CIL CONSULTING
Protection des Données

BIG DATA et PROTECTION DES DONNÉES PERSONNELLES: ENJEUX ET LIMITES DU REGLEMENT EUROPEEN

Limites des normes impératives face aux enjeux éthiques, économiques et sociaux liés au développement des technologies de l'information et des communications dans le cadre des Big Data.

Florence BONNET

15 Octobre 2015 - Séminaire #BigDataAristote Ecole Polytechnique

Le Big Data est une révolution

Révolution : Changement **brutal et violent** pour l'économie et pour les individus …

- **Déluge de données**

Il y a plus de systèmes connectés dans le monde que de personnes : connectivité aussi essentielle que l'air que nous respirons

Baisse des coût (ex. beacon gratuits) + augmentation des capacités de traitement (volume et temps réel) → « data driven world »

- **Règne de l'algorithme…..qui impacte directement l'individu**

Gerd Leonhard : « humarithms »

Ere du scoring cf. Chine: scoring social,

de l'ultra personnalisation de l'offre,

De la « gamification »

- Multiplication des **systèmes automatisés sans intervention humaine** cf. programmatic marketing

- **Imprévisibilité** : incompatible avec le lent processus de création de la norme juridique

http://www.computerworld.com/s/article/91109/Sidebar_The_Mosaic_Effect

CIL CONSULTING Protection des Données

Le Big data est source de risques plus nombreux et plus graves

Risque = **sources de risques** qui exploitent des **vulnérabilités** et qui rendent possible un **évènement plus ou moins grave**

- **Plus de sources de risques** cf. multiplicité des acteurs dans “l’espace numérique” , « IoE »…

Ecosystème de start up (pas sensibilisées, pas de budget)

- **Plus de vulnérabilités** (techniques, personnes…)

Cf. capteurs +/- fiables

Cf. Biais inconnus, représentativité des résultats difficile à évaluer, nouveaux types d’erreur

- Plus de données issues de sources diverses : facilite la **ré-identification et la révélation de données sensibles**

cf. data fusion peut faire naitre de nouvelles infos ; mais aussi plus de données erronées (biais)

- **Gravité des évènements**: Opacité de technologies toujours plus performantes pour suivre, profiler, surveiller, ré-identifier des données

CIL CONSULTING Protection des Données

Une nécessaire évolution du cadre juridique de la protection des données

- **Consensus mondial**: risques pour la vie privée

Mais pas de consensus universel: quoi? Comment?

- **Conflits d'intérêts**

  cf. Safe Harbor

  cf. Coexistence de stratégies conflictuelles au sein de l'UE: Promotion d'une « data-driven economy » vs GDPR

- **Le nouveau règlement**

  -Indispensable: 1995 vs 2015

  -Mais risques: frein à l'innovation - détournement de la loi

Challenge: Comment prévenir des menaces pas toujours identifiées et dues à l'utilisation de solutions encore inconnues (nouvelles données, nouveaux algorithmes) ?

CIL CONSULTING Protection des Données

Applicabilité territoriale

Directive	Commission – Parlement - Conseil
Traitement effectué dans le cadre des activités d'un établissement du responsable du traitement dans l'UE, Ou recourt à des moyens de traitement situés sur le territoire d'un Etat membre (ex. collecte via IoT).	Traitements de données dans le cadre des activités d'un établissement du responsable de traitement ou du sous-traitant dans l'UE, Traitement de données de citoyens européens par un responsable établi hors UE pour : -offre de biens et services -ou suivi des comportements des personnes

Objectif → harmonisation:

- Mais le règlement renvoie à la loi (cf. profilage)
- Risques d'interprétations divergentes entre les CNIL européennes
  cf. « consistancy mechanism » : complexe et long.
  cf. Quid divergences au sein du EDPB?
- Globalisation: contrôles impossibles sans relais locaux

Des acteurs du net plus puissants que les états

Question :
Is Google Too Powerful Too Be Held Accountable?

