Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Préposé fédéral à la protection des données et à la transparence
PFPDT

La protection des données à l'heure des mégadonnées «big data – big protection»

Jean-Philippe Walter, Préposé suppléant

GRIFES / GiTi / EPFL Alumni, Lausanne, 7 mai 2015

Introduction

- Environnement sociétal profondément modifié avec Internet, IoTs et TICs
- Société basée sur l'information:
- Tous nos gestes, interactions, comportements, déplacements, etc. susceptibles de générer des données qui alimenteront des algorithmes prédictifs
- Au travers des mégadonnées, pouvoir connaître tout ce qui est possible au sujet d'une personne pour comprendre ses prédispositions, prédire ses actions futures et décider de son futur
- Attrait des méthodes analytiques et analyses prédictives touche tous les secteurs: finance, assurances, émetteurs de carte de crédit, santé, impôts, justice et police, etc.
- Révolution qui se devrait d'être bénéfique pour la société, mais qui peut entrainer une nouvelle forme de domination par les algorithmes et de surveillance ubiquiste et omniprésente

Protection des données

Protection des droits et des libertés fondamentales, notamment droit à la vie privée lors de traitement de données personnelles

- Données personnelles: toutes informations qui se rapportent à une personne identifiée ou identifiable (y. c. adresse IP, numéro tél., tag RFID, adresse Mac, etc.)

Principes de protection des données

- Licéité
- Bonne foi, transparence
- Proportionnalité
- Finalité
- Exactitude
- Sécurité des données

Exigence d’un motif justificatif et légitime pour le traitement:

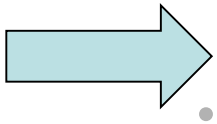
- Loi,
- Consentement,
- Intérêt privé ou public prépondérant

Droits des personnes concernées

- Information sur le traitement
- Droit d’accès
- Droit de rectification
- Droit de s’opposer au traitement
- Droit d’ester en justice

Qu’est-ce que les mégadonnées ?

- Pas une nouvelle technologie, mais une nouvelle manière d’exploitation des données
- Capacité d’analyser d’énormes quantités de données
- Provenant de sources variées
- Utilisant des algorithmes très puissants



- Dégager les constantes sous-jacentes et les corrélations afin de prédire les résultats à venir
- Renforcer les connaissances et la prise de décisions

Autrement dit

«un regard sur le passé pour en tirer des conclusions pour l'avenir» (*dirigeant de Swisscom*)

Celui qui a accès à une quantité massive de données et dispose des capacités d'analyse nécessaires, peut déterminer les besoins et les comportements des personnes cibles, mieux qu'elles ne le savent elles-mêmes.

Ou encore

«Nous sommes passés d'un contexte où des fragments d'information à notre sujet étaient stockés à plusieurs endroits différents, en ligne et hors ligne, à celui où il est possible d'obtenir une image pleinement détaillée de qui nous sommes, le tout étant enregistré et sauvegardé de façon numérique»
(*Terrence Craig et Mary E. Ludloff, Privacy and Big Data, 2011, p. 5*)

- Production d'un savoir non causal sur la base d'algorithmes et de profils auxquels on peut rattacher une personne (abandon de la causalité pour l'intuitif)
- Classification se fait de manière objective et non sur base de préjugés, donc difficile à contester

(*Antoinette de Rouvroy, CRID, Namur*)

D’où proviennent les données ?

Personnes concernées

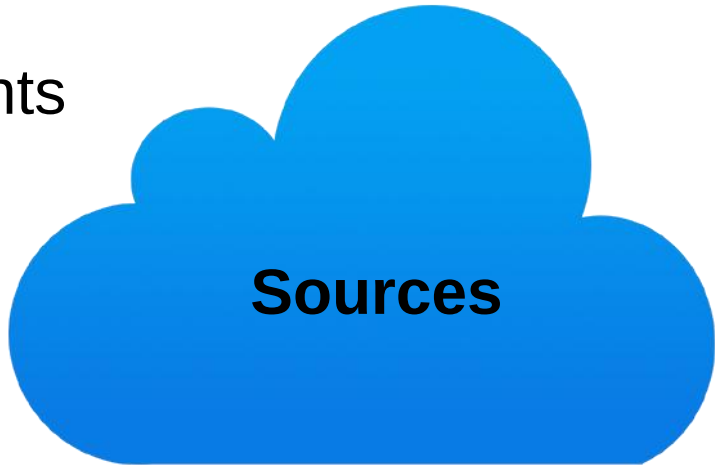
Programme de fidélité

Métadonnées

Dossiers patients

Internet des objets
(Smartphone, Smart Grid, Auto, appareil santé, Apps, etc.)

