Privacy protection in the digital age

Antoine Boutet

antoine.boutet@insa-lyon.fr

Privacy protection

- **Why should you care?**
- What legal means?
- What technical means?
- How to assess and reduce risks?

The visible part of the Web

- **Facebook** knows what you like, who you interact with, ...
- **Instagram** knows where you spend your vacation,...
- **Amazon** knows what you buy
- **YouTube** knows what you're watching
- **Google** knows what you're thinking

They think they know who you are, what you are going to do soon,...

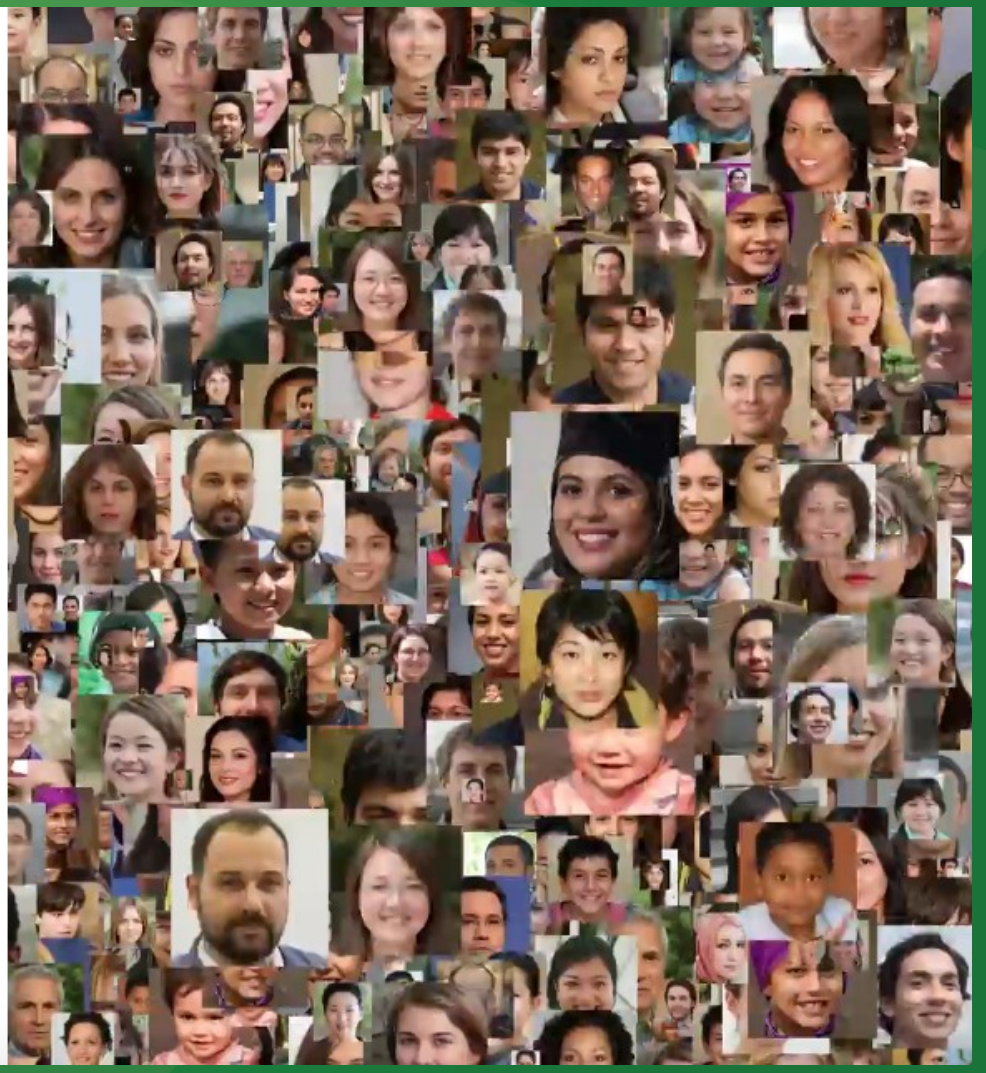
NB: growing possibilities for tracking (internet of things, “cross device tracking”, etc.) and cross-referencing of data

Reuse risks

The New York Times

The Secretive Company That Might End Privacy as We Know It

A little-known start-up helps law enforcement match photos of unknown people to their online images — and "might lead to a dystopian future or something," a backer says.



