échangées.

Par conséquent l'agent utilisateur compare automatiquement ces préférences avec la politique du SP. Pour cet objectif, les spécifications d'APPEL définie des algorithmes de comparaison (matching). Ces derniers, utilisent des connectives définies dans les spécifications APPEL afin de comparer les politiques. Ces connectives peuvent avoir six valeurs : Or, And (qui est la connective par défaut), Non-Or, Non-And, Or-Exact et And-Exact. La valeur Or-Exact signifie que la politique est conforme si une expression ou plusieurs expressions sont retrouvées dans la politique du SP. Si cette politique contient des éléments non listés dans la règle, la comparaison échoue. Au contraire la connective And-Exact signifie que la politique est conforme si toutes les expressions qu'elle contient peuvent être retrouvées dans la politique du SP. A l'instar de la connective précédente, si la politique contient des éléments non listés dans la règle, la comparaison échoue. Les préférences exprimées dans une politique APPEL sont formulées comme montré dans la **Figure 2.2** :

```
1  <appel:RULESET xmlns:appel="http://www.w3.org/2001/02/APPELv1" xmlns:p3p="http://www.w3.
       org/2000/12/P3Pv1">
2    <appel:RULE behavior="block" description="Service collects personal data for 3rd
         parties">
3      <p3p:POLICY>
4        <p3p:STATEMENT>
5      <p3p:DATA-GROUP>
6   <p3p:DATA>
7     <p3p:CATEGORIES appel:connective="or">
8       <p3p:physical/>
9       <p3p:purchase/>
10    </p3p:CATEGORIES>
11  </p3p:DATA>
12      </p3p:DATA-GROUP>
13      <p3p:PURPOSE appel:connective="or">
14  <p3p:telemarketing/>
15  <p3p:other-purposes/>
16      </p3p:PURPOSE>
17      <p3p:RECIPIENT>
18  <p3p:same/>
19  <p3p:unrelated/>
20      </p3p:RECIPIENT>
21      <p3p:RETENTION>
22  <p3p:business-practices/>
23  <p3p:undefinitely/>
24      </p3p:RETENTION>
25          </p3p:STATEMENT>
26      </p3p:POLICY>
27    </appel:RULE>
28 </appel:RULESET>
```

Figure 2. 2: Exemple d'une politique basique de préférences APPEL.

Dans cette politique APPEL, l'utilisateur ne permet pas au SP de collecter des données « Purchase » pour des buts de télémarketing ou des buts indéfinis. Il refuse également que ses données soient stockées pendant une période indéfinie. Il ne permet pas également que ses informations soient retenues par le SP sous les pratiques business de ce dernier. Le partage avec des SPs n'ayant pas la même politique que le SP courant est également refusé.

Quoique le langage APPEL a été premièrement définit avec un but limité, elle a fourni un schéma simple et attractive pour l'expression des préférences de l'utilisateur. Néanmoins, quelques travaux **[25], [26], [27],** montrent qu'APPEL contient des limitations.

Mis à part les limites listées dans les spécifications APPEL **[24]** elle-même, comme l'incapacité d'exprimer des règles sophistiquée **[25].** Ces travaux ont identifiés quatre limitations majeures :

- Difficulté de spécifier ce qui est acceptable.
- Une politique convenante peut être rejeté.
- Les extensions P3P ne sont pas supportées.
- Les combinaisons simples sont difficiles à exprimer.

Ces limitations ont eu lieu principalement suite à l'utilisation des connectives, et l'interopérabilité limitée avec P3P.

Pour pallier à ces défauts, un groupe de chercheurs ont suggéré en **[25].** Une amélioration partielle d'APPEL en **[25].** Ils ont proposé par la suite le langage XPref. Ce dernier réutilise deux éléments XML du langage APPEL : RULESET et RULE.

Pour synthétiser, P3P permet de résoudre de manière simple le problème de l'information et de consentement de l'utilisateur, en étant capable de décrire les divers aspects relatifs au traitement des données (justification, conservation, transmission). Toutefois, P3P exprime la politique d'un SP indépendamment de tous les types de réglementation que nous avons pu identifier dans le chapitre 1.

L'expression de la règlementation avec des éléments de politiques de protection de données à elle seule n'est pas suffisante, il faut l'intégrer à une solution d'autorisation offrant un niveau de granularité très fin permettant la gestion d'accès aux données personnelles avec des règles applicables. D'où la nécessité d'étudier les infrastructures de gestion d'accès à base de politiques.

3. INFRASTRUCTURES DE CONTROLE D'ACCES :

3.1. Définition :

Le contrôle d'accès consiste à vérifier si une entité demandant d'accéder à une ressource a les droits nécessaires pour le faire.

Un contrôle d'accès offre ainsi la possibilité d'accéder à des ressources physiques ou logiques.

Le contrôle d'accès comprend généralement 3 composantes :

3.1.1. L'authentification :

L'authentification est la procédure qui consiste, pour un système informatique, à vérifier l'identité d'une entité (personne, ordinateur...), afin d'autoriser l'accès de cette entité à des ressources (systèmes, réseaux, applications...). L'authentification permet donc de valider l'authenticité de l'entité en question.

L'identification permet donc de *connaître* l'identité d'une entité alors que l'authentification permet de *vérifier* cette identité.

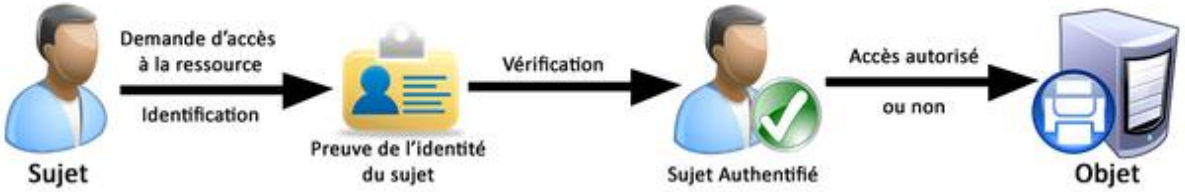

Figure 2. 3: exemple d'authentification.

3.1.2. Autorisation :

Cette phase consiste à vérifier que l'utilisateur maintenant authentifié dispose des droits nécessaires pour accéder au système. Elle est parfois confondue avec la précédente sur de petits systèmes, mais sur des systèmes plus importants, un utilisateur peut tout à fait être authentifié (ex : membre de l'entreprise) mais ne pas avoir les privilèges nécessaires pour accéder au système (ex : page réservée aux gestionnaires).

3.1.3. Traçabilité (Accounting) :

Pour lutter contre les usurpations de droits, il est souhaitable de suivre les accès aux ressources informatiques sensibles (heure de connexion, suivi des actions, …).

3.2. Modèles de contrôle d'accès :

La politique de contrôle d'accès est une vue haute et abstraite du contrôle d'accès. Un niveau intermédiaire permettant de faire le pont entre la politique et son implémentation est le modèle de contrôle d'accès. Le modèle va permettre de supporter la politique définit. Un modèle de contrôle d'accès remplit différents propriétés de sécurité. Par exemple le modèle RBAC (Role Based Access Control) donne la possibilité de séparer les rôles qu'une personne peut endosser à un instant t. Le modèle MAC (Mandatory Access Control) quant à lui permet de garantir la confidentialité des données. Il existe ainsi différents modèles qui sont plus ou moins adaptés à la politique de sécurité de l'entreprise. Le modèle DAC (Discretionnary Access Control) donne la possibilité au propriétaire de la ressource de gérer lui-même les autorisations.

Depuis 50 ans, plusieurs modèles de contrôle d'accès ont été successivement proposés. On peut citer DAC, MAC, RBAC, TBAC, TMAC ou encore ORBAC.

3.2.1. Contrôle d'accès discrétionnaire DAC :

Le contrôle d'accès discrétionnaire ou « Discretionary Access Control » *(DAC)* permet à un sujet d'attribuer des permissions à d'autres sujets. Ce contrôle d'accès est flexible mais il peut générer des erreurs.

3.2.2. Contrôle d'accès obligatoire MAC :

Mandatory access control (MAC) ou **contrôle d'accès obligatoire**

Le contrôle d'accès obligatoire est utilisé lorsque la politique de sécurité des systèmes d'information impose que les décisions de protection ne doivent pas être prises par le propriétaire des objets concernés, et lorsque ces décisions de protection doivent lui être imposées par le système. Le contrôle d'accès obligatoire doit permettre d'associer et de gérer des attributs de sécurité relatifs à cette politique, sur les fichiers et processus du système.

3.2.3. Contrôle d'accès basé sur l'identité :

Le contrôle d'accès basé sur l'identité (IBAC - Identity Based Access Control) [Lampson, 1971] est parmi les premiers modèles de contrôle d'accès. Ce modèle introduit les concepts fondamentaux de sujet, d'action et d'objet.

L'objectif de ce modèle IBAC est de contrôler tout accès direct des sujets aux objets via l'utilisation des actions. Ce contrôle est basé sur l'identité du sujet et l'identificateur de l'objet, d'où le nom du modèle IBAC.

Dans IBAC, une permission a le format suivant : un sujet a la permission de réaliser une action sur un objet. La politique d'autorisation qui permet de spécifier les permissions est définie grâce à l'utilisation d'une matrice de contrôle d'accès dans laquelle les lignes et colonnes de la matrice correspondent respectivement à l'ensemble des sujets et des objets du système d'information.

	Objet 1	Objet 1	Objet 1
Sujet 1	rw	r	w
Sujet 2	r	-	rw
Sujet 3	w	-	r

Tableau 2.1: matrice de contrôle d'accès de modèle IBAC

Cependant la limite de ce modèle est sa mise à l'échelle. En effet, la politique devient complexe à maintenir lorsque le nombre d'entités est important.

Le modèle RBAC introduit une abstraction de l'entité sujet qui devient rôle, permettant ainsi de réduire cette complexité.

3.2.4. Contrôle d'accès à base de rôle :

Le modèle de contrôle d'accès à base de rôle (RBAC) a été proposé pour présenter une nouvelle organisation des droits centrée sur le concept de rôle.

L'objectif principal de ce modèle est de permettre la structuration de l'expression de la politique d'autorisation autour du concept de rôle (un concept organisationnel) : des rôles sont affectés aux utilisateurs conformément à la fonction attribuée à ces utilisateurs dans l'organisation. Le principe de base du modèle RBAC est de considérer que les autorisations sont directement associées aux rôles.

Dans RBAC, les rôles reçoivent donc des autorisations pour réaliser des actions sur des objets. Comme IBAC, le modèle RBAC ne considère que des autorisations positives (permissions) et fait donc l'hypothèse que la politique est fermée. Un autre concept introduit par le modèle RBAC est celui de session. Pour pouvoir réaliser une action sur un objet, un utilisateur doit d'abord créer une session et, dans cette session, activer un rôle qui a reçu l'autorisation de réaliser cette action sur cet objet. Si un tel rôle existe et si cet utilisateur a été affecté à ce rôle, alors cet utilisateur aura la permission de réaliser cette action sur cet objet une fois ce rôle activé. Lorsqu'un nouveau sujet est créé dans le système d'information, il suffit d'affecter des rôles au sujet pour que ce sujet puisse accéder au système d'information conformément aux permissions accordées à cet ensemble de rôles.

3.2.5. Contrôle d'accès à base d'attributs (ABAC) :

Le modèle ABAC défini par L. Wang, D. Wijesekera, S. Jajodia, propose de définir les droits d'accès en fonction des caractéristiques des identités. A l'instar du modèle IBAC, la politique des droits d'accès peut être matérialisée par une matrice, mais en ne se basant pas sur les identités. De ce fait, les droits d'accès à une ressource ou un service sont définis pour un ou plusieurs attributs que les identités sont susceptibles de posséder. Ce paradigme offre donc plus de flexibilité. De plus, en définissant un attribut se rapprochant de la notion de rôle, ABAC permet de simuler le comportement d'un modèle RBAC, mais le généralise en ne limitant pas les droits d'accès aux seuls utilisateurs présents dans l'organisation. Il permet notamment de déterminer des droits d'accès avec une granularité plus fine. De plus, en définissant un rôle comme un ensemble d'attributs, il est plus facile de gérer les conflits. Par ailleurs, la gestion des droits d'accès est facilitée, car elle ne nécessite pas d'informations supplémentaires.