Données clients



Internet
(réseaux sociaux, moteurs de recherche, sites divers, etc.)

Capteurs

Statistique

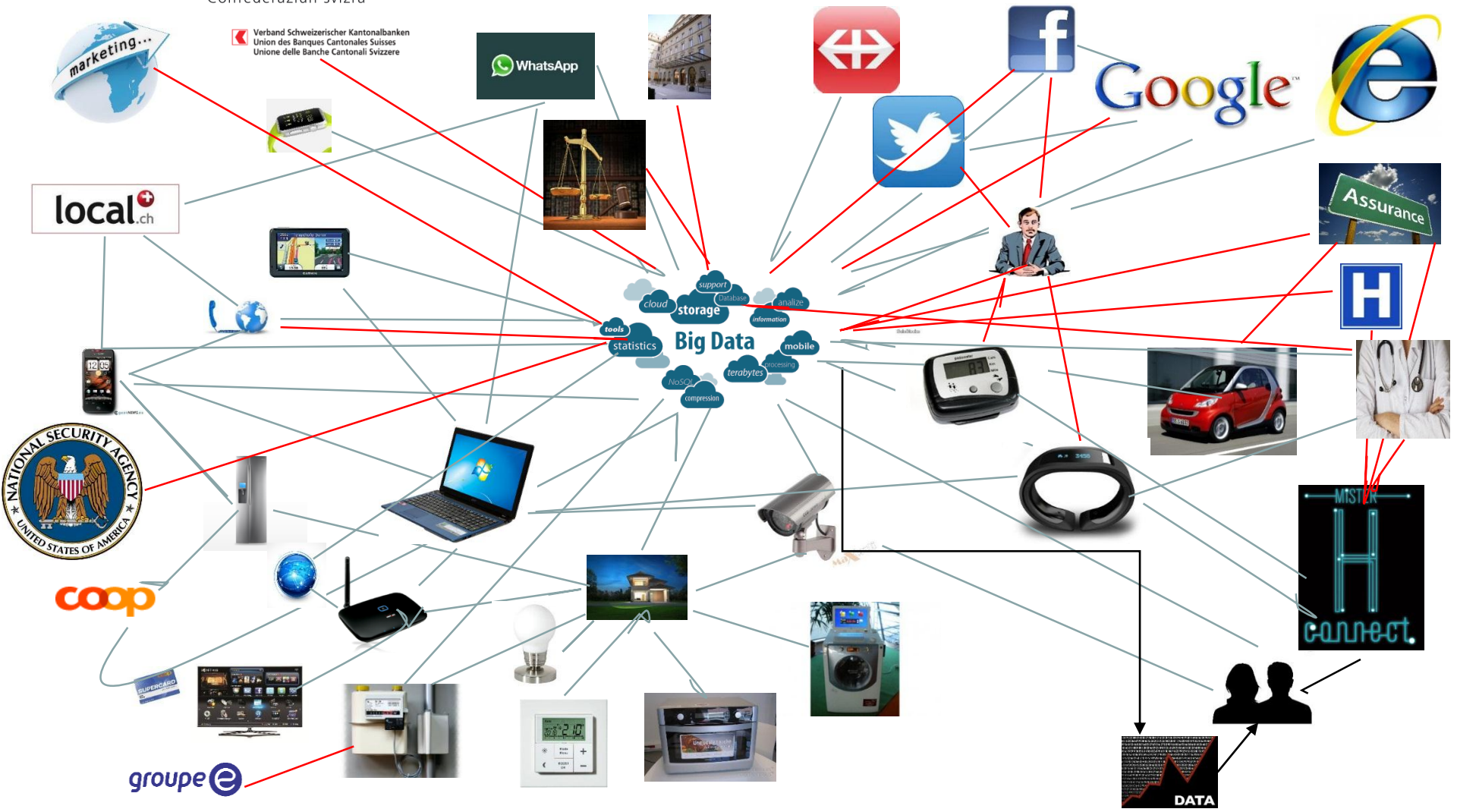
Commerce d’adresses

Administration

Crédit

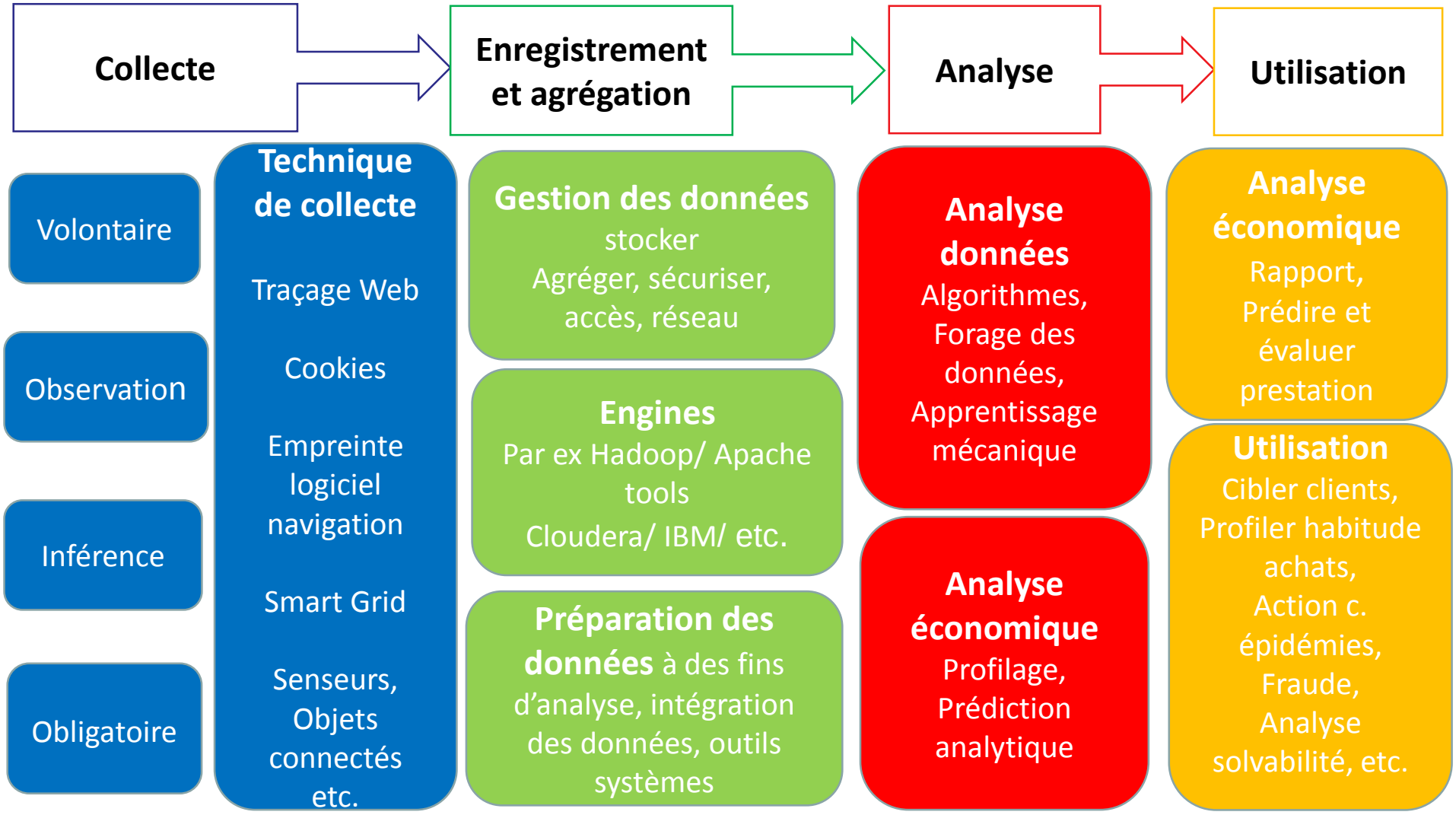
Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Préposé fédéral à la protection des données et à la transparence
PFPDT



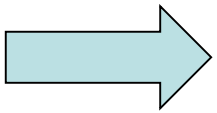
Le parcours des mégadonnées

source: Working Paper on Big Data and Privacy, Privacy Principles under pressure in the age of Big Data analytics, Berling Group, Data Protection Commissionner, May 2014



Analyse mégadonnées

- Prédire des événements et des comportements
- Influencer des décisions et des comportements
- Décider pour autrui
- Etc



- Applications comportent des risques non négligeables pour respect des droits et libertés fondamentales, droit vie privée, dignité humaine
- Changement de paradigme: de présomption d'innocence à présomption de culpabilité
- Applications sans risque, voire bénéfiques pour personnes

Applications non problématiques ?

Recours à des données non personnelles ou anonymes ne permettant pas de réidentifier une personne, par exemple:

- Lutte contre les épidémies, promotion santé,
- Pollution dans les grandes agglomérations, gestion du trafic, développement des transports publics,
- Examen des effets secondaires des médicaments,
- Prévision météo, prévision de cataclysmes naturels, etc.