Reuse risks

Le Monde

Consulter le journal

Se connecter | S'abonner

ACTUALITÉS | ÉCONOMIE | VIDÉOS | OPINIONS | CULTURE | M LE MAG | SERVICES

PIXELS · VIE PRIVÉE

Partage

Reconnaissance faciale : une start-up analyse les photos des réseaux sociaux pour la police américaine

Le « New York Times » a publié une enquête sur cette entreprise financée par Peter Thiel (Palantir, Facebook) qui fournit son logiciel à la police américaine.

Publié le 20 janvier 2020 à 12h41 - Mis à jour le 20 janvier 2020 à 15h05

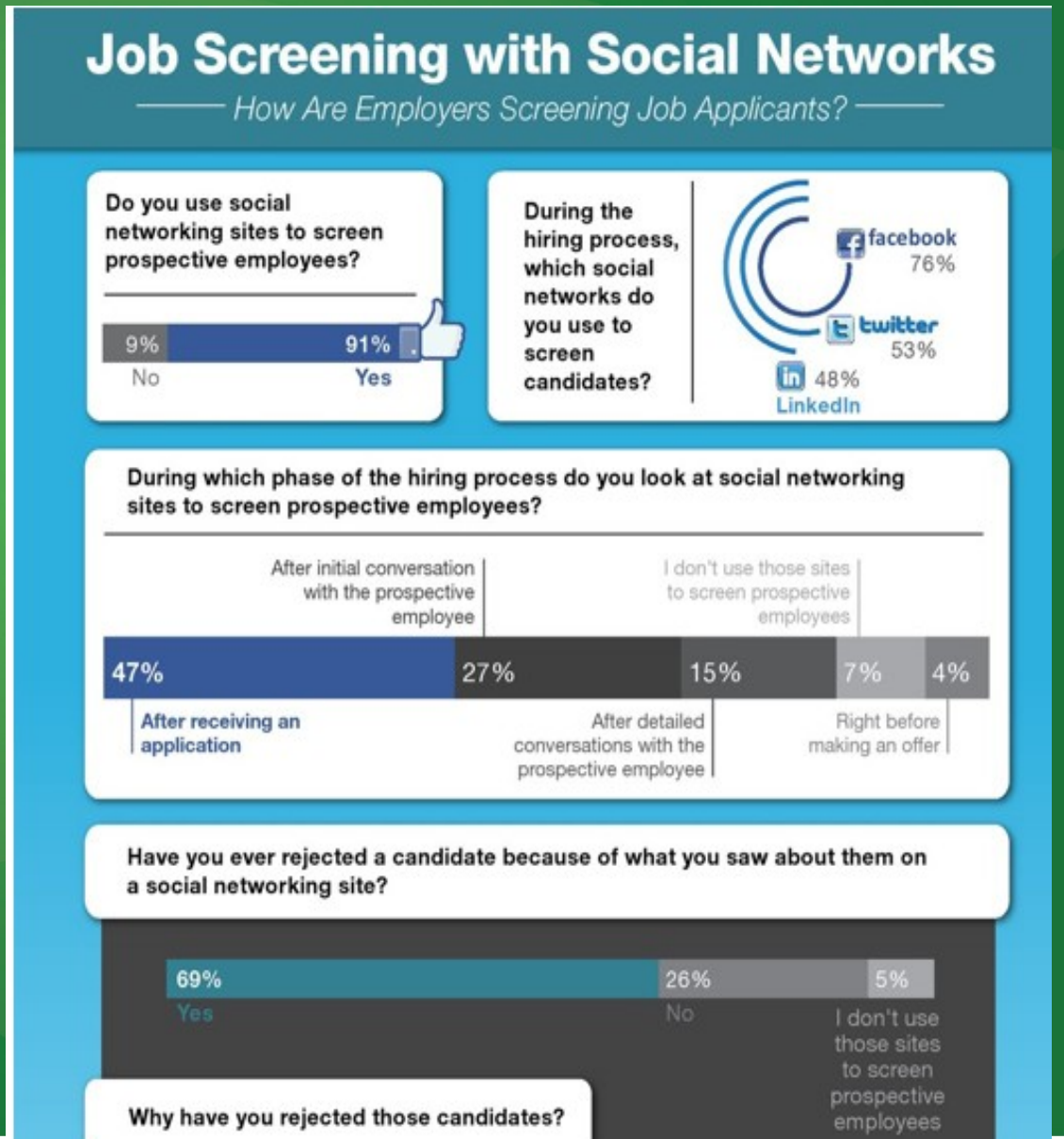
Lecture 2 min.