CIL CONSULTING
Protection des Données

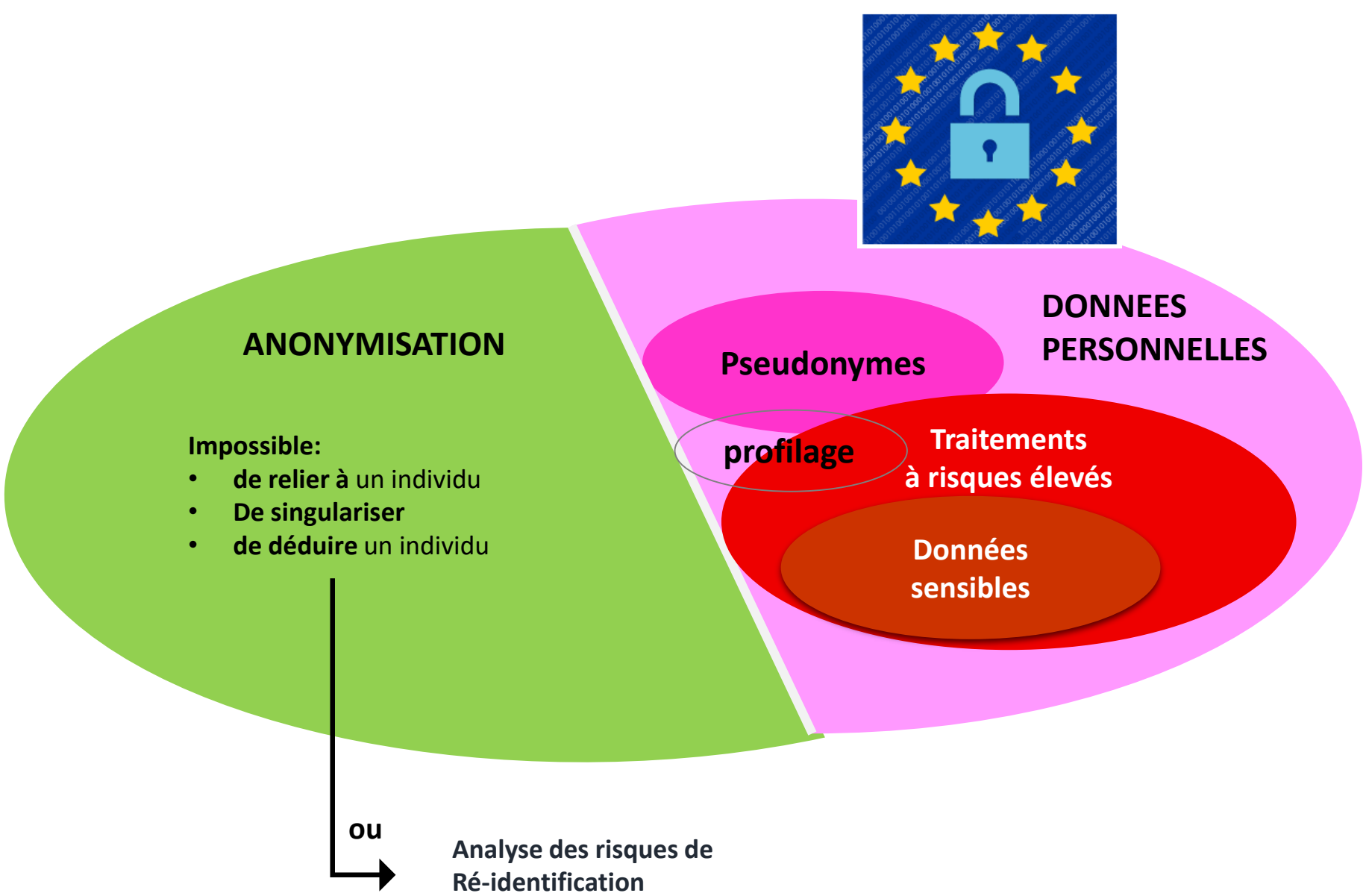
La notion de données à caractère personnel au cœur des débats

