# Protection des données personnelles dans les Systèmes Multi-Agents centrés Utilisateurs

Guillaume Piolle, Yves Demazeau

Laboratoire IMAG-Leibniz, Equipe MAGMA

17 octobre 2006 - Annecy
D2A2

# Introduction

La protection de la vie privée est un problème général dans le domaine du commerce électronique et de l'informatique distribuée.

**Quel est le point de vue spécifiquement multi-agent que l'on peut avoir sur la question ? Y a-t-il une approche spécifiquement multi-agent pour résoudre les problèmes de protection des données personnelles ?**

Compromis habituel : qualité ou fluidité du service, versus protection de la vie privée

**Est-il possible de déplacer ce compromis, ou de déplacer les rapports de confiance ?**

# Introduction

Domaines sensibles pour la protection des données personnelles :

- E-commerce, assistants personnels
- Traçabilité dans un réseau (Internet)
- Dossier médical distribué
- Sondages et vote électronique
- *Privacy* sur le lieu de travail
- Informatique pervasive et ambiante

# Traductions de *Privacy*

Désigne à la fois, suivant le contexte, "vie privée" et "protection de la vie privée". Mention de la "protection des données personnelles (nominatives)" dans la législation francophone.

Définition anglaise sur Wikipedia : *Privacy is the ability of an individual or group to keep their lives and personal affairs out of public view, or to control the flow of information about themselves.*

Spécialisation de la définition par champ d'application (domaine politique, médical, génétique, gouvernemental, sur internet, en entreprise)

Terme anglais : portée beaucoup plus technique et appliquée que la "vie privée" en français (article francophone très sociologique). D'où le choix de "données personnelles".

# Autres interprétations

Publications de recherche en informatique : réduction de la protection de la vie privée à un problème de contrôle d'accès local. Selon la législation européenne, ce n'est qu'un point parmi d'autres.

Inclusion du "SPAM" (communications non sollicitées) dans le problème : plutôt la conséquence d'une faille qu'une composante en elle-même. Conséquence importante et répandue cependant, traitée à part dans la législation.

# Définition implicite, contexte légal

Loi 78-17 "Informatique et Libertés", Directive européenne EC/02/58. Aucun de ces textes ne donne de définition précise pour la protection de la vie privée (*cf* problème de traduction), mais une caractérisaion de ses propriétés :

- **Information**
- **Consentement** de l'utilisateur
- **But** et **justification** de la collecte de données
- Droit de l'utilisateur à la **modification** et à la **Suppression** des données
- Restriction de la **communication à des tiers** des informations collectées ou traitées
- Limitation de la **durée de conservation** des données

Cadre similaire au Canada, mais pas aux Etats-Unis par exemple (ECPA et législations par état)

# Le concept de vie privée dans un environnement multi-agent

Les applications "typiquement multi-agents" génèrent un contexte riche et complexe du point de vue du respect de la vie privée : multiplication des contacts, collaborations inter-agents souvent nombreuses et dynamiques, motivations divergentes.

Cette situation met l'accent sur certaines propriétés en particulier de la protection de la vie privée, notamment la conservation et la transmission des données. Un simple contrôle d'accès devient insuffisant d'une manière évidente.

Un agent autonome peut être conscient du caractère sensible de ses propres informations personnelles (problématiques de traçabilité en informatique ambiante, agents mobiles).

# Des méthodologies typiquement multi-agent ?

Le paradigme multi-agent (et les approches distribuées en général) peut proposer des approches intéressantes pour assurer (ou tout au moins augmenter le niveau de) la protection des données personnelles :

- Division d'une application de service en plusieurs agents de traitement spécialisés ;
- Déportation de traitements sensibles sur des agents plus contrôlés et plus spécialisés (agents matériels ?), aux fonctionnalités limitées ;
- Développement d'agents capables de raisonner sur le contexte normatif ;
- Spécialisation du moteur de gestion de la connaissance des agents cognitifs ;
- Automatisation de la mise en place de schémas de communication protégeant les données personnelles.

# Platform for Privacy Preferences

P3P : projet du W3C, visant à l'établissement d'un standard pour communiquer sur les politiques de protection des données personnelles entre applications web.