Pouvoir comparer en quelques instants une photographie avec une base de données de plus de trois milliards de photographies publiées par tout un chacun sur les réseaux sociaux : la promesse a séduit six cents services de police aux Etats-Unis. Cet outil est proposé par Clearview, discrète start-up américaine à laquelle le *New*

Les plus lus

Risks of self-exposure

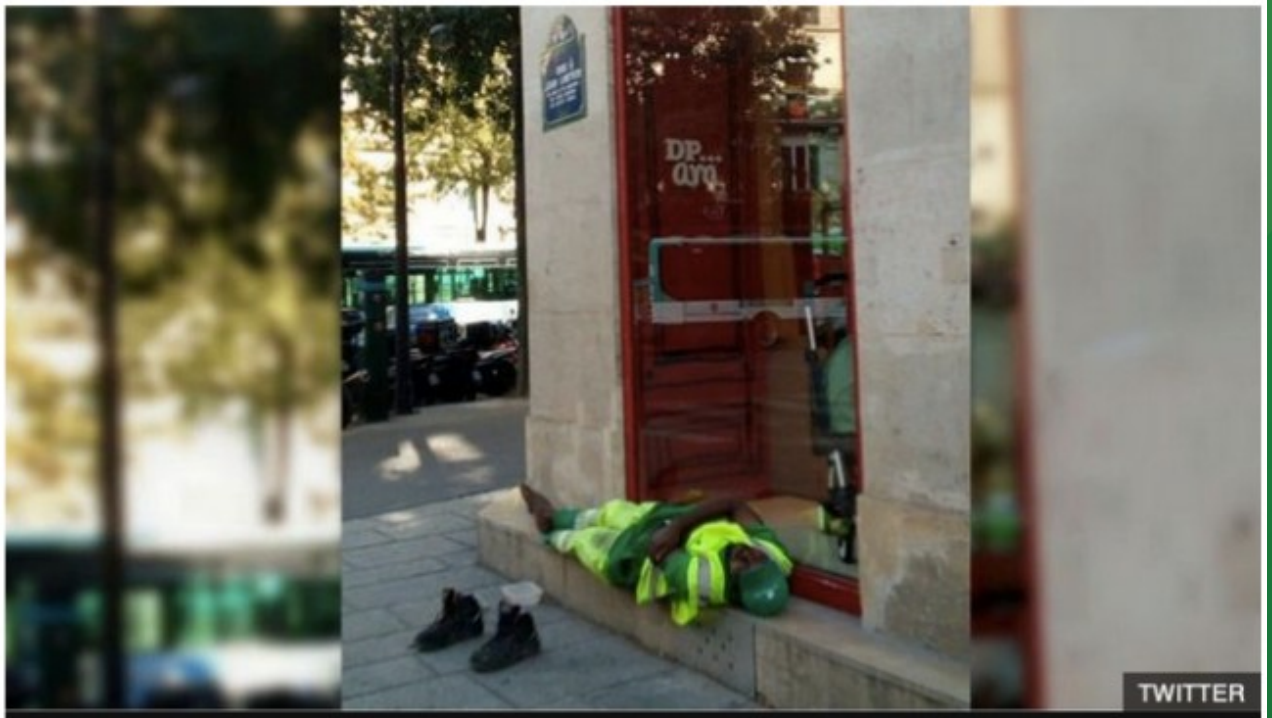
(Reppler, 2011)



