

PLDAC 2016/2017

Système de contrôle de données personnelles

Encadrants

- Bernd AMANN (LIP6)
- Ali HASSAN (post-doc LIP6)

Contexte

Depuis quelques années les web s'est de plus en plus transformé en réseau mondiale de collecte de données personnelle. Il n'existe pratiquement plus de site web qui ne collecte pas des données sur leurs utilisateurs, en commençant par leur adresse email, leur activité sur le site (suivi des liens), leurs liens sociaux avec d'autres utilisateurs jusqu'à leur géo-localisation, leurs habitudes de sommeil et leur sentiment. Ces données sont ensuite partagées avec d'autres sites sans contrôle particulier de l'utilisateur [2].

Objectifs

L'objectif de ce projet PLDAC est de choisir quelques sites représentatifs (par exemple, facebook.com, sncf.fr et marmiton.fr) et d'étudier le type de données personnelles qu'ils collectent et les techniques de collecte utilisées (cookies). A partir de cette étude, il faut ensuite définir un modèle de données qui permet de stocker et interroger des informations (métadonnées) sur les sites web visités par les utilisateurs (adresse, politique de confidentialité, ...) et l'utilisation et l'échange de cookies entre ces sites pour le partage de données personnelles. Enfin, ce modèle sera utilisée pour construire un programme qui permet de visualiser les données collectées et échangées entre les différents sites [1] et de mieux contrôler la collecte et l'échange de cookies [6].

Plan de travail

- étude de modèles et de techniques d'acquisition de données personnelles sur le web (eg. CookieViz)
- définition modèle de données personnelle générique
- développement d'un système de contrôle de données personnelles
- création d'un prototype

Compétences attendus:

- bonne maîtrise des langages de programmation de haut niveau comme Java, Python, PHP
- modèles de données avancés (niveau MLBDA)
- une expérience en programmation web est un plus

Contact: bernd.amann@lip6.fr

Références bibliographiques

[1] <https://linc.cnil.fr/fr/cookieviz-une-dataviz-en-temps-reel-du-tracking-de-votre-navigation>

[2] <http://binaire.blog.lemonde.fr/2017/01/04/a-la-recherche-des-donnees-perdues/>

[3] https://en.wikipedia.org/wiki/Ad_exchange

[4] <http://www.cookie-checker.com/>

[5] <http://www.editthiscookie.com/>

[6] <http://www.youronlinechoices.com/fr/>