

LES «QUATRE PÉCHÉS» SOCIAUX DU BIG DATA

BIG DATA ET INTERNET DES OBJETS

GRIFES / GiTi / EPFL Alumni
Jeudi 7 mai 2015 @ EPFL



Dr Sami Coll
Université de Lausanne
<http://samicoll.wordpress.com>

Qui suis-je?

h e p i a

Haute école du paysage, d'ingénierie
et d'architecture de Genève



UNIVERSITÉ
DE GENÈVE
FACULTÉ DES SCIENCES
ÉCONOMIQUES ET SOCIALES

FNSNF



The City
University
of
New York

SURVEILLANCE
STUDIES CENTRE

Unil
UNIL | Université de Lausanne



Qu'est-ce que l'Internet des objets ?

- Ubiquitous computing (WEISER, 1991)
 - Après les mainframes, puis les PC, l'informatique «explose», se répand, et devient «invisible»
 - Ce n'est plus à l'utilisateur de faire l'effort d'aller vers la machine, mais l'inverse
- Principes
 1. Ubiquité
 2. Interconnexion globale (Internet IPv6)
 3. Intelligence, action et autonomie (algorithmes)



Qu'est-ce que l'Internet des objets ?

- Pervasive computing (USA), Disappearing computer (UE)
- Ambient intelligence (FP6, 2001)
- Internet of things (UIT, ...)
- Quantified self
- eHealth
- ...



Internet des objets ou pas ?



Internet des objets ou pas ?



<http://youtu.be/btOFclgDG14>

Qu'est-ce que le Big Data ?

- Désigne avant tout une technologie
- Puis, un ensemble de pratiques en voie de généralisation
- Un terme qui fait consensus pour qualifier ces pratiques s'inscrivant dans une réalité sociale, économique et technologique



Commande d'une Big Pizza



<http://youtu.be/RNJ9EEcsoE>

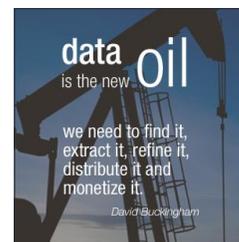
Caractéristiques du Big Data

1. Stockage en masse de données dans des DB non-structurées (≠ DB relationnelles)
2. Algorithmes capables d'un traitement exhaustif, parfois en temps réel, visant la prédiction
3. Secteur privé le plus actif : récolte, analyse de données personnelles pour établir des corrélations entre les profils et les produits



Caractéristiques du Big Data

4. Données utilisées pour satisfaire finalités non nécessairement annoncées pendant la récolte
5. Les données susceptibles d'être croisées avec celles d'autres DB
6. De grandes attentes des milieux économiques, « le nouvel or noir »



IoT/BD : Capitalisme informationnel

1. Pratiques: La capitalisation de données, accumuler pour accumuler, sans objectif préalable
2. Contexte: Capitalisme économique mondialisé reposant sur la collecte des données
3. Réification: Transférabilité des données, comme donnée d'échange. Neutralisation du contexte
4. Conception de l'individu: Hyper-individualisme et individualisation de masse

Les «péchés sociaux» du Big Data ?

1. Le Big Data n'est pas le savoir
2. Le Big Data est discriminatoire
3. Le Big Data n'est pas démocratique
4. La vie privée n'est pas la solution



1. Le Big Data n'est pas le savoir...



JOE PYNE, vétéran du Vietnam: «Vous avez des long cheveux, alors vous êtes une fille?»



FRANK ZAPPA, Musicien: «Vous avez une jambe en bois, alors vous êtes une table?»

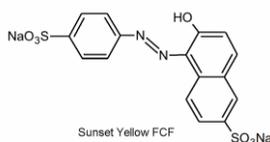
2. Le Big Data n'est pas démocratique

- Le Big Data donne le pouvoir à des acteurs spécifiques de prendre des décisions pour augmenter les profits, prévenir des maladies, réduire les risques, etc.
- Quels sont ces (nouveaux) acteurs ?
- Quelle légitimité ? Une «autorité rationnelle-légale» (WEBER 1968) ? Qui se transforme en autorité technocratique?
- Un pouvoir hybride: au-delà des frontières privé/public. Quelques exemples...