Comme son nom l'indique, le modèle ABAC définit les autorisations d'accès en se basant sur des caractéristiques de chaque entité, appelés attributs. Trois groupes d'attributs se distinguent selon le type de l'entité à laquelle ils s'appliquent :

- Les attributs des sujets : un sujet est une entité qui peut agir sur une ressource. A chaque sujet on associe des attributs qui définissent son identité et ses caractéristiques. Par exemple le rôle du sujet peut aussi être considéré comme un attribut, tout comme le nom, le prénom, ou le titre, etc.
- Les attributs des ressources : La ressource est un objet du système sur lequel un sujet peut agir. Autrement dit, c'est une entité qui peut être accessible à un sujet. Une ressource peut être un fichier, un service, etc. A chaque ressource est associée des attributs, variables selon sa nature, mais qui peuvent être : son type, le nom de son auteur, son propriétaire, la date de modification, etc.
- Les attributs d'environnement : l'environnement peut être décrit par des informations opérationnelles, techniques, liées à la situation ou encore au contexte dans lequel l'accès à l'information se produit. La particularité du modèle ABAC, est la prise en compte du contexte d'exécution du système, en définissant des attributs d'environnement, comme par exemple : la date, le niveau de sécurité du réseau, le débit de la connexion, etc.

3.3. Modèles de contrôle d'accès sensibles à la protection des données :

Les modèles de contrôle d'accès listés ne considèrent pas la protection des données personnelles comme un premier objectif. Les modèles MAC (Mandatory Access Control), DAC (Discretionary Access Control), et RBAC (Role-Based Access Control) ne remplissent que quelques besoins de protection des données personnelles **[2].** Actuellement, de plus en plus de modèles s'intéressent à inclure des éléments de protection, et principalement l'élément « purpose » comme élément principal de contrôle d'accès. Seulement, la spécification de cet élément n'est pas suffisante à elle-même pour définir des politiques de protection ou des préférences de l'utilisateur. Qui Ni et al., proposent des extensions au modèle RBAC **[2]** afin de prendre en compte des contraintes liées à la protection des données. Pour ce la, ils définissent une famille de modèles P-RBAC (Privacy-aware RBAC), où les politiques de protection sont exprimées à travers des affectations de permissions. Ces affectations ne correspondent pas aux permissions définies dans RBAC à cause de la présence des nouveaux éléments « purposes » et « conditions » dans les politiques d'accès. Un langage spécifique LC0 a été proposé afin de permettre la définition des conditions. Une permission explicite définie alors les buts de l'action (l'élément « purpose »), sous quelles conditions,

ainsi que les obligations qui doivent être accomplies après l'accès. Ce travail développe des algorithmes d'analyse de conflits, permettant de détecter les conflits entre les différentes affectations de permissions. Donc, les trois extensions principales sont : l'élément « purpose », la définition des obligations, et un langage dédié aux conditions. Une permission sensible est ainsi modélisée selon des privilèges qui ont la forme générale suivante : role × action × data × purpose × conditions × obligation.

Le modèle de contrôle d'accès à base de « purpose » proposé en **[2]**, tend à appliquer des éléments de protection dans les domaines qualifié de non-confiance. Son objectif est de maintenir la consistance entre la politique de protection et les pratiques déclarées. Ils s'attachent à la formalisation des différents types de buts (purposes). Ensuite, les auteurs spécifient des invariants correspondants aux besoins de protection dans une politique de protection. L'objectif de représentation de ces invariants est de fournir une interprétation claire et non-ambigüe des politiques. Les buts (purposes) sont divisés en deux classes : des buts désirés, et des buts d'accès. Les buts désirés sont liés à l'objet d'accès, spécifient ainsi les usages pour lesquelles une donnée est accédée. Les buts d'accès sont liés à l'accès aux données et spécifient les intentions pour lesquelles une donnée est accédée. Chaque demandeur doit déclarer le but d'accès, le système valide le but d'accès cité afin d'être sûr que l'utilisateur est permit pour le but d'accès. Le but d'accès doit être compatible avec le but désiré afin de permettre l'accès.

Dans chaque organisation, les objets (données) sont organisés en utilisant des types d'objets. L'entité « rôle » a été étendue afin de prendre en compte le rôle conditionnel, qui est basé sur les attributs du rôle et les attributs du système **[02].**

Le PuRBAC (Purpose-Aware Rôle-Based Access Control) proposé dans **[28]**, étend le modèle RBAC en modélisant des besoins liés à la protection de la vie privée. Les buts « purposes » représentent l'entité centrale. Ils représentent l'intermédiaire entre le rôle et les entités de permission. Le modèle définit des hiérarchies de rôles et de buts (pruposes). Il supporte l'expression des contraintes et des obligations, et les définit comme des conditions sur l'affectation des permissions aux différents buts.

Ensuite, les buts (purposes) sont assignés aux rôles. La requête d'un utilisateur est formée par un identifiant de session, un but, et une permission requise. L'autorisation peut être requise pour des buts liés au rôle actif. Il y a une autre différence majeure avec le modèle RBAC, quand une requête est soumise à l'ADF (Access Decision Function), il peut soit refuser l'accès, ou définir une autorisation conditionnelle. Les auteurs modélisent trois types de conditions : contraintes, pre-obligations, et post-obligations.

Les contraintes sont utilisées pour vérifier l'information basée sur les variables des données dans le système. Par exemple, le consentement du possesseur des données est considéré comme une contrainte. Les Pre-obligations désignent que l'utilisateur doit exercer quelques actions avant d'avoir accès. Ceci peut inclure par exemple : la réauthentification de l'utilisateur avant qu'il accède à des données sensibles, ou l'ajustement des données.

Les post-obligations, représentent les obligations après autorisation d'accès. Ceci peut concerner par exemple, la politique de rétention d'une donnée qui programme la suppression des données.

Les auteurs argumentent que les buts (purposes) d'utilisation sont considérés comme une entité séparée car la politique de protection habituelle dépend du but d'utilisation, deuxièmement ils supposent qu'il y a une relation entre la notion de « purpose » dans les politiques de protection et la notion de « rôle » en RBAC, troisièmement décomposer la politique à différentes entités et relations entre elles, fait de la gestion de différentes parties des politiques aussi indépendantes que possible. Ainsi, l'hypothèse que le « purpose » est une entité séparée dans une telle modélisation de politiques va augmenter la complexité de la politique. Par exemple, la gestion des hiérarchies et des conflits doit couvrir cette nouvelle entité.

Les auteurs proposent un nouveau modèle sensible à la protection de la vie privée.

ils identifient quatre types de politiques afin de répondre aux besoins de leur modèle sensible à la protection de la vie privée :

- politiques de contrôle d'accès : comme dans le contrôle d'accès traditionnel, ils gouvernent l'accès aux données et services gérés par le groupe,
- politiques de révélation : gouvernent la révélation des propriétés ou des informations personnelles du groupe et spécifient les conditions sous lesquelles elles peuvent être révélées,
- politiques de traitement des données : spécifient comment les données personnelles sont utilisées par l'autre partie après révélation,
- politiques de filtrage : filtrent la réponse retournée à l'autre partie pour prévenir la révélation des informations sensibles liées à la politique elle-même.

Une règle de contrôle d'accès ou de révélation possède la forme suivante :

subject WITH subject-expression CAN action FOR purpose ON object WITH object-expression IF conditions FOLLOW obligations.

Où : subject, object, et action sont des entités auxquelles la règle réfère. Subject-expression et object-expression réfèrent à des conditions que le sujet et l'objet doivent satisfaire respectivement. Purpose, spécifie comment la donnée va être utilisée. Conditions, est une expression booléenne que le serveur doit suivre quand il gère les données personnelles. Il n'y a pas d'indications sur le type des conditions qu'on peut exprimer. La tâche d'administration n'a pas été évoquée dans ce modèle.

Fisher-Hubner et Ott, se sont intéressés à la définition d'un modèle de contrôle d'accès généralisé **[2]**. Ce modèle inclue deux autres principes : la nécessité de traitement et la liaison au « purpose ». Les auteurs spécifient que l'utilisateur peut accéder aux données/services, si cet accès est nécessaire pour accomplir la tâche en cours, et seulement s'il est autorisé à faire cette tâche. L'accès utilisateur est contrôlé en appliquant une procédure de transformation pour laquelle la tâche courante de l'utilisateur est autorisée. Ensuite, le but de la tâche utilisateur en cours doit correspondre aux buts pour lesquels les données personnelles ont été collectées.

Pour synthétiser, différents modèles de contrôle d'accès sont proposés, mais aucun ne permet de prendre en compte les axes règlementaires introduis dans le chapitre 1. Seulement, le modèle ABAC adopte le concept d'attributs ce qui permet d'exprimer toutes les caractéristiques des sujets, ressources, actions et environnement sous la forme d'attributs, et ainsi apporter des extensions en se basant sur ce concept. Ce même concept offre une flexibilité et une granularité très fine des politiques issues du modèle.

Nous nous basons principalement sur XPACML (mettre référence) issue du XACML, l'implémentation standardisée la plus utilisée du modèle ABAC. XACML fournit une architecture et des concepts définissant les grandes lignes de conception d'un système de contrôle d'accès. L'architecture de contrôle d'accès est extensible et modulaire, elle définit différents points sur lesquels se base la gestion d'accès qui ont été adaptés dans XPACML (référence).

4. EXTENSIBLE PRIVACY ACCESS CONTROL MARKUP LANGUAGE (XPACML):

4.1. Introduction :

XPACML est un outil proposé pour l'utilisateur, et considère les données personnelles comme des ressources à protéger. Comme déjà mentionné dans la section précédente,

XACML est un langage qui adopte le concept d' « attribut » du modèle ABAC, ce qui permet d'exprimer toutes les propriétés des éléments d'une règle d'accès à une donnée avec des attributs. XPACML profite de cette caractéristique fondamentale pour apporter les extensions nécessaires au modèle de base du langage XACML, afin que ce dernier puisse répondre aux besoins de présentation des politiques et de contrôle sur les données.

4.2. Principes du langage XPACML :

La **Figure 2.4** illustre le modèle du langage XPACML. Les extensions apportées par XPACML au modèle de base du langage XACML sont représentées par le biais de classes en bleu. Le langage XPACML est utilisé à la fois pour l'expression des politiques de traitement des données par SP, comme pour l'expression des préférences de l'utilisateur en termes de protection des données.

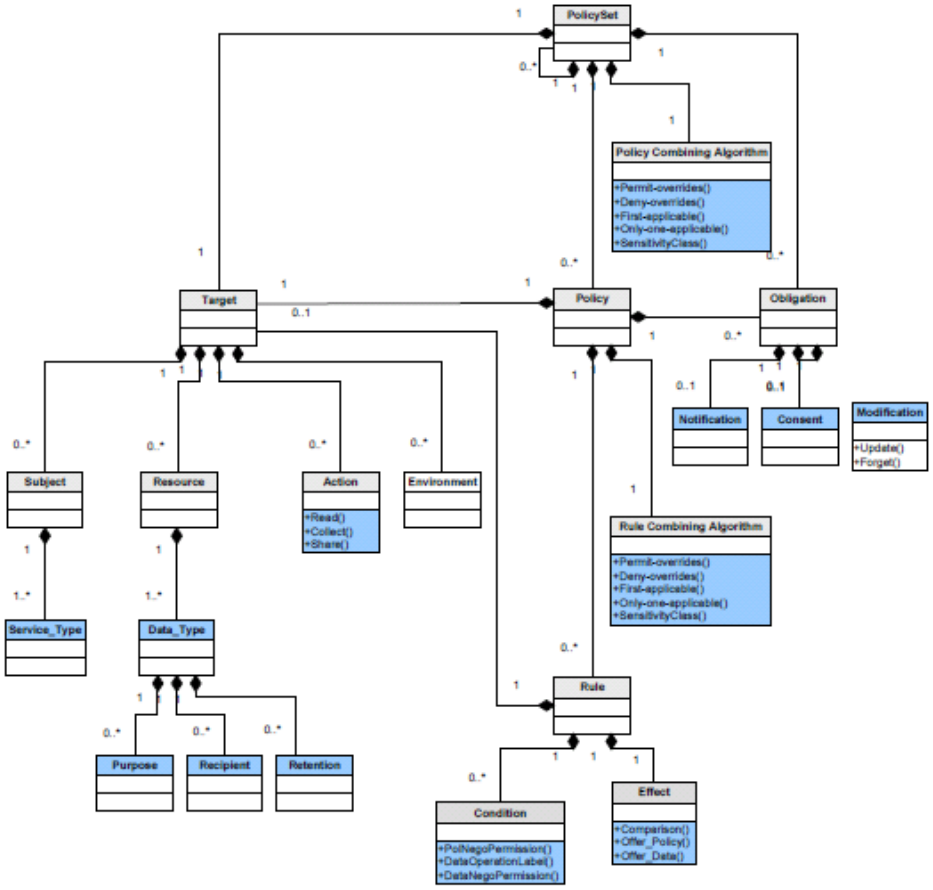

Figure 2. 4: Structure du langage XPACML (modèle de langage de politique).