Directive 95/46	GDPR
Toute information relative à une **personne physique identifiée ou qui peut être identifiée, directement ou indirectement**, par référence à un numéro d'identification ou à un ou plusieurs éléments spécifiques, propres à son identité physique, physiologique, psychique, économique, culturelle ou sociale.	Information concernant une **personne qui peut être identifiée** par exemple par un identifiant, des données de localisation…ou par une ou plusieurs caractéristiques physiques, génétiques, mentales, économiques, culturelles, par son identité sociale. **NB** Art.10: Si le Traitement ne permet pas d'identifier une personne, le RT n'a pas à collecter des info. supplémentaires pour se conformer au règlement

La réglementation ne s'applique pas aux données anonymes; or le Big Data remet en cause la distinction entre données personnelles et non personnelles : **les enjeux sont éthiques**

Paradoxe: anonymat et tracking
Il reste pratiquement toujours un risque d'individualisation ou de corrélation ou d'inférence cf. avis G29

Le Big data peut cibler des groupes - La cible n'est pas l'individu en tant que tel mais le groupe auquel il est assimilé selon probabilités ex. surpoids – publicité au moment où la personne est la plus vulnérable devant son écran


ANONYMISATION
Impossible:
• de relier à un individu
• De singulariser
• de déduire un individu
ou
Analyse des risques de
Ré-identification
DONNEES
PERSONNELLES
Pseudonymes
profilage
Traitements
à risques élevés
Données
sensibles

Le principe de finalité spécifique

Les données sont collectées pour des **finalités déterminées**, explicites et légitimes et **ne sont pas traitées ultérieurement de manière incompatible avec ces finalités.**
Exceptions: finalité scientifique, statistique ou historique ne sont pas incompatibles

Les données sont adéquates, **pertinentes et non excessives** au regard de ces finalités

Le Data mining permet de découvrir ou de déterminer des inférences ou des modèles qui n'étaient pas connus au préalable (ne savent pas ce qui va en sortir): **résultats ni intuitifs ni prévisibles**

Big Data : n = all c/ n=sample
Vise à traiter le plus de données possible pour leur donner un sens (nouveau) donc **toutes les données sont potentiellement pertinentes.**

Avis du G29: base légale + finalité compatible

Statistiques: peuvent être à des fins commerciales
Test de compatibilité : rigoureux dans le cadre du big data
2 scénario : tendances ou s'intéresse à l'individu (analyse ou prédiction préférences) → pas compatible → Opt In (ex. tracking et profilage pour marketing direct, pub. comportementale, courtage de données, publicité localisée, recherche marketing sur la base de tracking localisé)
Exception si loi nationale ex. lutte c/fraude vav du service
Cependant un traitement ultérieur n'est pas forcément incompatible (rôle des mesures de sauvegarde techniques et organisationnelles pour une séparation fonctionnelle ex. droits des personnes, anonymisation, pseudonymisation, aggrégation de données, PETs')

Commission	Parlement	Conseil
Si finalité du traitement ultérieur non compatible → **nouvelle base légale** (**consentement**, **contrat**, loi, vie de la personne, intérêt public) SAUF intérêt légitime	Pas de traitement ultérieur non compatible **Traitement ultérieur incompatible possible?**	Si finalité du traitement ultérieur incompatible → Nouvelle base légale **y compris intérêt légitime** du responsable de traitement ou d'un tiers **s'ils priment sur les intérêts des personnes** cf. marketing?