Risks of self-exposure

Un éboueur licencié pour une photo sur Twitter

16 janvier 2020 Partager



Adama Cissé assoupi lors de "sa pause"

Risks of self-exposure

Un licenciement pour des propos tenus sur Facebook jugé légal

LEMONDE.FR avec AFP | 19.11.10 | 10h38 • Mis à jour le 19.11.10 | 10h59

Le licenciement de trois salariés d'une entreprise d'ingénierie de Boulogne-Billancourt (Hauts-de-Seine) pour dénigrement de leur hiérarchie sur Facebook a été jugé fondé vendredi par le conseil des prud'hommes, ont annoncé les avocats des deux parties.

Fin 2008, trois salariés avaient été licenciés pour *"incitation à la rébellion"* contre leur directrice des ressources humaines et faute grave. Les trois salariés étaient notamment accusés d'avoir incité leurs collègues à *"rendre la vie impossible"* à la hiérarchie de l'entreprise, lors de conversations tenues sur leurs comptes Facebook personnels, depuis leur domicile. Ces propos avaient été rapportés à la direction par l'un de leurs "amis Facebook".

Les salariés avaient plaidé le fait que ces échanges avaient un caractère strictement privé ; une vision contestée par la direction, qui a expliqué que Facebook est *"un site social ouvert"*.

The case of Facebook

- **Privacy policy** longer (5,830 words) than the American constitution
- 50 privacy settings, 170 options
- Modifications to the conditions of use sometimes inconsiderate

Mobile user tracking

- **Telecom operators** can easily **track users** (sending/receiving calls or messages) and **infer very intrusive information** (religion, political opinions, dating, favorite pubs, hobbies, etc.)

- Many **mobile applications collect personal information**: geo-location, identification number, contacts, etc.

Illustration

- Demo: Inspect What Your Location History Reveals About You

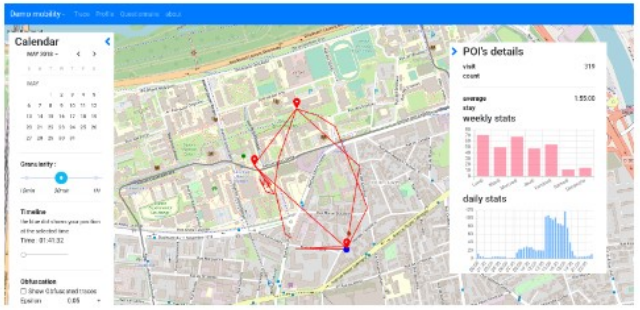


Figure 1: Screenshot of the demonstration's interface. Clicking on a POI yields attendance statistics.