4.2.1. Requête :

L'accès à une donnée s'effectue par le biais d'une requête (Request). Il doit être possible de formuler une requête avec les éléments suivants :

- L'identité des demandeurs qui seront appelés les sujets (Subjects). Une requête dans XPACML peut être initiée par plusieurs sujets ;
- La ressource à accéder (Resource) ;
- L'action à effectuer (Action).

Ces éléments peuvent posséder des propriétés.

4.2.2. Règles de contrôle d'accès :

Une règle de contrôle d'accès (rule) XPACML se base sur les éléments définis dans la requête XPACML représentant sa cible d'évaluation (target) pour en générer une décision (EFFECT) :

- L'identité des demandeurs qui seront appelés le (les) sujet (s) concerné(s) (SERVICE_TYPE),
- La (les) ressource (s) à accéder (DATA_TYPE) avec les déclarations d'usage autorisées (tags , , et de l'élément POLICY),
- Les actions autorisées pour exécution sur la ressource (ACTION),
- La décision à renvoyer (EFFECT), après évaluation des trois premiers paramètres.

En effet, la cible (target) avec les trois premiers éléments, représente une première étape pour savoir si une règle XPACML peut être appliquée à une requête donnée ou pas. Ainsi une décision d'une règle XPACML peut traditionnellement être soit positive (Permit), pour permettre l'accès à la donnée requise, soit négative (Deny), pour en refuser l'accès, en se basant sur les tags de l'élément POLICY fourni par le SP.

Une décision globale est générée à partir d'un ensemble de règles (rule). Les règles sont regroupées en politiques (Policy). Les politiques peuvent être regroupées en ensembles de politiques (policySet).

4.2.3. Politiques de contrôle d'accès (POLICY) :

Une politique XPACML regroupe plusieurs règles de contrôle d'accès concernant la même ressource (donnée). Ainsi, si nous avons des règles de contrôle de l'accès la donnée

AddressCity du type Address, et d'autres qui concernent la donnée Localisation, il sera plus judicieux de les séparer en deux ensembles distincts.

Ainsi, quand une demande de lecture de AddressCity parviendra à notre système de contrôle d'accès, seules les règles concernées seront utilisées. Une politique XPACML possède une cible (target) qui restreint son champ d'application à un ensemble limité de requêtes qui satisfont des conditions bien particulières.

Une fois une politique est appliquée à un contexte de requête, toutes les règles qui sont contenues dans la politique sont appliquées. Une sélection plus fine est alors obtenue. Les cibles des règles limitent le nombre de règles appliquées dans une politique.

4.2.4. Ensemble de politiques (PolicySet) :

Pour structurer les politiques et règles, XPACML reprend l'élément PolicySet du langage XACML **[2],** Aussi, et du fait que dans un même ensemble de politiques, plusieurs politiques peuvent s'appliquer et générer des réponses différentes, les algorithmes de combinaisons des règles XPACML sont encore utilisés par les ensembles de politiques, mais cette fois-ci nous parlons de combinaison de politiques (Policy Combining Algorithms). Enfin, nous pouvons associer les mêmes obligations définies pour les politiques XPACML, afin d'être exécutées en plus de la décision d'un PolicySet.

Enfin la **Figure 2.5** représente un exemple de politique XPACML avec les différents éléments discutés au suivant :

```
<? xml version="1.0" encoding="UTF-8" ?>
<xpacml:PolicySet xmlns:xpacml="urn:xpacml:policy"
xpacml:p3p="http://www.w3.org/2002/01/P3Pv1"
xpacml:xml="http://www.w3.org/XML/1998/namespace"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.w3.org/2001/XMLSchema-
instance">
<xpacml:Description>XPACML example SP's privacy
policy</xpacml:Description>
<xpacml:Policy PolicyId="eCommerce">
<xpacml:Target>
<xpacml:Subject>
<xpacml:Service_Type ApplyToDescendent="No">
<xpacml:eCommerce />
</xpacml:Service_Type>
</xpacml:Subject>
</xpacml:Target>
<xpacml:Rule Effect="Permit" Category_Id="CCNumber_Value"
ApplyToDescendant="Yes">
<xpacml:Description>One rule describes a policy for a specific
```

```
      category</xpacml:Description>
      <xpacml:Target>
      <xpacml:Resources>
      <xpacml:Resource                ResourceId="CCNumber_Value"
DataCompostion="Composed">
      <p3p:PURPOSE xmlns:p3p="http://www.w3.org/2002/01/P3Pv1">
      <p3p:current />
      <p3p:admin />
      <p3p:develop />
      <p3p:historical />
      </p3p:PURPOSE>
      <p3p:RECIPIENT
xmlns:p3p="http://www.w3.org/2002/01/P3Pv1">
      <p3p:ours />
      </p3p:RECIPIENT>
      <p3p:RETENTION
xmlns:p3p="http://www.w3.org/2002/01/P3Pv1">
      <p3p:no-retention />
      </p3p:RETENTION>
      </xpacml:Resource>
      </xpacml:Resources>
      <xpacml:Actions>
      <xpacml:read />
      <xpacml:collect />
      <xpacml:share />
      </xpacml:Actions>
      </xpacml:Target>
      </xpacml:Rule>
      </xpacml:Policy>
   </xpacml:PolicySet>
```

Figure 2. 5: Exemple d'une politique XPACML.

4.3. Architecture de contrôle d'accès XPACML :

L'architecture XPACML permet d'interfacer les échanges entre les utilisateurs, les SPs et d'autres entités intermédiaires comme les autorités législatives. Elle suit l'idée d'un « courtier de vie privée ». Une hypothèse basique de son approche architecturale est que chaque SP doit s'interfacer avec l'architecture qui constitue un proxy entre lui et l'utilisateur. La gestion des données personnelles au sein de l'architecture est gouvernée par un ensemble de règles (de politiques), qui reflètent les besoins législatifs. Ces règles permettant l'expression des données, les éléments reflétant leur protection (règlementaire), et les types de services ayant des interactions avec le système. Les politiques XPACML issues sont exécutées au sein de l'architecture. XPACML utilise le même modèle de langage de politique XACML pour l'expression des préférences utilisateur en termes de protection des données, comme pour l'expression des politiques du SP. L'idée principale derrière, est que chaque donnée personnelle injectée dans le système est sujette à des politiques qui spécifient quelle partie de la donnée doit être exposée, à quels types de services et sous quelles contraintes d'usage.

L'architecture XAPCML dispose nativement d'une couche de données, dans laquelle sont stockées toutes les données personnelles liées à l'utilisateur et ses préférences ainsi qu'aux normes législatives, une couche de traitements qui constituent les procédures d'actions spécifiques du système utilisateur, et d'une couche interface responsable d'intercepter les actions et préférences de l'utilisateur. Un schéma de la structure interne de l'architecture avec une vue composants, est fourni dans la **Figure 2.6**:

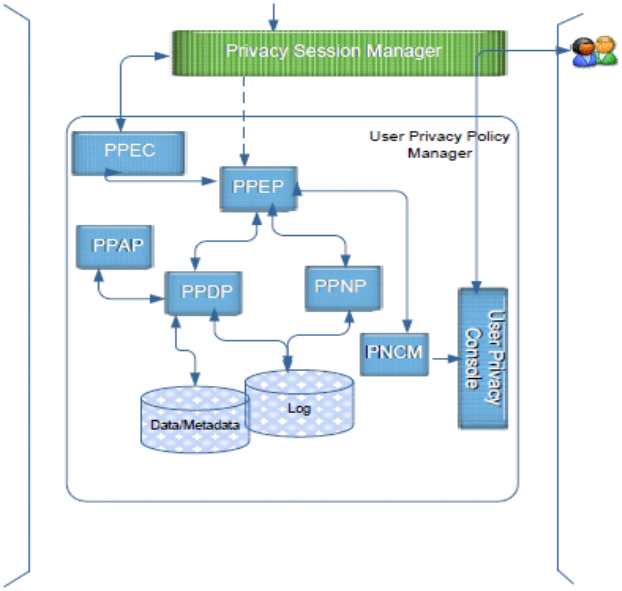

Figure 2. 6: la structure interne de l'architecture XPACML.

ARCHITECTURE:

L'architecture est bâtie autour des éléments suivants :

Un système de politiques : il comprend les composants PPAP/PPDP/PPNP/PPEP dont le rôle est expliqué ci-dessous.

- ***PPEC (Pricay Policy Enveloppe Constructor):*** il lie les données personnelles de l'utilisateur avec les méta-données de vie privée qui leur sont liées, dans une enveloppe chiffrée et signée qui sera transmise au SP.

  Un gestionnaire de notification et de consentement (PNCM) : il gère toutes les tâches liées à la notification et au consentement de l'utilisateur.

- ***Le gestionnaire de session de vie privée PSM (Privacy Session Manager)***

  Ce composant représente l'interface vie privée du système utilisateur avec les agents externes. Il est conçu pour interagir avec l'agent du SP pour récupérer sa politique d'usage des données en XPACML. Dès que cette dernière est obtenue, elle

est transférée au composant PPEP puis PPDP pour comparaison avec les préférences de l'utilisateur.

- ***Le gestionnaire de politiques de protection des données UPPM (User Privacy Policy Manager)*** UPPM est le composant qui gère les interactions de l'utilisateur avec les composants internes de l'architecture. Ceci inclut des interactions comme l'ajustement du niveau de protection des données souhaité, l'édition des préférences de l'utilisateur pour chaque type de données, et l'information de l'utilisateur des politiques de protection activés.
- ***Le gestionnaire de notification et de consentement PNCM (Privacy Notification and Consent Manager)***

  Ce composant prend en charge toutes les tâches liées à la notification ou le consentement de l'utilisateur. Il prend en charge la fonction de consentement en utilisant les fenêtres de consentement. Les alertes de violation des termes légaux comme les préférences utilisateur peuvent être également fournies via ce composant. Il est également utilisé afin de demander l'avis d'acceptation ou pas de l'utilisateur suite à un conflit non résolu par le processus de négociation
- ***Le point d'administration des politiques PPAP (Privacy Policy Administration Point)***

  Ce composant présenté définit les règles de contrôle d'accès et les politiques XPACML.

- ***Le point d'application des politiques PPEP (Privacy Policy Enforcement Point)***

  Ce composant reçoit toutes les requêtes d'accès aux données personnelles et les décompose en plusieurs requêtes XPACML, chacune est dédiée à une donnée particulière. Chaque requête élémentaire est envoyée par la suite au composant PPDP pour obtenir une autorisation de délivrance de la donnée requise. En effet, le PPEP établit une requête d'autorisation, spécifiant l'identifiant du SP, le type de service faisant la demande, le type de la donnée requise et d'autres informations requises par le PPDP (ex : conditions) afin d'établir une décision. Une fois les décisions élémentaires reçues, le PPEP les regroupe pour en faire une globale.

- ***Le point de décision des politiques PPDP (Privacy Policy Decision Point)***

Le PPDP est le composant central de l'architecture, chargé de sélectionner les règles/politiques et ensembles de politiques applicables à une requête donnée. Le PPDP prend en compte plusieurs paramètres afin d'établir une décision.

- ***Le point de négociation des politiques PPNP (Privacy Policy Negotiation Point)***

Ce composant effectue la négociation des termes de politique du SP avec les termes des préférences de l'utilisateur. Un terme d'une règle ou d'une politique de contrôle d'accès est composé de la paire (terme, politique). En effet, pour toute requête refusée dont le conflit porte sur un ou plusieurs terme(s) de la politique de traitement de données, le PPDP envoie le terme en question au PPNP à travers le PPEP afin de trouver un arrangement avec les préférences éditées par l'utilisateur. Le composant PPNP entame à travers le PPEP un processus de négociation de politiques avec le SP jusqu'à ce que la négociation aboutisse ou un signal de terminaison arrive d'une des deux parties. La réponse finale de la négociation est ainsi transférée au composant PPEP. Ce dernier applique la décision globale en autorisant ou pas l'accès aux données personnelles.