Droit d'opposition aux mesures basées sur le profilage

Directive 95/46/EC (art.15)	Commission	Parlement	Conseil
Toute personne a le droit de ne pas être soumise à: -une décision produisant des **effets juridiques à son égard ou l'affectant de manière significative**, -prise **sur le seul fondement d'un traitement automatisé de données**, -destiné **à évaluer certains aspects de sa personnalité**, tels que son rendement professionnel, son crédit, sa fiabilité, son comportement, etc.	Toute personne a le droit de ne pas être soumise à: -une mesure qui produit des **effets juridiques à son égard ou qui l'affecte de manière significative**, -basée sur un **traitement automatisé**, -destiné à **évaluer certains aspects de sa personnalité ou pour analyser ou prédire** en particulier, ses performances au travail, sa situation économique, sa localisation, sa santé, ses préférences personnelles, sa fiabilité ou son comportement.	Toute personne a le droit de s'opposer à: -Un **traitement automatisé** de DCP, - Destiné à **évaluer ses caractéristiques ou pour analyser ou prédire** ses performances au travail, économiques, sa localisation, sa santé, ses préférences personnelles, sa fiabilité ou son comportement.	Toute personne a le droit de ne pas être soumise à: -une décision qui produit des **effets juridiques à son égard ou l'affecte de manière significative**. -**basée uniquement sur un traitement automatisé et notamment au profilage** (traitement automatisé de DCP **destiné à évaluer les caractéristiques d'une personne ou pour analyser ou prédire** ses performances au travail, sa situation économique, sa localisation ou ses déplacements, sa santé, ses préférences personnelles, sa fiabilité ou son comportement),
Exception : -Décision prise dans le cadre de la conclusion ou de l'exécution d'un contrat, -Décision autorisée par une loi.	**Exception** : -Contrat, -Loi, -**Consentement**	**Exception** : -Contrat, -Loi, -**Consentement** **Droit à intervention humaine**	**Exceptions :** -Contrat, -Loi, -**Consentement explicite.** **Droit à intervention humaine**

La personne peut consentir à son profilage

Limites au profilage basé sur le traitement de données sensibles

Loi I&L	Commission	Parlement	Conseil
Article 25 - autorisation de la CNIL : -Traitements automatisés **susceptibles d'exclure des personnes** du bénéfice d'un droit, d'une prestation ou d'un contrat; -Les traitements automatisés ayant pour objet **l'interconnexion de fichiers** et dont les finalités principales sont différentes.	**Interdiction** : Traitement basé uniquement sur des données sensibles art.9 (+ données génétiques, condamnations, mesures de sécurité) **Quid des données qui deviennent sensibles?** **Sauf autorisation** par UE ou loi nationale: -sécurité publique -prévention, investigation, détection et poursuite d'infractions criminelles -intérêt éco, financier y compris taxe -détection, prévention, investigation, de failles éthiques pour des professions réglementées -contrôle, inspection d'une autorité -protection des données ou droits et libertés des personnes. **Actes délégués**	**Interdiction**: Ou traitement basé uniquement sur des données sensibles art.9 (+ données génétiques, biométriques, sanctions admin. , jugements, condamnations, mesures de sécurité) Ou si effets discriminatoires **Aucune autorisation possible** **Profilage données sensibles pas exclu** **Lignes directrices du board**	**Interdiction** : Traitement basé uniquement sur des données sensibles art.9 (+ génétique) **sauf consentement explicite** (si pas interdit par loi nationale) ou intérêt public. **Sauf autorisation** par UE ou loi nationale: -sécurité nationale -prévention, investigation, détection et poursuite d'infractions criminelles ou exécution de peines criminelles ou prévention des menaces à la sécurité publique -objectifs importants d'intérêt public en particulier économique, financier, santé publique et SS, -protection de l'indépendance judiciaire -détection, prévention, investigation, de failles éthiques pour des professions réglementées -contrôle, inspection d'une autorité -protection des données ou droits et libertés des personnes -demande de droit civil

Art.33 Etude d'impact sur la protection des données – Notion de « risque »

CJUE: la directive 95/46 vise à garantir la **protection des libertés et des droits fondamentaux** des personnes physiques notamment du **droit au respect de la vie privée et du droit à la protection des données à caractère personnel** (art.7 et 8 Charte Droits Fondamentaux)