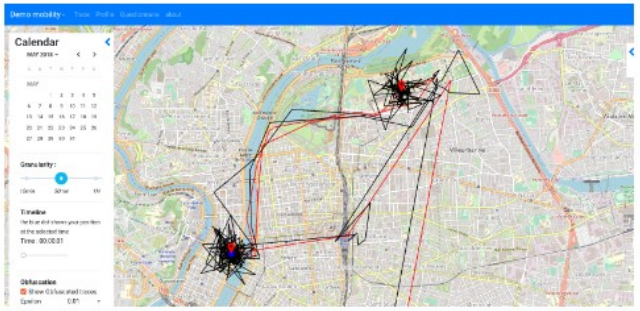
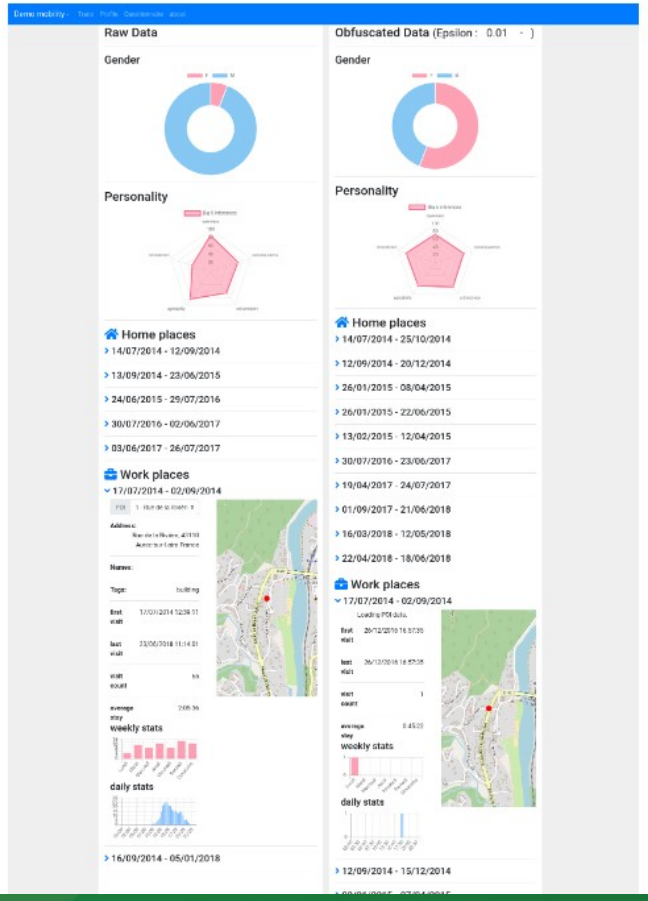
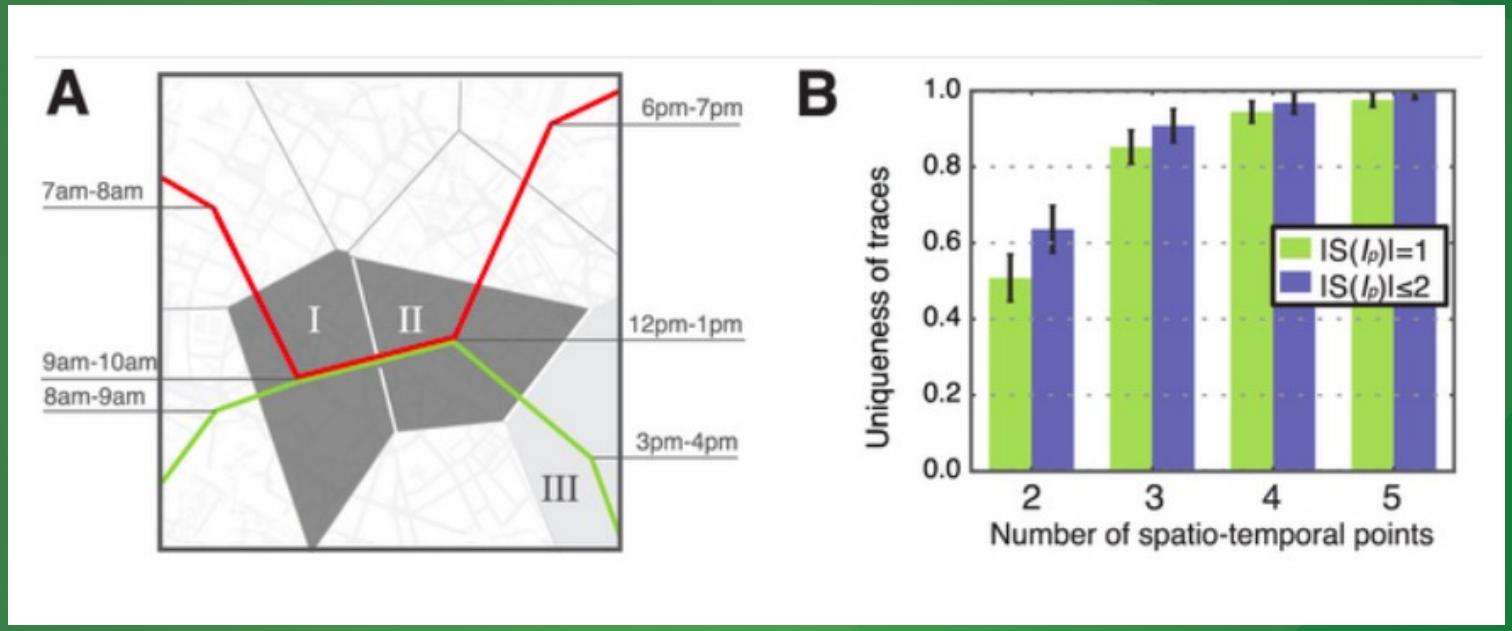