- ***Le composant enveloppe de vie privée PPEC (Privacy Policy Enveloppe Constructor***

Ce composant assemble les données à révéler avec les politiques de protection associées, dans une enveloppe chiffrée et signée qui sera transmise dans sa totalité au SP.

- ***Les bases de données Base de données utilisateur***

Le répertoire des données personnelles est l'endroit de stockage pour les données de l'utilisateur qui sont à fournir pour un but spécifique.

5. CONCLUSION :

Les langages de politiques comme P3P proposent des solutions aux problèmes de formulation des pratiques d'usage faites des données personnelles, en offrant une boite à outils pour l'expression des éléments de politiques de protection des données. Néanmoins, ces langages présentent des limitations en termes de prise en compte des réglementations. En essayant de chercher des mécanismes mettant en œuvre des politique de contrôle d'accès aux données (ressources de l'utilisateur), nous avons établi un état de l'art des modèles de contrôle d'accès proposant des éléments de protection des données et nous avons présenté leurs limitations. On s'est intéressé plus particulièrement à ABAC et son implémentation standardisée XACML. Ce dernier adoptant le concept d'attributs pour la spécification des règles d'accès, permet de définir des politiques d'accès fines et mieux adaptées aux cas de figures réels. Ensuite, nous sommes intéressés aux outils techniques intégrant les concepts de protection des données couvrant un périmètre plus large que P3P, APPEL et XPref à ceux des langages de contrôle d'accès [29]. De ce fait, nous avons présenté XPACML (sur lequel nous nous basons dans le présent travail), définissant des extensions au langage XACML pour couvrir les aspects de protection des données personnelles. XPACML répond à la problématique de présentation des axes règlementaires, et celle de contrôle d'usage associé aux données.

La présentation des aspects réglementaires à elle seule, et la comparaison de la politique du SP avec les préférences de l'utilisateur ne sont pas suffisantes. En effet, XPACML n'est applicable que si la (les) donnée(s) demandée (s) par le SP existe (s) avec la politique accompagnante. Mais que se passe t–il en cas où le SP demande un niveau de granularité de données non présenté dans les BDDs de XPACML ?

contextes d'usage des données. Nous présentons dans le chapitre suivant un modèle d'information sémantique dédié à cet effet.

CHAPITRE III

1. INTRODUCTION:

Nous avons montré tout au long du chapitre précédent les motifs qui nous conduit à choisir la solution XPACML, comme langage d'expression des politiques d'usage appliquées sur les données personnelles de l'utilisateur, afin de contrôler l'accès du SP à ces dernières.

Seulement le langage XPACML est un langage statique, où, tous les éléments d'une règle d'accès doivent être explicités afin que cette dernière puisse être appliquée. Ainsi il n'est pas applicable dans le cas où le SP demande un niveau de granularité de données non présenté dans les BDDs de XPACML ?
La question est comment manipuler, filtrer les données personnelles d'une manière intelligente de telle façon à pouvoir trouver des règles applicables pour des requêtes de données non explicitées dans le modèle du langage XPACML.

Pour ce faire, nous nous somme dirigé vers les outils de manipulation des informations sur le Web, et son évolution avec le temps.
Nombre de travaux démontrent que le Web actuel comprend le contenu des pages et peut relier les données entre elles d'une page à une autre. Ceci est facilité avec le *web sémantique*, qui est à la base d'interprétation des données circulant sur le Web.

Dans ce chapitre on va présenter un modèle d'information sémantique conçu pour apporter des solutions aux limitations du modèle du langage XPACML, en adoptant les concepts du Web sémantique (plus particulièrement les ontologies).

2. LES QUATRES PRINCIPAUX STANDARDS DU WEB SEMANTIQUE:

Le Web sémantique est une extension du Web standardisée par le World Wide Web Consortium (W3C). C'est un ensemble de connaissances basées sur XML et RDF(S) qui permettent et facilitent la recherche d'information en se basant sur les concepts et les liens entre eux.
L'architecture du web sémantique s'appuie sur une pyramide de langages pour représenter des connaissances sur le web en satisfaisant les critères de standardisation, d'interopérabilité et de flexibilité. Cette architecture en couches (**Figure 3.1**) permet une approche progressive dans les processus de standardisation et d'acceptation par les utilisateurs. En ce qui suit les éléments principaux de sa structure:

- **RDF:** un modèle de triplets pour décrire et connecter des ressources anonymes ou identifiées par un URI. Ces ressources sont dans notre cas les données personnelles de l'utilisateur et les politiques associées (préférences de l'utilisateur). **SPARQL**:(Protocol And RDF Query Language) : un langage de requêtes sur les graphes RDF, c'est un outil de manipulations des annotations et ontologies. Il nous permet dans notre cas d'interroger l'ontologie et obtenir les préférences de l'utilisateur appliquées sur les données requises par le SP.
- **RDFS:** est un langage de descriptions légères.
- **OWL:** Le langage OWL *(Web Ontology Language)* a été conçu pour être utilisé par les applications qui traitent le contenu de l'information au lieu de la présenter seulement aux êtres humains. Dans notre cas on s'est basé dessus pour décrire les différentes classes de données de l'utilisateur en concepts, ainsi que pour l'expression des éléments de politiques applicables sur les données en tant que préférences de l'utilisateur.

a

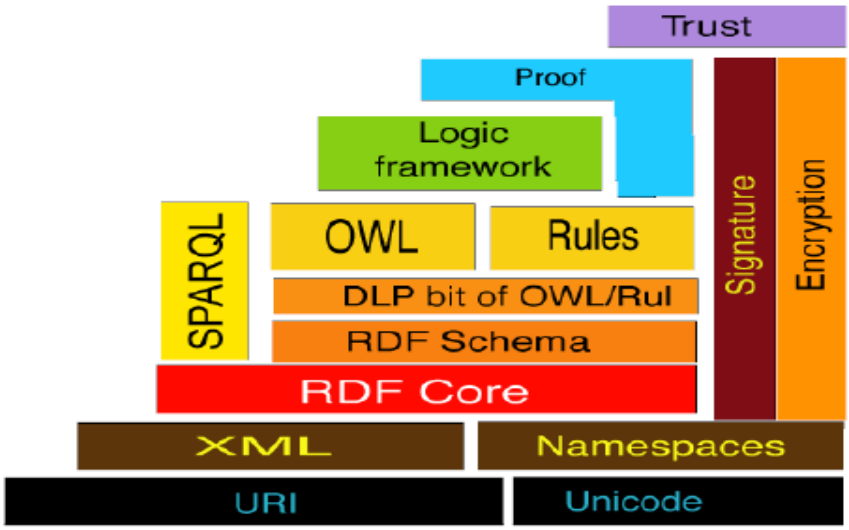

Figure 3. 1: structuration de web sémantique.

3. QU'EST CE QU'UNE ONTOLOGIE:

L'un des concepts de base du web sémantique ce sont les ontologies. Les ontologies sont apparues au début des années 90 dans la communauté Ingénierie des connaissances, dans le cadre des démarches d'acquisition des connaissances pour les systèmes à base de connaissances (SBC). **[30].**

Définition 1: Ensemble des objets reconnus comme existant dans le domaine. Construire une ontologie c'est aussi décider de la manière d'être et d'exister des objets **[30].**

Définition 2: « En informatique, une ontologie est un ensemble structuré de concepts. Les concepts sont organisés dans un graphe dont les relations peuvent être :

- ✓ des relations sémantiques.
- ✓ des relations de composition et d'héritage (au sens objet).» **[31].**

4. POUR QUELLES RAISONS DEVELOPPER UNE ONTOLOGIE ?

Les quatres standards du Web sémantique sont conçus à l'origine pour les applications Web basés sur le XML (eXtensible Markup Langage) et le RDF (Resource Definition Framework), ils permettent de déduire des connaissances en utilisant le principe de recherche par concept sur des applications reliées entre elles par des liens sémantiques

Nous pensons que ces outils sont bien adaptés pour notre contexte de travail, pour les raisons suivantes :

- Une ontologie exprimée avec un langage du Web sémantique fournit des moyens

pour être développée indépendamment du système auquel elle appartient. Le partage des informations est possible, permettant ainsi une minimisation du coût de la redondance.

- RDF et OWL sont des langages de représentation des connaissances, avec une forte expressivité, permettant la modélisation de différents types d'informations (exemple: les informations liées à un utilisateur, des évènements,…etc).
- L'ontologie formalisant les éléments de protection fournit une représentation explicite de leurs sémantiques, qui peuvent être raisonnées par les moteurs d'inférence logiques existants.
- En appliquant un moteur d'inférence du Web sémantique, il nous est possible d'utiliser des règles logiques spécifiques pour déduire depuis les éléments constituant le modèle de base, des données liées à un contexte donné.
- En adoptant cette approche, il ainsi est possible de personnaliser ou généraliser les inférences faites sur le modèle.
- Les langages du Web sémantique peuvent être utilisés comme des méta-langages pour définir d'autres langages plus spécifiques comme les langages pour la communication et le partage de connaissances sur la protection de la vie privée.
- Partager la compréhension commune de la structure de l'information entre les SPs et les utilisateurs.
- Analyse du savoir sur un domaine et permettre son réutilisation.

- Expliciter ce qui est considéré comme implicite sur un domaine.

5. NOTRE TRAVAIL:

Dans cette partie nous allons présenter le travail d'implémentation que nous avons accompli, et qui consistait premièrement à l'édition de l'ontologie sous le langage OWL, suivi par son exploitation dans une application qui a été développée sous Eclipse.

5.1. Choix des éditeurs et de langage de programmation :

5.1.1. Choix de l'éditeur de l'ontologie :

Parmi les éditeurs des ontologies existant nous avons choisis d'utiliser *protégé. Protégé* est un éditeur d'ontologies. Il a été créé à l'université de Stanford.

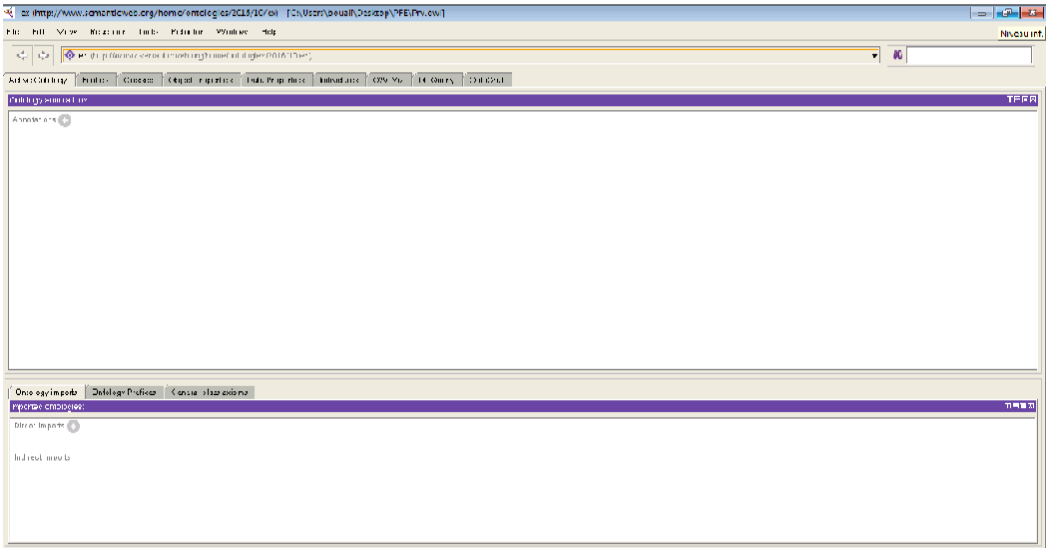

Figure 3. 2: Page d'accueil de protégé.

5.1.2. Choix de langage de programmation Java et l'éditeur eclipse:

Java:

Java est un langage de programmation orienté objet, développé par Sun Microsystems. Il fut présenté officiellement en 1995. Selon les développeurs de Sun, Java est un langage : simple, orienté-objet, distribué, interprété, robuste, sécurisé, neutre vis à vis de l'architecture, portable, multi-plateforme, à haute performance, multi-threaded et dynamique.

Netbeans:

Figure 3. 3: Page d'accueil de netbeans.