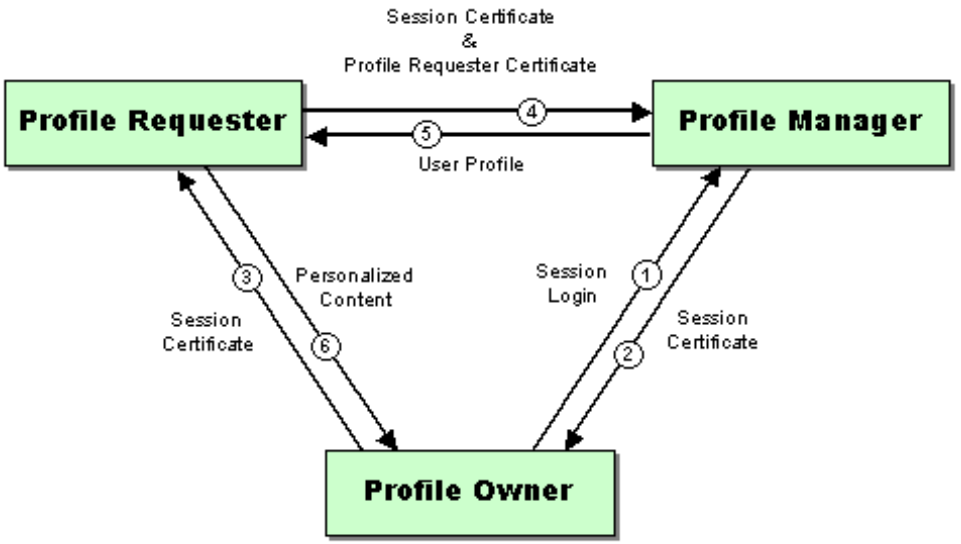
Fonctionnalités :

- Spécification des politiques côté service/serveur ;
- Spécification des exigences côté utilisateur ;
- Détection de la compatibilité P3P des services ;
- Transmission de la politique du serveur au client ;
- Acceptation ou refus automatique par le client (par exemple pour le dépôt d'un cookie).

Aucun seuil minimum n'est fixé concernant les politiques, et aucune garantie n'est donnée sur le respect de ces politiques.

# IETF IDsec

IDsec est un protocole permettant de gérer des identités virtuelles (associées à des profils utilisateur) par le biais d'un service distant. Le système est basé sur l'utilisation de jetons d'accès, délivrés au service désirant accéder à un élément de profil. Le protocole présente plusieurs risques, principalement liés à la concentration des informations personnelles.



# Architectures de type TCPA

Trusted Computing Platform : une plate-forme utilisateur + une puce cryptographique (le Trusted Computing Module, TPM) sur la carte-mère.

Les profils sont générés sur la plate-forme utilisateur, et associés à une identité virtuelle (anonyme). Les identités virtuelles sont ensuite communiquées au service distant.

Le TPM stocke les clés privées et le code des agents responsables de la gestion du profil utilisateur. D'où une certification crypto de la plate-forme, du processus de création du profil et de l'association identité virtuelle / profil par une autorité de certification (Privacy Certification Authority, PCA).

A suivre : dans les spécifications v1.2, un processus sans PCA, basé sur un protocole zero-knowledge (cf rapport du groupe de travail de l'UE sur la protection des données).

# Gestion d'un profil distribué

Approche intéressante pour son centrage utilisateur important.

Stéphanie Riché et Gavin Brebner (HP labs Grenoble) proposent une architecture permettant la gestion d'un profil utilisateur distribué sur un réseau ad hoc de terminaux, centré sur l'utilisateur.

Les profils sont distribués et répliqués de manière dynamique sur les différents terminaux, de la même manière qu'une politique de contrôle d'accès et de consistance des données.

# Synthèse

Questions posées par l'état de l'art :

- Est-ce seulement possible de garantir chacune des composante de la protection des données personnelles ?
- Si ce n'est pas le cas (indécidabilité), comment peut-on contourner les difficultés pour augmenter le niveau de confiance malgré tout ?
- Quels composantes de la protection des données personnelles sont les plus spécifiques aux SMA ?

La protection des données personnelles ne se résume pas au contrôle d'accès, mais s'assurer des "propriétés distantes" apparaît comme très difficile.

# Perspectives de recherche

P3P : peut être intégré tel quel dans une application

TCPA : idées intéressantes sur la certification des processus. Est-il possible de renverser l'architecture ? Est-il possible de s'appuyer sur une certification logicielle, ou de limiter l'usage de la certification matérielle ?

Faire en sorte de pouvoir traiter des profils distribués côté utilisateur.

Recherche future : développement de plusieurs modèles de protection des données (basés sur la certification cryptographique, la gestion de connaissance adaptative...) et d'un banc de tests approprié.

# Questions and comments

Merci de votre attention

Question, commentaires, suggestion et références bienvenus !


Internet Engineering Task Force.
Idsec : Virtual identity on the internet.
http://idsec.sourceforge.net/.


Pearson, S. (2002).
Trusted agents that enhance user privacy by self-profiling.
In *AAMAS-02 Workshop on Deception, Fraud and Trust in Agent Societies*, pages 113–121, Bologna, Italy.


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


World Wide Web Consortium.
Platform for privacy preferences specification 1.0.
http://www.w3.org/P3P.