Commission	Parlement	Conseil
Le Responsable de Tt ou le Sous-traitant Si **risques spécifiques** i.e.: -Profilage -Traitement à large échelle de données sensibles, recherche épidémiologique, études maladies mentales ou infectieuses, données relatives aux enfants, génétiques, ou biométriques -Contrôle d'espaces publics en particulier via vidéosurveillance -Cas où l'autorité nationale doit être consultée	Le Responsable de Tt ou le Sous-traitant, Si **risques spécifiques** i.e. -traitements concernant + 5000 pers. sur 12 mois, -traitement à large échelle de données sensibles, de localisation, d'enfants, d'employés -profilage, -soins, recherche épidémiologique, enquêtes maladies mentales ou infectieuses -Contrôle d'espaces publics à large échelle, -violation de DCP avec impacts pour le sujet, -Cœur de métier = contrôle régulier et systématique du sujet -Cas prévu par l'autorité nationale	-Le RT -Si **risques élevés** tels que : Discrimination, Usurpation d'identité ou fraude, Perte financière, Atteinte à la réputation, Violation des pseudo, Violation secret professionnel Désavantage économique ou social significatif Concerne en particulier : le profilage Liste établie par chaque autorité nationale

En pratique:
-a minima une description du projet, **évaluation des risques** et des mesures existante - Complexité, temps, coût → doit débuter en amont des projets
-intérêt si évite la demande d'autorisation préalable
-WP29 avis 203: comment faire une EIVP pour toute nouvelle application? d'ici 2020, il y aura 50 billions de systèmes connectés?
-Harmonisation entre autorités nationales?