NetBeans est un environnement de développement intégré (EDI), placé en open source par Sun en juin 2000 sous licence CDDL (Common Development and Distribution License) et GPLv2. ... NetBeans constitue par ailleurs une plate forme qui permet le développement d'applications spécifiques. NetBeans permet la prise en charge native de divers langages tels le C, le C++, le JavaScript, le XML, le Groovy, le PHP et le HTML. **[32].**

5.2. Création de l'ontologie:

Les concepts de base constituant notre ontologie doivent refléter les données à protéger de l'utilisateur, l'ensemble des SPs qui sont susceptibles d'interagir avec l'utilisateur et les éléments de politique que l'utilisateur pourra utiliser afin de spécifier ses préférences personnelles. .

Dans ce qui suit on va présenter quelques interfaces de l'ontologie crée;

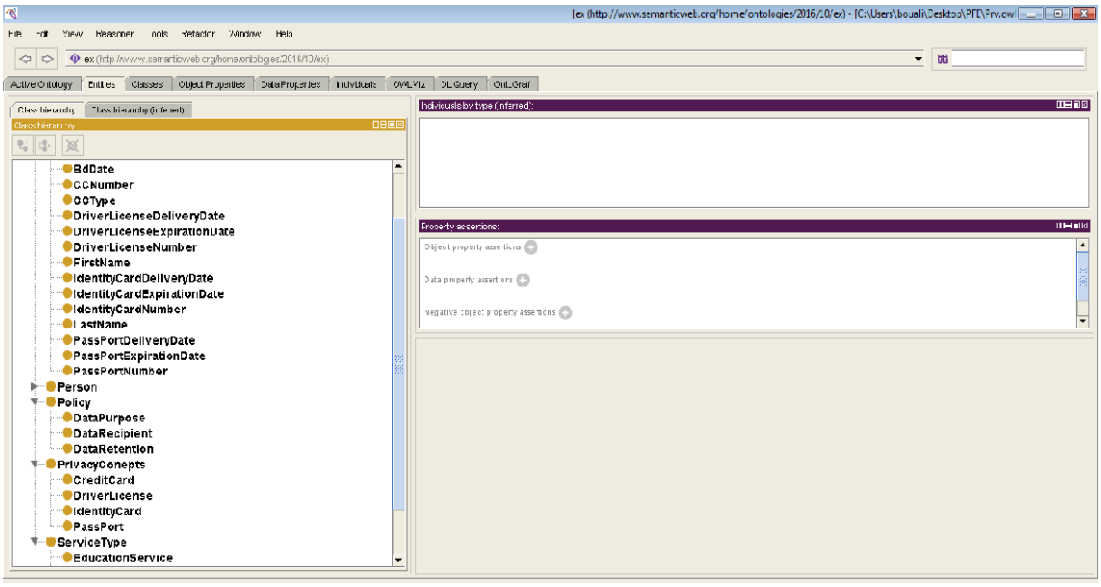

Figure 3.4: Représentation des concepts DataType et Policy de l'ontologie.

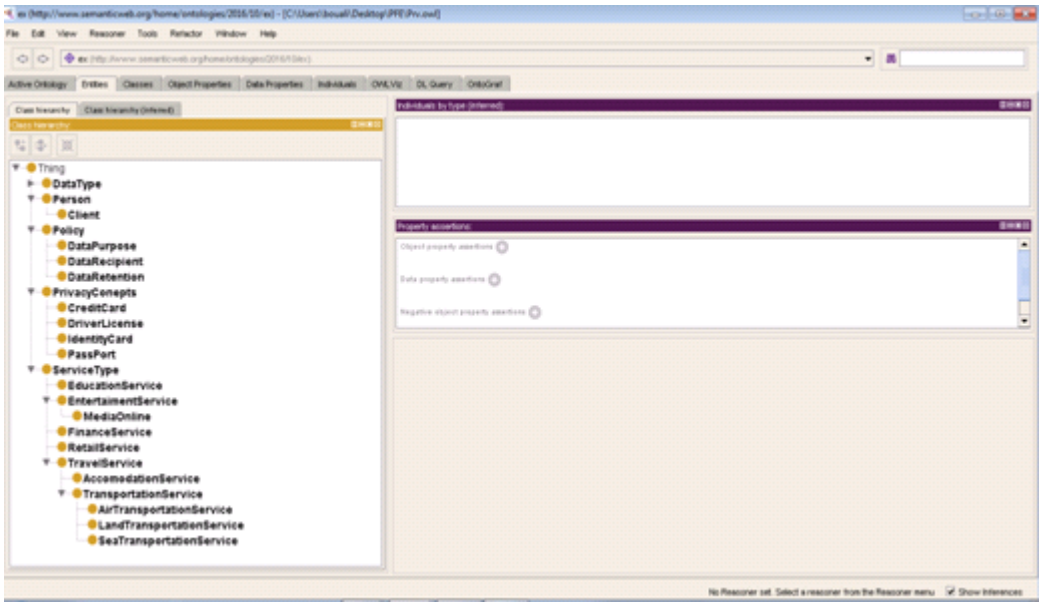

Figure 3.5 : Représentation des concepts ServiceType.

Nous avons lié les concepts de l'ontologie avec des liens sémantiques comme illustré dans la **figure 3.6**.

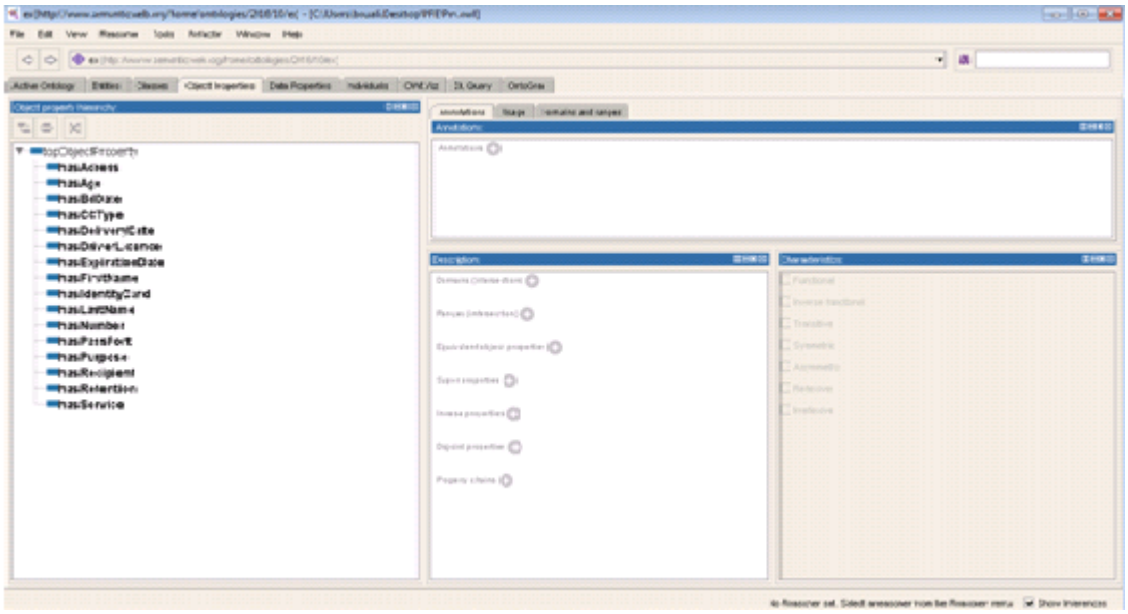

Figure 3. 6 : Représentation des attributs de l'ontologie.

Une représentation générale des concepts liés entre eux est illustrée dans **la figure 3.7**.

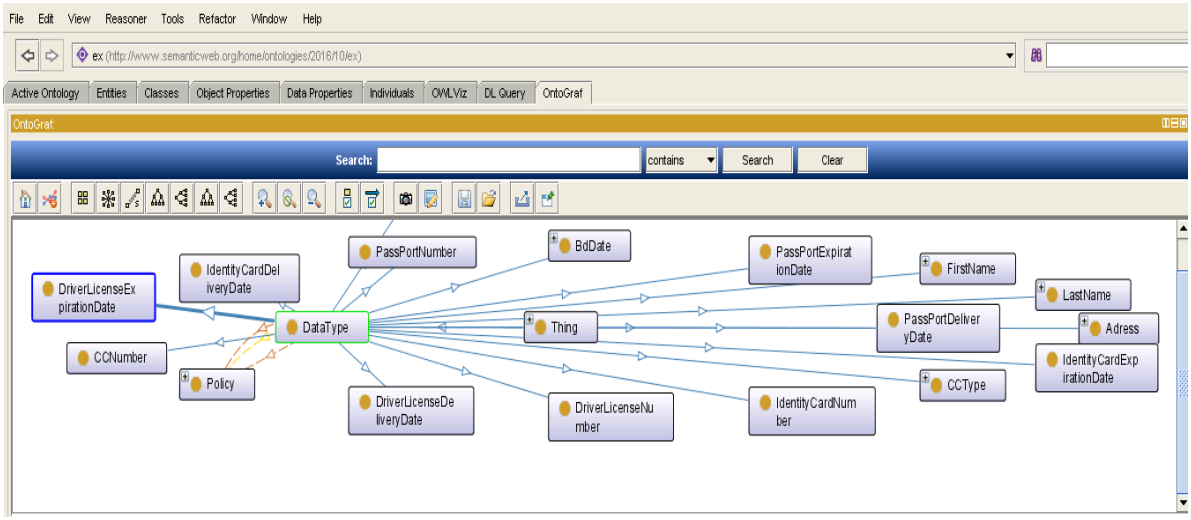

Figure 3. 7 : Représentation des liens sémantique entre les concepts.

A titre d'exemple la figure suivante représente un lien d'héritage entre service type et éducation service.

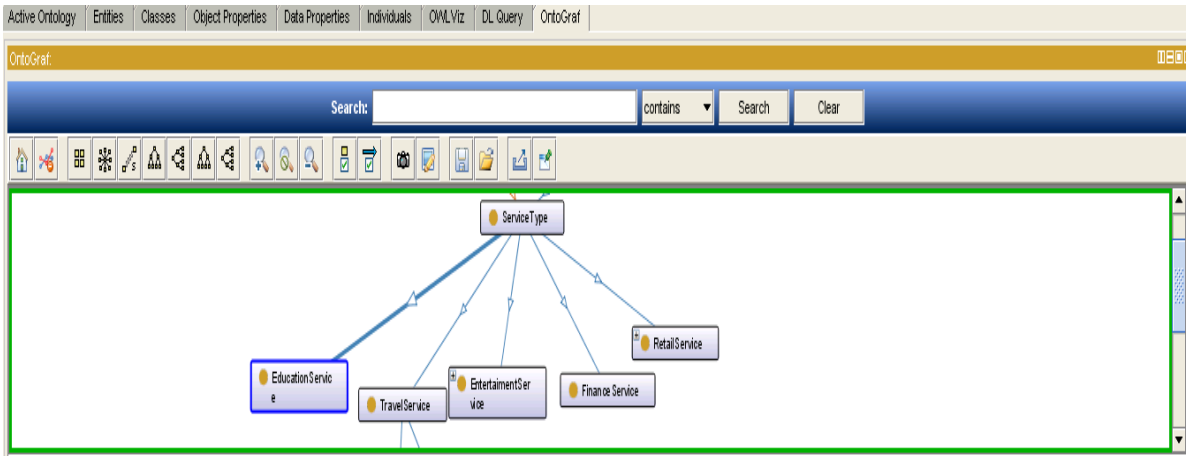

Figure 3. 8 : Exemple d'un lien d'héritage entre service type et education service.

5.3. Conception de système :

Dans ce qui suit nous allons détailler la conception de notre système via un scénario E-Commerce (Commerce Électronique), qui illustre sur un exemple concret des concepts abstraits que nous avons mis dans le modèle sémantique déjà présenté (**figure 3.7**).

Figure 3. 9 : Première page d'accueil de l'utilisateur.

Après le lancement de l'application, l'utilisateur choisit le fournisseur de service tout en protégeant ses données.

Pour des raisons de mise en œuvre du scénario, nous avons mis dans le même système le SP également. Ceci afin d'illustrer les différentes étapes de spécifications des éléments des

politiques (préférences de l'utilisateur), et de comparaison des données, où le SP est un élément principal de la requête.

Supposons que le fournisseur de service en ligne offre plusieurs sous services. Parmi les services offerts il y'a le service de location de voiture en ligne.