... mais un savoir spécifique

- Type particulier de production de savoir ≠ voies traditionnelles (sciences naturelles, médecine ou sciences sociales)
- Recherches de corrélation, de similitudes... expliquer le lien de causalité n'est plus central. Algorithmes de type datamining et KDD
- «Data-ism» comme idéologie et croyance envers 1/ une «neutralité» des données, 2/ le fait que l'exhaustivité pallierait l'inexactitude des données

Le cas du colorant sunset Yellow



3. Le Big Data est discriminatoire

- «La discrimination totale est la réponse à la discrimination» (un «Big Data Scientist»)
- «Vous en avez marre d'être discriminé? Prouvez-nous que nous avons tort!»
- Risques personnalisés, CRM, médecine personnalisée... le règne de la personnalisation

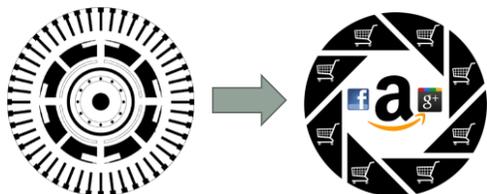


Cassavetes aime-t-il Chuck Norris?



Systèmes de recommandation

- Basés sur le profil des utilisateurs, une «machine de classification» (ELMER 2004)
- Une aliénation culturelle? Risque de réduction de la mobilité sociale (LAHIRE 2004)



4. La vie privée n'est pas la solution

- Protection de la vie privée et des données reposent sur le consentement des utilisateurs
- Les utilisateurs doivent savoir ce qui va être fait de leurs données (LPD art. 4 al. 4), ne doivent pas être utilisées à d'autres fins (al. 3), pas de récolte superflue (al. 2)
- Réaliste dans le contexte des projets de l'IoT (censé être invisible !) et du Big Data ?



La «nouvelle star» aime-t-elle Debussy?



Systèmes de recommandation



- Douglas Stanhope @ Louie C.K. Saison 3 Episode 9

Paradoxes de la vie privée

- Les gens veulent que l'on conserve leurs données personnelles !

- «Pourquoi vous n'avez pas cette information? Je fais quoi maintenant avec mon mixer tout foutu?»



Paradoxes de la vie privée

- Les questions de contrôle, censées protéger la vie privée des clients, produisent les sentiments d'invasion
- «C'est n'importe quoi! Vous êtes une banque ou quoi? Vous voulez aussi mon numéro de slip?»



La vision légale de la vie privée

- Une bulle informationnelle qui entoure chaque personne. Chacun contrôle qui peut y entrer et qui ne peut pas y entrer (STALDER, 2002)



Vision des citoyens et utilisateurs

- Liberté, émotions, limites espace/temps, liberté de décision, information mais non nécessairement informatisées
 - «Mes relations privées»,
 - «Mes décisions personnelles»,
 - «Mon corps, ma santé»,
 - «Mon temps libre», etc.



La vie privée comme expérience

- La vie privée comme expérience vécue au quotidien
- Ca gêne: montrer sa CI, répondre aux questions de contrôle, donner son âge, etc.
- Ca ne gêne pas: transmission automatique des achats quotidiens, analyse approfondie de ces achats par les entreprises



La vision des GAFAs

«If you have something that you don't want anyone to know, maybe you shouldn't be doing it in the first place»
(Schmidt, *The Guardian*, 11/1/10)



- "Privacy no longer a social norm"
(Zuckerberg, *CNBC*, 3/12/09)



Définir la vie privée est politique!

- Qui peut définir ce qu'est la vie privée?
 - «Vous avez une vie privée, vous devez la protéger et nous allons vous dire comme le faire»
- Vie privée est un outil de gouvernance (COLL 2014)

1. Formation d'un savoir sur la vie privée
2. Les systèmes de pouvoir qui en règlent la pratique
3. Les formes dans lesquelles les individus peuvent et doivent s'en reconnaître comme les sujets



Conclusions

1. Du Big Data émerge un savoir instrumental visant l'efficacité. Pouvoir technocratique exercé par des acteurs économiques privés. Avec IoT, pouvoir potentiellement intrusif
2. Algorithmes informatiques, en proposant un horizon de consommation conforme aux profils, qui renforcent l'homogénéité sociale. IoT qui prendra des décisions pour nous, les achat ?
3. Vie privée est outil de gouvernance incitant à adopter une « pratique de la vie privée », compatible avec le capitalisme informationnel



Merci beaucoup !



samicoll.wordpress.com
sami.coll@unil.ch