CIL CONSULTING
Protection des Données

Transparence, consentement libre et éclairé…

Plus d'information,
Langage clair
Et adapté notamment pour les enfants

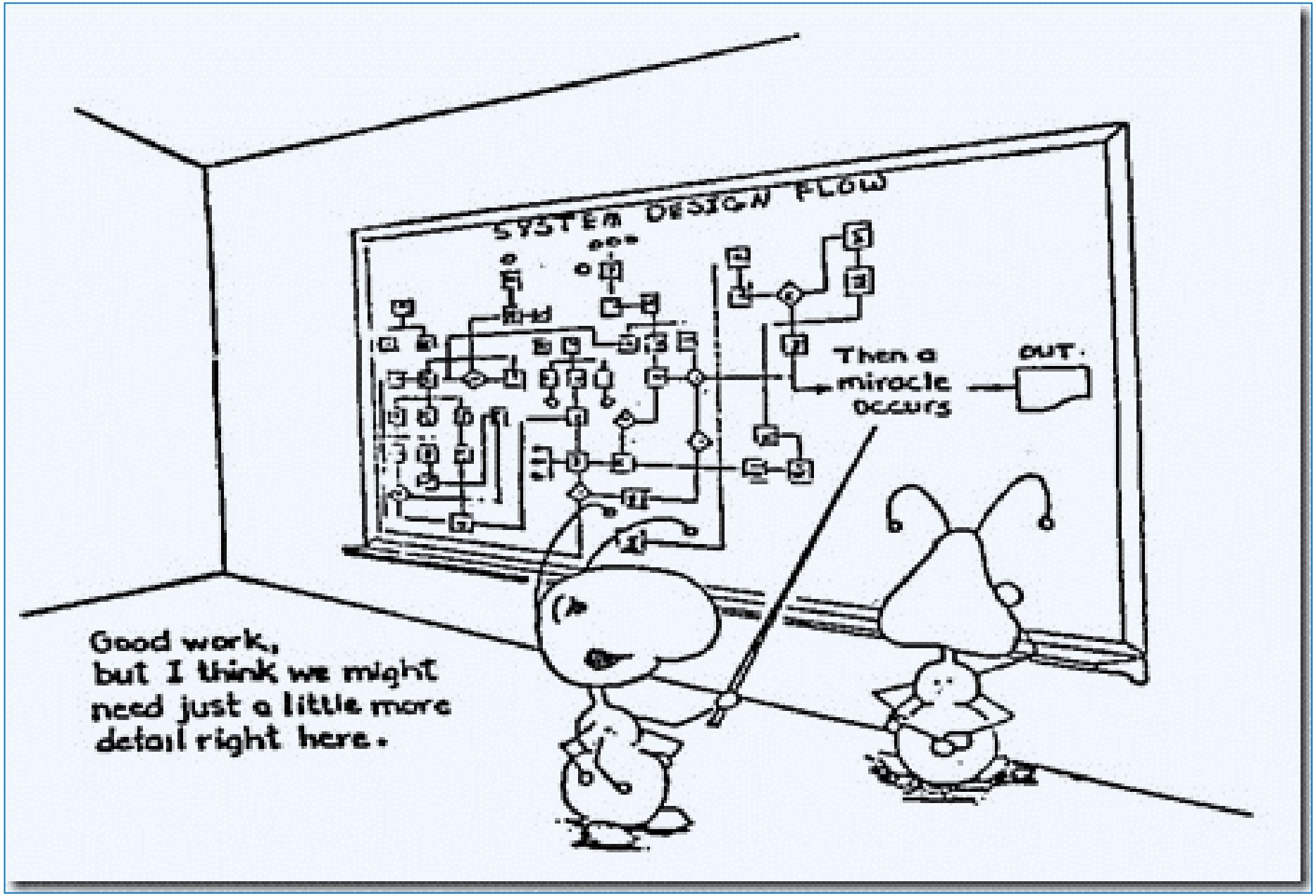
-Choix de refuser?
-Personnes vulnérables: ex. parents des enfants ??
-Pas de sensibilisation aux éventuelles corrélations, inconscience….d'ailleurs qui sait cf. black box?
-Politiques qui induisent en erreur ou pas respectées (Cf. Nomi opposition au retail tracking)

-En pratique:
IoT, Beacon : où ? comment ? (Cf. CNIL 16/072015 JC Decaux) – QRcode?
Qui lit les politiques?
Dialogue direct entre les apps (ex.Uber et Hilton HHonors) Qui va lire toutes les politiques de confidentialité pendant son trajet?
Quid décisions prises grâce à mes données vis-à-vis d'autres personnes?
Combien de personnes savent exploiter leurs données ex. Google?

Commission – Parlement- Conseil : Informations sur le profilage

-mesures basées sur le profilage
-éventuels effets pour le sujet
-la logique , signification (conseil)

CIL CONSULTING
Protection des Données


SYSTEM DESIGN FLOW
Then a miracle occurs
OUT.
Good work, but I think we might need just a little more detail right here.

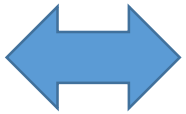
Droit à l'oubli et à l'effacement….mission impossible