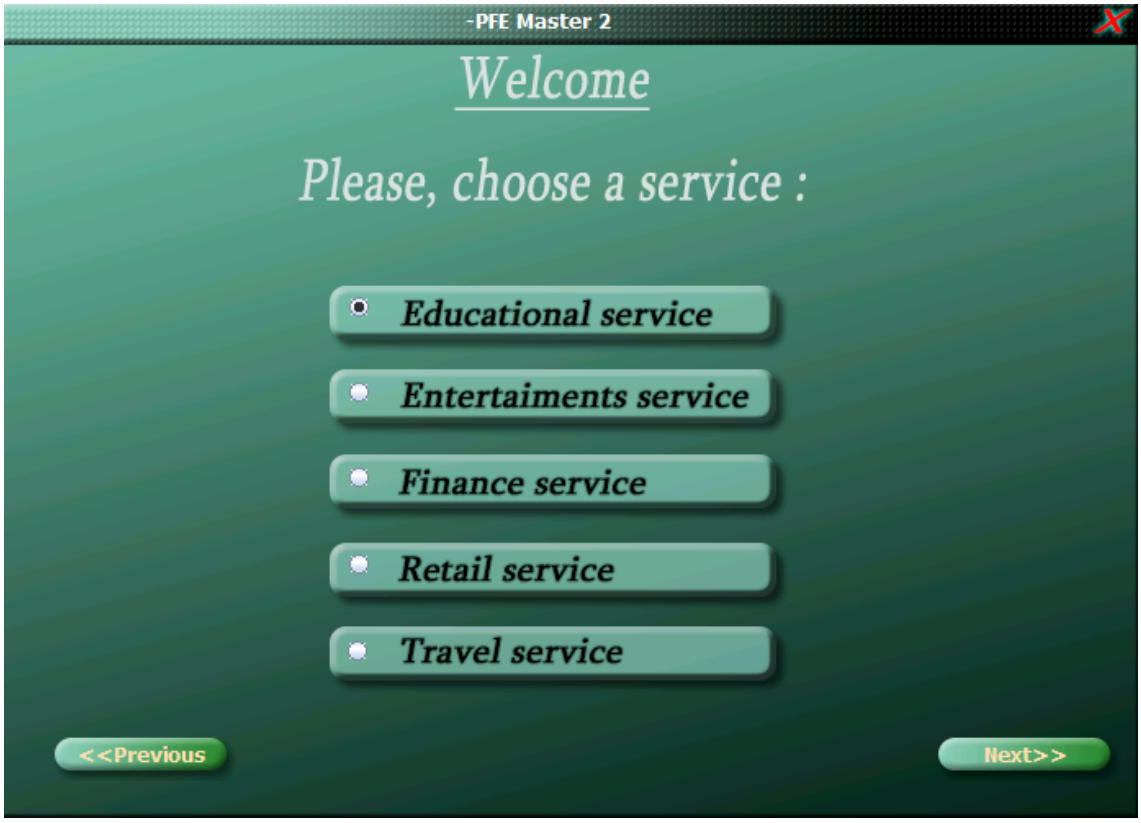

Figure 3. 10 : Représentation des services offerts

Le scénario se déroule comme suit : au début l'utilisateur choisi un service en ligne parmi les services existants. Nous supposons qu'il a choisi le service de transport. Puis il choisit le mode de transport (pour notre exemple nous choisissons le transport terrestre avec ses différents sous services).

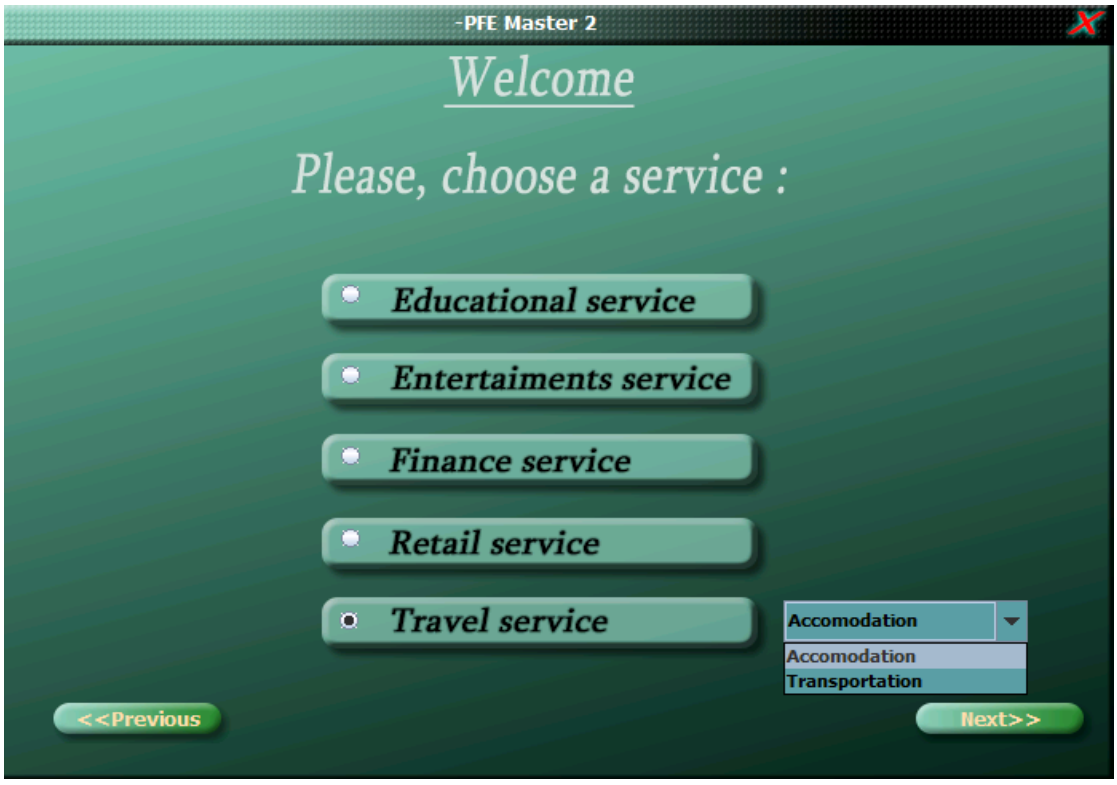

Figure 3. 11 : Représentation des sous services (exemple : travel service).

Ensuite l'utilisateur (client) choisi le service location de voiture comme l'illustre la figure suivante.

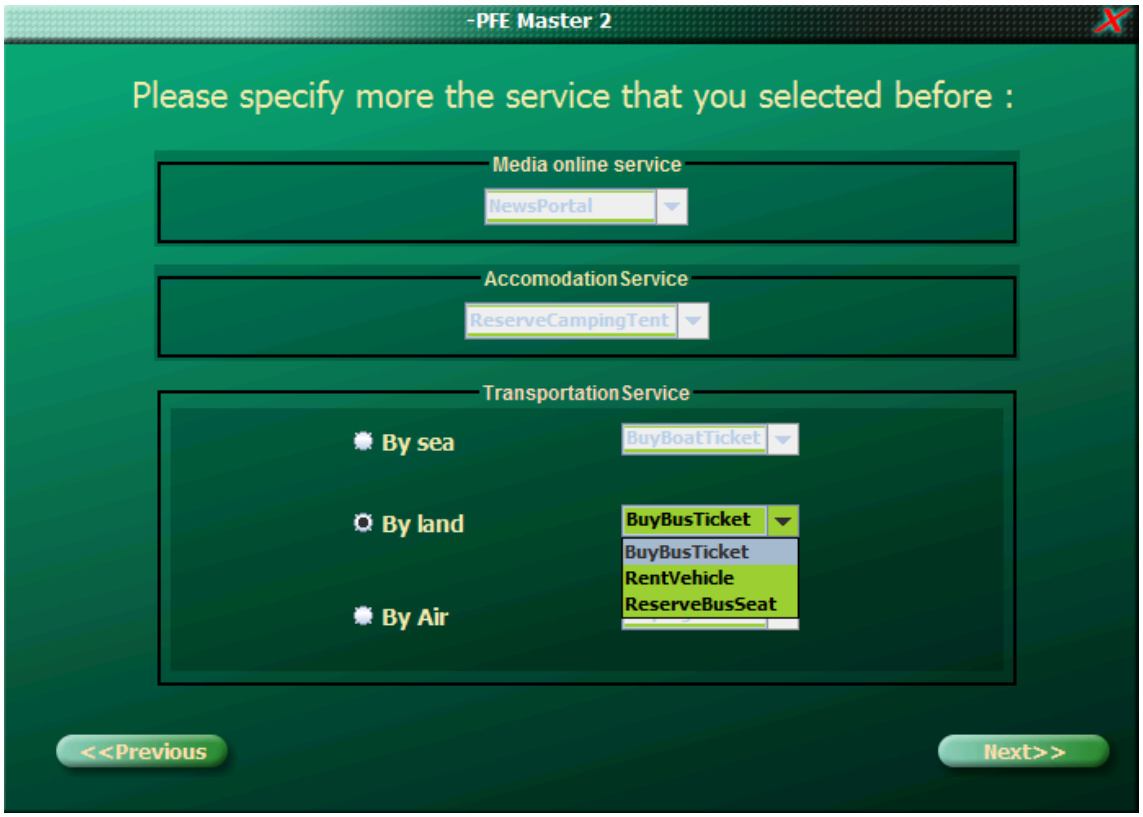

Figure 3. 12 : Choix du mode de transport.

Après l'utilisateur spécifie les politiques (ses préférences personnelles) qui seront appliquées sur ses données personnelles (données requises pour ce service). Le SP applique également des politiques sur ces données personnelles de l'utilisateur, dont il les fournit avec le contrat du service. Ces politiques qui lui sont spécifiques, afin de pouvoir les comparer avec les préférences de l'utilisateur.

Si toutes les données requises sont présentes; et les politiques appliquées dessus par le SP coïncident avec les préférences spécifiées par l'utilisateur, l'utilisateur pourra accéder au sous service location de voiture.

Dans ce qui suit on va montrer les interfaces de l'application implémentant la suite de notre scénario.

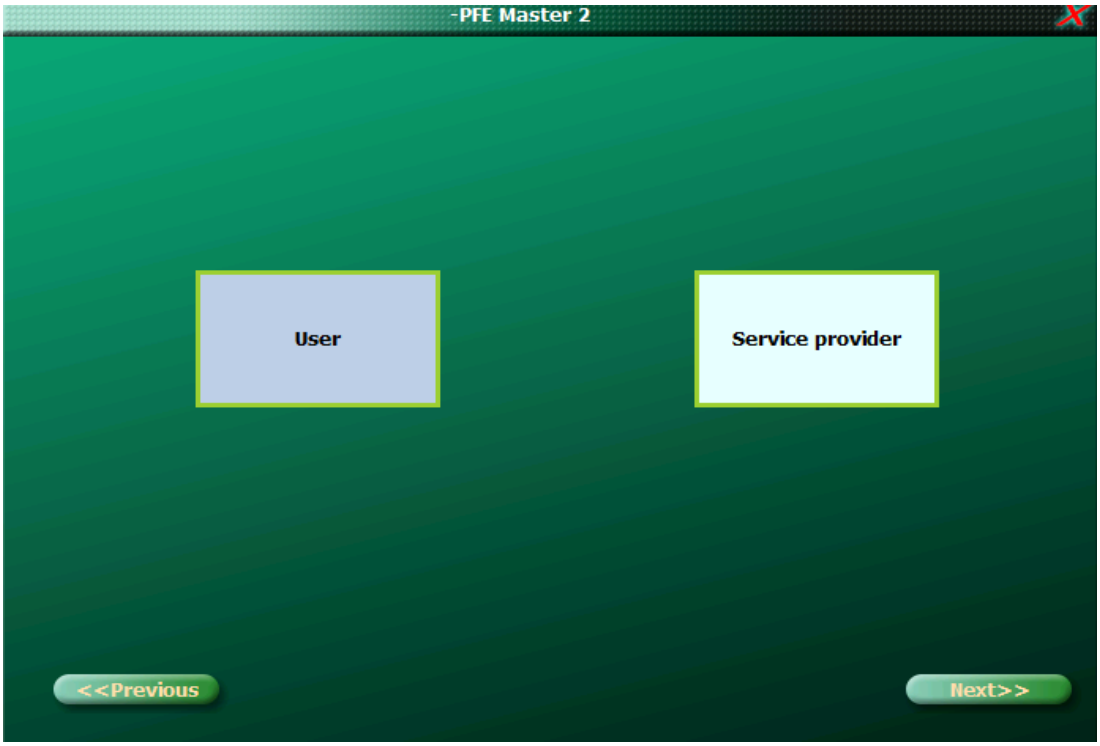

Figure 3. 13 : Choix du mode utilisateur.