Figure 2: Illustration of the defense mechanism. Here the red lines show the initial traces, and the black lines represent the traces after the application of the noise.

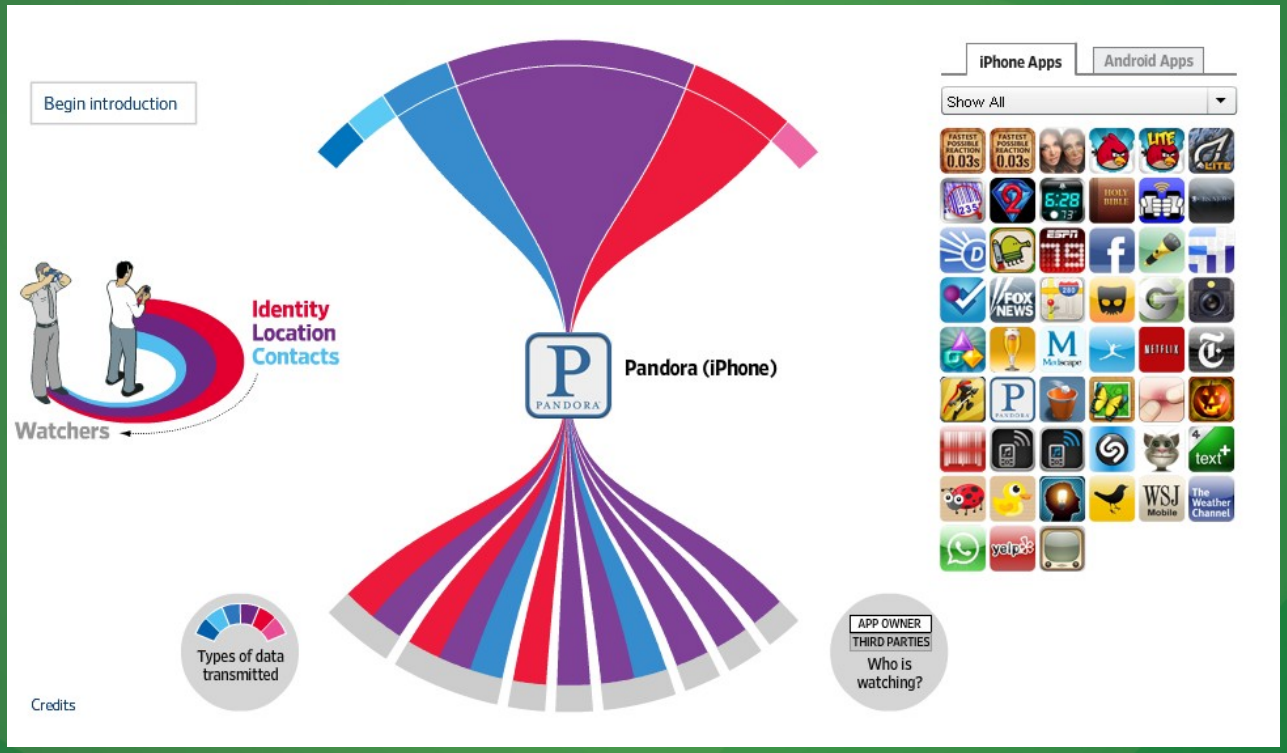


Unique in the Crowd [Nature 2013]