Applications non problématiques ?

Applications basées sur données personnelles ou qui pourraient être rattachées à des personnes identifiées ou identifiables, mais qui n'ont pas pour objectifs d'identifier la personne, par exemple:

- Gestion des stocks d'une entreprise basée sur données consommation, sans volonté d'orienter choix consommateur
- Développement des offres touristiques sur base analyse métadonnées du téléphone portable
- Prévention des infractions sur base des données récoltées dans le cadre enquêtes police
- Initiative Global Pulse des N.U: prévision des pertes d'emploi, réduction des dépenses, etc. sur base analyse messages postés sur réseaux sociaux

Applications à risque ?

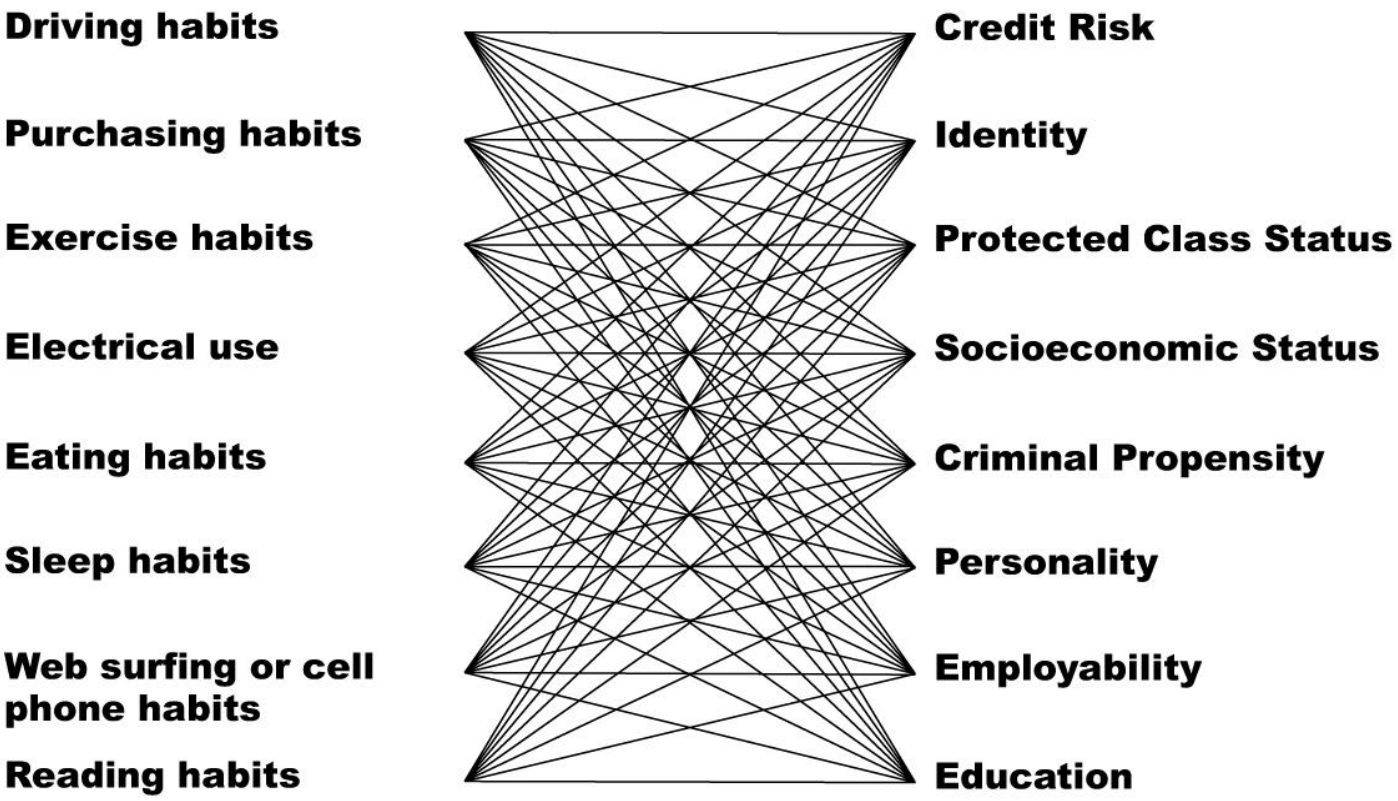
Exploitation de mégadonnées en vue d'influencer le comportement d'une personne identifiée ou identifiable