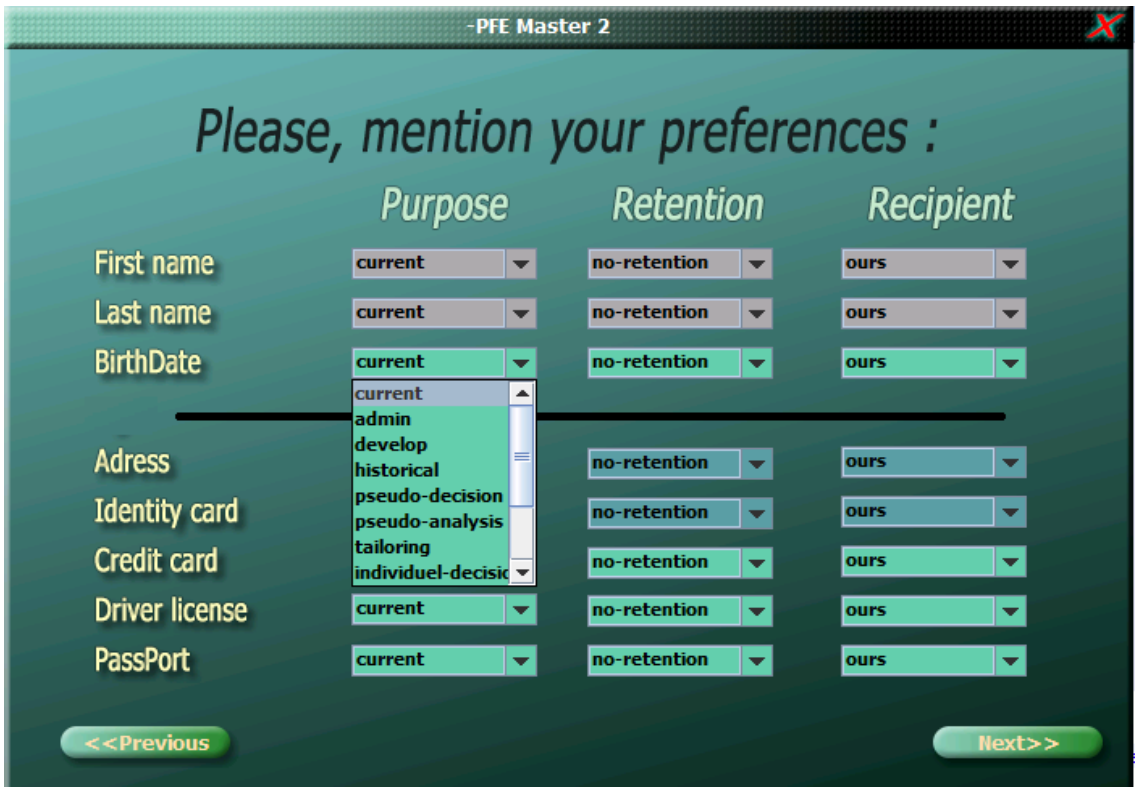

Figure 3. 14 : Spécification des préférences de l'utilisateur avec ses trois éléments de politiques.

Ensuite l'utilisateur est invité à rentrer les valeurs de ses données personnelles dont il vient de spécifier les préférences.

Figure 3. 15 : Représentation des données personnelles de l'utilisateur.

Une fois les valeurs des données renseignées par l'utilisateur, ces dernières sont chargées dans l'ontologie comme le montre la figure 3.15:

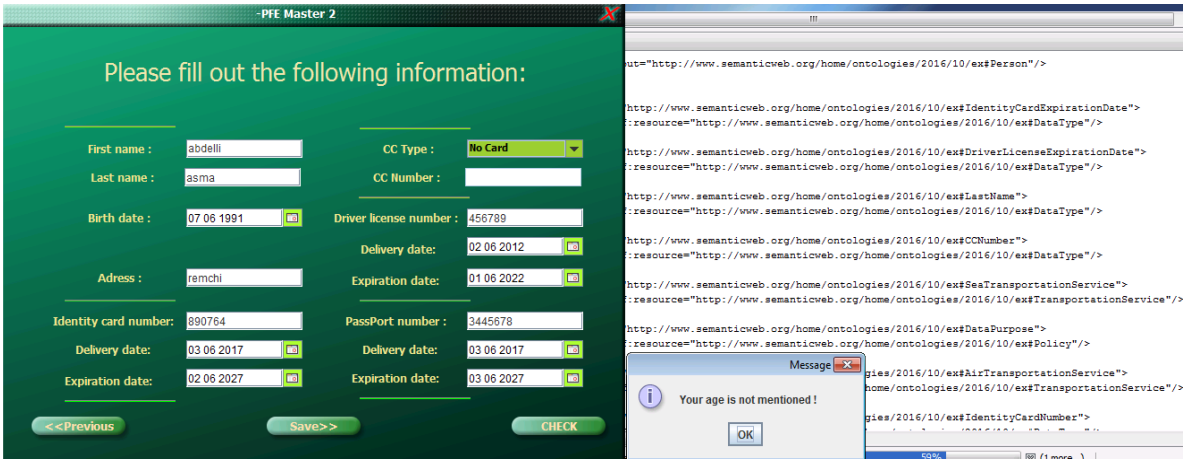

Figure 3. 16 : Chargement des valeurs des données personnelles dans l'ontologie.

Seulement le SP exige deux contraintes supplémentaires sur quelques données pour que l'utilisateur ait accès au service (figure 3.17 et figure 3.18):

- Première contrainte: la duré de permet de conduire doit être plus de 12 mois.
- Deuxième contrainte: l'utilisateur doit avoir l'âge entre 25 et 75.

Le modèle sémantique que nous avons définit pour le language XPACML, ne contient pas ces deux données explicitement dans le modèle de données déjà présenté dans la figure 3.4. Ainsi que les politiques les accompagnants.

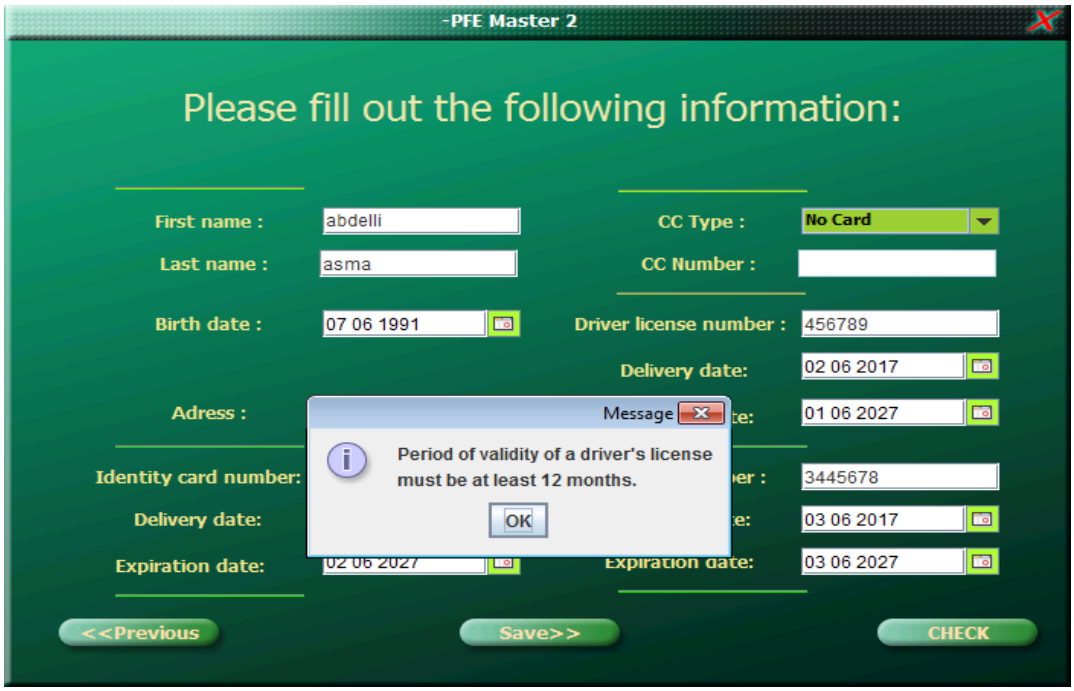

Figure 3. 17 : Représentation de contrainte de permis de conduire.

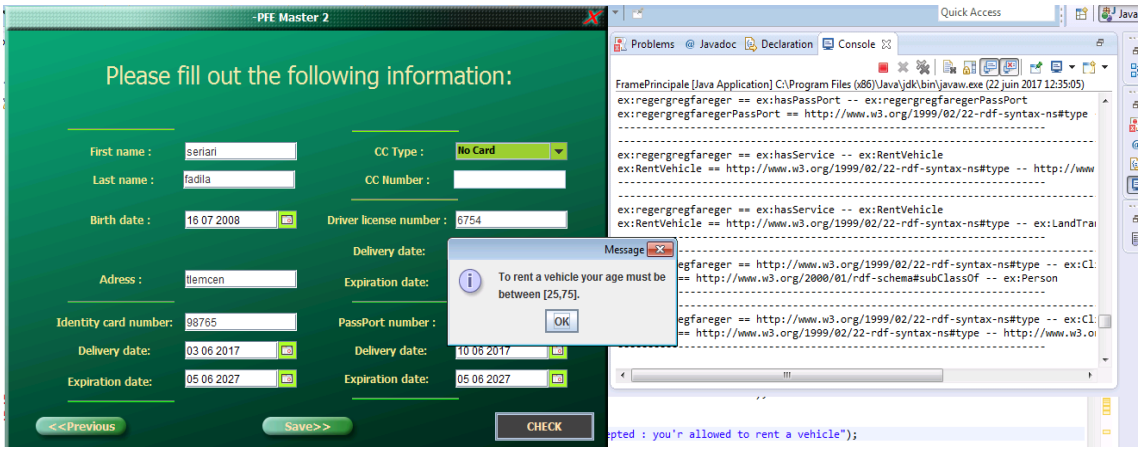

Figure 3. 18 : Représentation des contraintes âge.

Afin de donner suite à la transaction, et pouvoir comparer les politiques du SP avec les préférences de l'utilisateur, le modèle sémantique de données (figure 3.7) doit être exploité avec un moteur d'inférence afin de déduire et vérifier les données exigées par le SP et les préférences qui leur sont associées également.

Notre solution réside dans la déduction des valeurs de données implicites demandées par le SP, depuis l'ontologie. Ceci est possible grâce à notre moteur qui utilise le langage SPARQL.

Ce dernier est un langage de requêtes pour l'interrogation de méta données RDF de l'ontologie.

Sachant que SPARQL est un langage de requête et que JENA est une API java qui permet d'interroger un document RDF grâce a ses classes (QueryFactory, QueryExecutionFactory) on peut créer une requête recuperer le resultat et le traiter a notre guise.