- 4 spatio-temporal data are enough to identify a user with a probability of 95%

MOBILITICS Project Cnil - Inria



Traced events

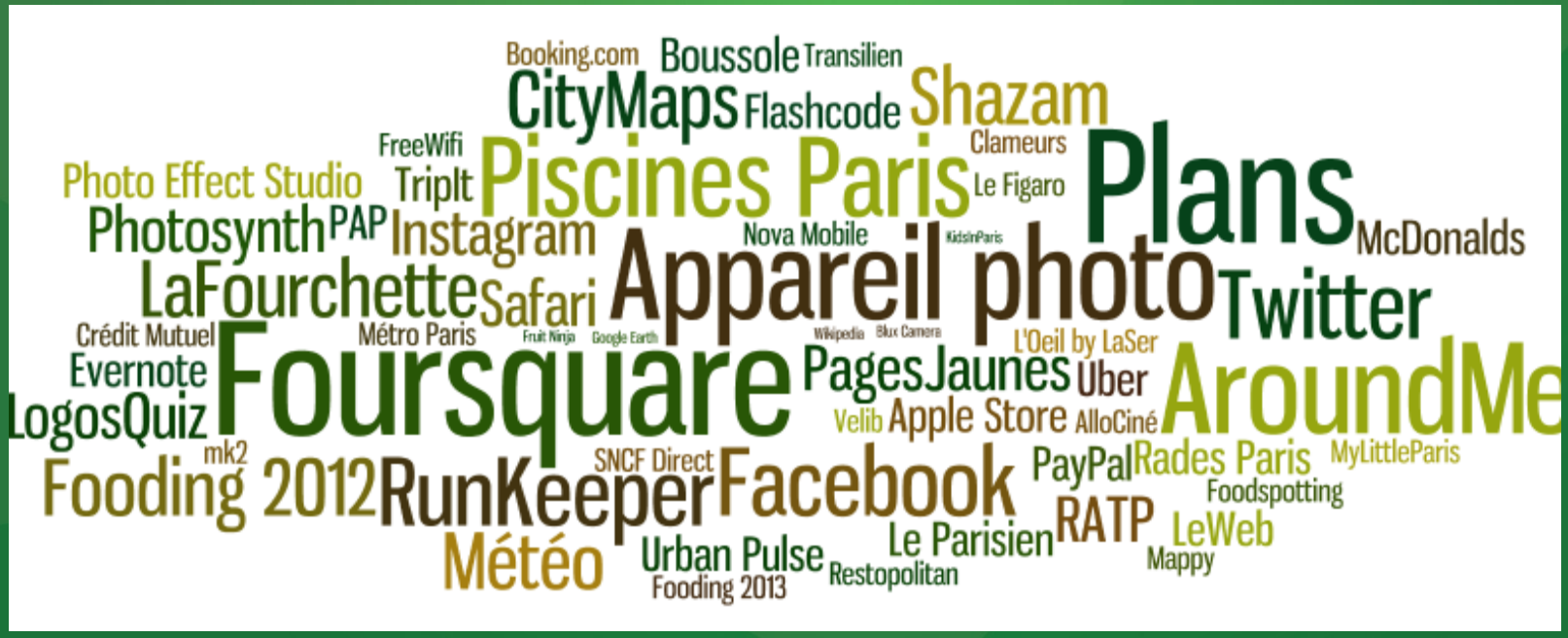
- Access to personal data: contact, geo-location, identification numbers, calendar, photos, videos, memos, etc.
- Incoming and outgoing calls
- Sending and receiving SMS, emails
- Internet browsing
- Taking photos

In vivo experiment

- 6 Cnil volunteers for 3 months
- 9 GB of data collected
- 7 million events analyzed
- 189 apps used

Geolocation: the queen of data

- 31% of apps used tried to access location
- 41,000 geolocation “events” in total, or an average of 76 events per day per volunteer



Invasive identifiers

- UDID: identifier integrated into the device by Apple and which cannot be modified or deleted by the user
- This UDID is very “in demand”: 87 applications out of the 189 accessed the UDID, or almost 50%

Why access the device name?