Directive 95/46 et CJUE	Commission	Parlement	Conseil
Toute personne physique justifiant de son identité peut exiger du responsable d'un traitement que soient, verrouillées ou effacées les données à caractère personnel la concernant, **si:** **Les données sont inexactes, incomplètes, équivoques, périmées,** **ou si le traitement est interdit.** **CJUE 13052014:** droit au déréférencement sur les moteurs de recherche Exception: droit à l'information du public si personne publique	**1-Droit à l'effacement sans délai et suppression des divulgations à venir,** Particulièrement si la personne était enfant, **Si:** -ne sont plus nécessaires -**retrait du consentement** -durée expirée -**droit opposition** -non-conformité **2-Données rendues publiques par le RT**: **mesures raisonnables pour informer** les tiers concernés Exceptions: -Liberté d'expression -Intérêt public secteur santé, -Recherche Hist. Scient. Stat. -Loi 3-Restriction (contestation, preuve) **4-Mécanismes garantissant l'effacement à expiration des durées ou revue périodique**	**1-1-Droit à l'effacement sans délai**, si le RT est à même de s'assurer qu'il s'agit de la personne, **2-Si données rendues publiques par le RT** sans justification: le RT prendra les **mesures raisonnables pour effacer** les données Exceptions: idem 3-restrictions 4-idem	**1-Droit à l'effacement sans délai**, particulièrement si était enfant, **Si:** -ne sont plus nécessaires -**retrait du consentement** et s'il n'existe pas d'autre base légale, -**opposition de la personne et il n'y a pas d'intérêt légitime du RT qui prime**, -**opposition Trait. marketing direct** -traitement illégal, -loi, -si données collectées pour offre de services de société de l'information **2- données rendues publiques par le RT:** en tenant compte des techno. disponibles et du coût, le RT prendra les **mesures raisonnables pour informer** les tiers Sauf impossibilité ou efforts disproportionnés Exceptions: idem 3-Restrictions

La carotte et le bâton

Besoin d'incitations:
Politique publiques, marchés publics cf.
Data Protection ByDesign : quel intérêt pour les entreprises? GDPR: selon état de l'art et coût d'implémentation)

- Collecte et durée de rétention minimum
- Accès limités



Compatibilité Big Data **?**

Besoin de contrôles…a posteriori:
Documenter, c'est bien mais très insuffisant.
Responsabilité algorithmique – audit de codes

Besoin de sanctions dissuasives:
Economie de marché → « qu'est-ce que je risque? » - Jusqu'à 2% CA global ?

Impliquer les professionnels ≠ lobbies

Encourager les professionnels « responsables »:
Solution jamais 100% bonne ou 100 % mauvaise, c'est une question de choix d'utilisation : on tend à interdire par crainte de ce qui pourrait arriver c/ analyse des risques cf. « cette société française n'a qu'à aller vendre sa solution en Amérique du Sud »
Conséquences dommageables pour l'économie et pour la société:

- Détournement et opacité (pour les moins responsables),
- autocensure (pour les plus responsables),
- Frein à l'innovation

CIL CONSULTING
Protection des Données

Délivrer des gages de confiance

Certification , labels et codes de conduites : mais à quelles conditions ?

Garanties (indépendance, contrôles, sanctions…) : critères par actes délégués – harmonisation via EDPB - codes soumis à l'avis de l'autorité nationale et codes adoption par la Commission

Besoin de mécanismes de reconnaissance (cf. standards, normes)

Prendre le lead au niveau européen

Mais attention - corégulation ≠ autorégulation

Eco system de confiance

Protocles, sécurité

Maitrise par l'individu: impliquer et responsabiliser les individus mais on ne peut pas faire peser la charge sur l'individu – Besoin d'indicateurs de confiance

Intermédiaires de confiance: ex. tiers indépendants pour classifier les appli selon des profils « privacy » (profil pouvant être interdit selon le contexte)

Sécurité et gestion des risques

Quels contrôles ? Cf. Limites vav transferts

Sujet à interprétation: cf. pseudonymes et anonymisation

Harmonisation et globalisation

- Outils/ressources communes à l'ensemble des DPA: lutter contre le fardeau administratif
- Un EDPB fort, des CNIL « indépendantes », avec des moyens renforcés
- Une application effective et uniforme des règles

"Google policy is to get right up to the creepy line and not cross it…. We don't need you to type at all. We know where you are. We know where you've been. We can more or less know what you're thinking about."[1]

–Google CEO (now Executive Chairman) Eric Schmidt (2010)

Merci de votre attention

Questions ?

Florence BONNET – florence.bonnet@cil-consulting.com

@FlorenceBonnet

CIL CONSULTING – www.cil-consulting.com
www.protection-des-donnees.com

171 avenue Charles de Gaulle – 92200 NEUILLY SUR SEINE

@CILCONSULTING