La Class QueryFactory est une classe qui implémente le design pattern factory qui nous offre une requête (Query), cette classe nous permet d'écrire des requête SPARQL qui sont Formatable (vérifiable par un moteur de syntaxe SPARQL intégré)

Une fois notre requête ecrite on a besoin de l'executer pour cela on se sert de la QueryExecutionFactory, cette classe nous offre un objet qui va nous permettre de communiquer la requête avec notre base de données OWL, enfin cette objet va nous rendre le resultat sous forme d'un resultSet contenant nos information nécessaire au traitement de la partie métier de notre application

Exemple:

```
String query = "PREFIX  ex:   <http://www.semanticweb.org/home/ontologies/2016/10/ex#>"
+"PREFIX ns: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>"
+ "SELECT  * "
+ "WHERE"
+ " { ?s  a  ex:Client ."    // chercher que des clients
+ " ?s ex:hasFirstName ?firstName " // ce client X a une propreité  qu on stoqué dans la variable firstName
+ " OPTIONAL" // une jointure sauf si firstName a la propreité hasValue
+ "    { ?firstName  ex:hasValue  ?o } ."  // on stoque le FirstName Dans la variable 'o'
+ " ?s ex:hasLastName ?lastName " // client X a une propreité hasLastName qu on stoqu dans la variable firstName
+ " OPTIONAL"
+ "    { ?lastName  ex:hasValue  ?o1 } ."
+ " ?s ex:hasBdDate ?bdProperty "
+ " OPTIONAL"
+ "    { ?bdProperty  ex:hasValue  ?bd } ."
+ " ?s ex:hasDriverLicense ?drvLProperty "
+ " OPTIONAL"
        + "    { ?drvLProperty  ex:hasDeliveryDate  ?PdelevryDate } ."
        + "     OPTIONAL"
        + "    { ?PdelevryDate  ex:hasValue  ?delevryDate } ."
        + "}";
    Query qry = QueryFactory.create(query);  // créer la requête
    readOnthologie("src/Prv.owl"); // spécifie l'ontologie de lecture ( bdd)
    QueryExecution exe = QueryExecutionFactory.create(qry, ontModel); // créer l objet qui va exécuter la requête
     ResultSet rs = exe.execSelect(); // exécuter requête et récupérer le résultat dans un ResultSet
     System.out.println("begin");
```

```
while (rs.hasNext()) { // tanque on a un resultat
QuerySolution querySolution = (QuerySolution) rs.next(); // obtenir le résultat courant
```

Comme le montre la figure suivante le moteur d'inférence déduit l'âge de l'utilisateur depuis sa date de naissance renseignée dans l'ontologie et donne suite à la comparaison de politique par la suite.

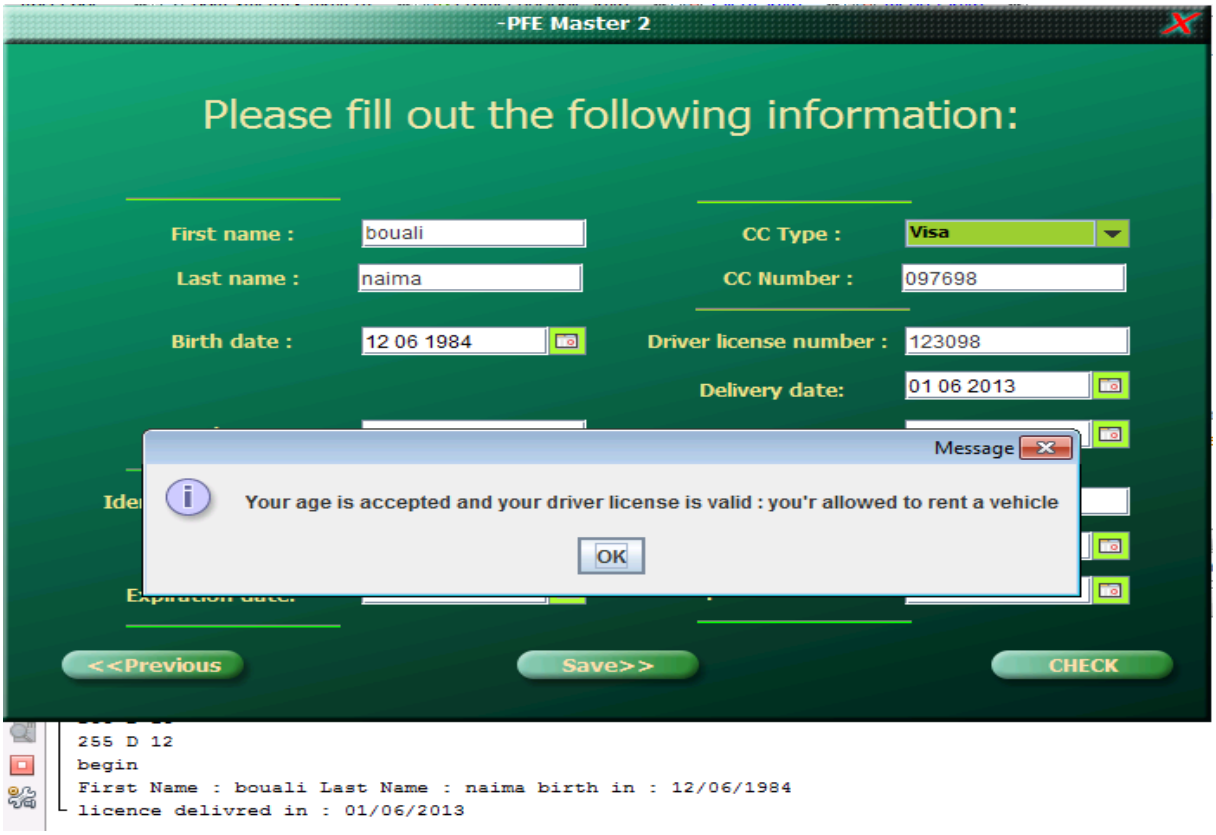

Figure 3.19: Déduction et vérification de l'âge et de validité de permis de conduire de l'utilisateur depuis l'ontologie.

6. CONCLUSION :

Ce chapitre est consacré à résoudre la problématique statique du XPACML avec des outils du web sémantique (ontologie).

Nous avons présenté dans ce chapitre les quatre standards principaux du web sémantique en nous basant particulièrement sur les ontologies. Nous avons présenté les raisons qui nous ont conduit à faire ce choix d'implémentation.

Ensuite nous avons présenté notre implémentation dans laquelle nous avons essayé d'exprimer le déroulement sous forme de scénario e_commerce courant afin de mettre en exergue les opportunités offertes par le modèle de données exprimé en ontologie.

Les liens sémantiques utilisés entre les différents concepts permettent de déduire des pseudo-données qui n'étaient pas explicitement présentées dans le BDD du langage XPACML.

Ainsi le processus de négociation entre l'utilisateur et le SP a pu se dérouler en évitant l'impasse de take it or leave it.

CONCLUSION GENERALE ET PERSPECTIVES:

Dans une optique d'avoir des solutions centrées utilisateur afin de lui donner un maximum de contrôle sur ses données personnelles et les pratiques d'usage appliquées sur ces dernières.

Nous avons présenté dans ce mémoire la problématique principale dans le chapitre1 ;la notion de la vie privée et la protection des données personnelles ,d'ont on a présenté la définition de la sphère privée. Nous avons étudié également comment le service provider peut gérer la protection des données personnelles collecté et stockés, ainsi que l'aspect légal de droit à la sphère privée et les différents technologies de protection des données personnelles.

Ensuite dans le chapitre2 nous avons étudié différents modèles d'architecture de protection, et différents langages d'expression des politiques d'usage (P3P, APPEL, XPref … etc).

On se fixant comme objectif de donner à l'utilisateur le contrôle sur ses données personnelles, nous nous sommes dirigées vers les modèles de contrôle d'accès (MAC, DAC, RBAC, ABAC …etc). On s'est focalisé principalement sur ceux ayant une implémentation largement adoptées (XACML implémenté sur le modèle ABAC)

Enfin nous nous sommes appuyées sur des outils du web sémantique pour résoudre la problématique de l'inflexibilité de XPACML, afin de pouvoir inférer des données à base de celles déjà existantes dans le modèle de données.

Les ontologies étant l'outil du Web Sémantique le plus fiable, on l'a adopter en guise de base pour définir le modèle des données de l'utilisateur et inférer sur ceci. Elle nous a servi également pour définir les préférences de l'utilisateur.

REFERENCES BIBLIOGRAPHIQUES

[1] (Samuel D. Warren and Louis D. Brandeis. The right of privacy. Harvard Law Review, 4 :193–195, 1890).

[2] (Kheira Dari Bekara. Protection des données personnelles cote utilisateur dans le e-commerce; Economies et Finances. Institut National des Télécommunications, 2012. Français.).

[3] (République française. Article 9. In Code civil, 1803).

[4] (République française. Loi numéro 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés; In Journal Officiel de la République Française, January 1978).

[5] (République française. Loi numéro 2004-801 du 6 août 2004 relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel. In Journal Officiel de la République Française, August 2004).

[6] (The European Parliament and the Council. Directive 1995/46/EC of the european parliament and of the council of 24 october 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. In European Union, editor, Official Journal of the European Communities, October 1995).

[7] (The European Parliament and the Council. Directive 2002/58/EC of the European parliament and of the council of 12 july 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector. In European Union, editor, Official Journal of the European Communities, July 2002).

[8] (World Wide Web Consortium. "Platform for Privacy Preferences specification1.1., 2006. http://www.w3.org/P3P/ »).

[9] (S. Riché and G. Brebner. "Storing and accessing user context". In 4th International Conference on Mobile Data Management, pages 1–12, Melbourne, Australia, january 2003).

[10] (European Union's Sixth Framework Programme. PRIME - Privacy and Identity Management for Europe. IST-507591, 2004–2008. https://www.primeproject.eu/).

[11](J.J. Borking. "Privacy incorporated software agent (pisa) : proposal for building a privacy guardian for the electronic age". In International Workshop on Design Issues in Anonymity and Unobservability. volume 2009/2001 of LNCS. pages 130–140, Berkeley. California. USA. 2000).

[12] (World Wide Web Consortium. "A P3P Preference Exchange Language 1.0 (APPEL 1.0)", http://www.w3.org/TR/P3P-preferences/.2002).

[13] (G. Karjoth and M. Schunter. "A privacy policy model for enterprises". IEEE Computer Security Foundations Workshop. IEEE, IEEE Computer Society Press, 2002).

[14](European Union's Sixth Framework Programme. PRIME - Privacy and Identity Management for Europe. IST-507591, 2004–2008. https://www.primeproject.eu/).

[15] (C. Ardagna, J. Camenisch, M.Mont, S. Clauss, S. Crane, Y.Deswarte, G.Hogben, Siani Pearson, L. Pimenidis, T. Roessler, and D. Sommer. "Anarchitecture for privacy-enhancing identity management". LAAS report 05206, LAAS-CNRS. Toulouse. France. 2006).

[16] (M.C Mont, S.Pearson, and P. Bramhall. "Towards accountable management of identity and privacy : Sticky policies and enforceable tracing services". HPL-2003-49. HP Laboratories Bristol. 2003).

[17] (http://www.oasis open.org/committees/tc_home.php?wg_abbrev=security).

[18] (Internet Engineering Task Force. IDsec: Virtual identity on the internet. http://idsec.sourceforge.net/).

[19] (J. Ferber. "Les systèmes multi-agents, vers une intelligence collective". Inter-Editions. Paris. France. 1995).

[20] (J.J. Borking. "Privacy incorporated software agent (pisa): proposal forbuilding a privacy guardian for the electronic age". In International Workshop on Design Issues in Anonymity and Unobservability. volume 2009/2001 of LNCS. pages 130–140, Berkeley. California. USA. 2000).

[21] (S. Riché, G. Brebner, and M. Gittler. "Client-side profile storage". NETWORKING Workshops on Web Engineering and Peer-to-Peer Computing. pages 127–133. Pisa. Italy. 2002).

[22] (S. Riché and G. Brebner. "Storing and accessing user context"In 4th International Conference on Mobile Data Management, pages 1–12, Melbourne, Australia, january 2003).

[23] (P3p: Platform for privacy preferences. http://www.w3.org/TR/P3P11/).

[24] (Appel 1.0: A p3p preference exchange language 1.0. http://www.w3.org/TR/P3P-preferences/).

[25] (Xpref: a preference language for p3p. http://www.sciencedirect.com/).

[26] (Youakim Badr Layth Sliman, Frdrique Biennier. A Security policy framework for context-aware and user preferences in e-services. Journal of System Architecture 55, 2009).

[27] (G. Hogben. A technical analysis of problems with P3P 1.0 and possible solutions.W3C Workshop on the Future of P3P, November 2002).

[28] (M.Hecker, T.S. Dillon, and E.Chang. "Privacy ontology support for ECommerce". IEEE Internet computing Journal.2008).

[29] (K. Bekara, Y. Ben Mustapha, and M. Laurent. "XPACML eXtensible Privacy Access Control Markup *Language* ". Second International Conference on Communications and Networking (ComNet'2010). ISDN 978-1-4244-8840-7.Tozeur. Tunisia. 2010).

[30]('Jean Charlet, Bruno Bachimont, Raphaël Troncy'.Ontologies pour le Web sémantique).

[31] ('http://fr.wikipedia.org/wiki/Ontologie_(informatique)').

[32] (https://fr.wikipedia.org/wiki/NetBeans).