- **Prédictions préférentielles**: anticiper préférences et inclinations personnelles, orienter les choix: domaine de la consommation et du marketing («TARGET»)
- **Prédictions préventives**: anticiper et prévenir certaines actions susceptibles d'engendre des risques sur le plan social:
  - Risque en termes de crédit et d'assurances
  - Fraude fiscale, à l'aide sociale ou à l'assurance
  - Lutte contre le crime: prédire ou identifier activités criminelles
- **Prédictions prédictives et préférentielles**

Aspects problématiques

- Non respect du principe de proportionnalité et minimisation des données
- Violation du principe de finalité
- Non transparence des traitements
- Exactitude des données: qualité des sources, probabilités ?
- Consentement libre, informée et explicite ?
- Profilage des individus et prédiction sur des événements ou des agissements futurs (maladie, crime, insolvabilité, etc.)
- Non garantie de l'anonymisation des données ou désanonymisation, Risque de réidentification
- Discrimination, stigmatisation
- Droits de personnes concernées (accès, rectification, …)
- Aucune donnée ne peut être qualifiée d'anodine
- Inégalité entre exploitants et personnes concernées

Problèmes Big Data



«Regulating the Internet of Things»
Prof. Scott Peppet, University of Colorado Law School

Protection des données: respect des principes existants

- Licéité doit couvrir l'ensemble des opérations de traitement, y. c. le forage et l'exploitation des mégadonnées
- Loyauté (bonne foi), transparence:
  - Élément permettant de déterminer si l'analyse des mégadonnées est conforme
  - Nécessaire à déterminer les effets du traitement sur les personnes concernées («TARGET») : correspondre aux attentes légitimes et raisonnables des personnes concernées, prévisibilité, éviter risque discrimination
- Proportionnalité (nécessité, minimisation)
- Finalité: déterminée et licite, finalité subséquente pas incompatible avec finalité originale (défi !)
- Exactitude: fiabilité, qualité algorithmes et données

Consentement

- Consentement pour l'utilisation de données personnelles à des fins d'analyse et de profilage doit être effectif, c.à.d libre, informé, spécifique et explicite:
  - pas clause générale (contrat d'adhésion),
  - diversification des consentements en fonction applications,
  - mise en place d'une architecture de véritable choix:
    - Maîtrise sur nos données implique le droit de choisir si communique ou non des données !
    - Neutralité du design pour ne pas fausser le choix
- En l'absence de consentement et de choix, nécessité d'un motif justificatif et légitime prépondérant

Anonymisation

- Anonymisation des données pour diminuer ou éliminer risque de violation protection des données si:
  - Techniquement irréprochable
  - Réidentification (quasi-) impossible : randomisation, généralisation des données, agrégation

Renforcement de la transparence

- Information sur quelles données sont collectées, comment elles vont être exploitées, pour quelles finalités et à qui elles vont être communiquées

- Information sur le fonctionnement des algorithmes utilisés pour établir profil et prise de décision

- Renversement du fardeau de la preuve

Renforcement des droits des personnes concernées

- Droit de connaître son profil et ses données, y. c. les sources d'où proviennent ces données
- Droit de rectification et d'opposition à l'utilisation des données à des fins d'analyse et de profilage
- Droit de faire valoir son point de vue lors de décision automatisée affectant de manière significative une personne
- Droit d'obtenir raisonnement qui sous-tend au traitement lorsque résultats / profils sont appliqués à personne
- Droit à l'oubli numérique, de portabilité ou droit de restitution des données

Autres garanties

- Privacy by design, Privacy by default, Privacy Enhancing Technologies
- Evaluation et analyse de risques «protection des données»
- Certification ou procédure d’autorisation des applications sensibles
- Autoréglementation avec procédure d’agrément
- Audit régulier, notamment des algorithmes
- Encadrement du profilage
- Obligation de vérifier et démontrer le respect des exigences de protection des données
- Contrôle et sanctions

Conclusions

- **L'ère des mégadonnées et de l'analyse prédictive est une réalité !**
- **De nombreux bénéfices pour les personnes et les collectivités sont attendus de l'exploitation des mégadonnées et des applications qui en découleront**
- **L'exploitation des mégadonnées liée au traitement exponentiel de données à caractère personnel soulève des questions sociales, juridiques, éthiques et économiques**
- **Dans ce contexte, le droit à la protection des données joue un rôle crucial: les bénéfices attendus des mégadonnées passent par des garanties fortes des droits des personnes concernées**
- **Au centre des préoccupations doit impérativement figurer la personne en tant que sujet des données et non objet des données**

La technologie se doit d'être au service de l'Homme et non l'inverse !


Merci !