- 36 applications (>15%) accessed it
- What use?

Many invisible economic actors

- **Multiple “third party players” very present within the applications:**
  - Provide technical or monetization solutions to the developer (statistics tools, advertising, etc.).
  - Classic tracking players: Google, Criteo, Xiti, etc. but also emerging players, specialized in mobile (Flurry)
- The means of information and control by users, already limited in the “classic” web, are non-existent on mobile

Online advertising: simplified model

- **Advertiser**: jeweler, car manufacturer, swimming pool seller, etc.
- **Publisher** (advertising host): online newspaper, dictionary, etc.
- **Broker**: places advertisers' advertisements with publishers (doubleclick, outbrain, etc.)

Illustration

Annonceur

Broker
(doubleclick.com)

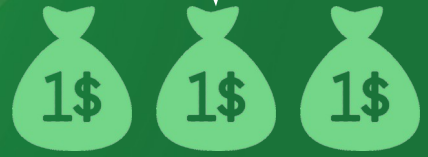
Publisher

Illustration

Annonceur

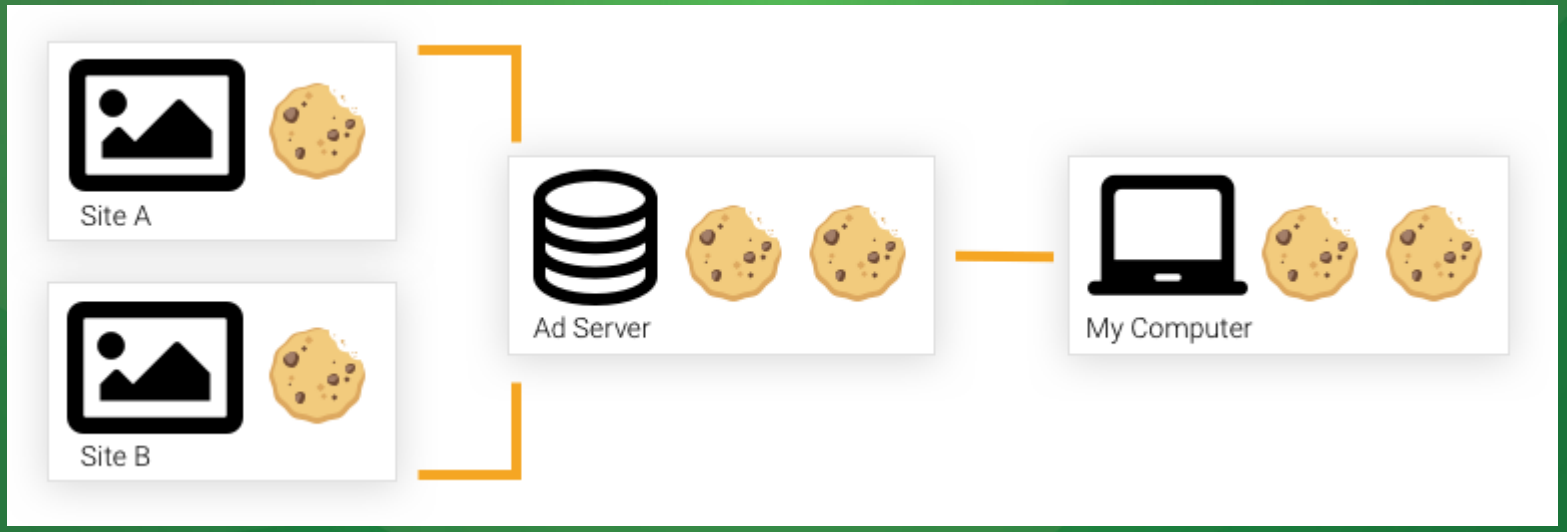
Publisher

Broker
(doubleclick.com)



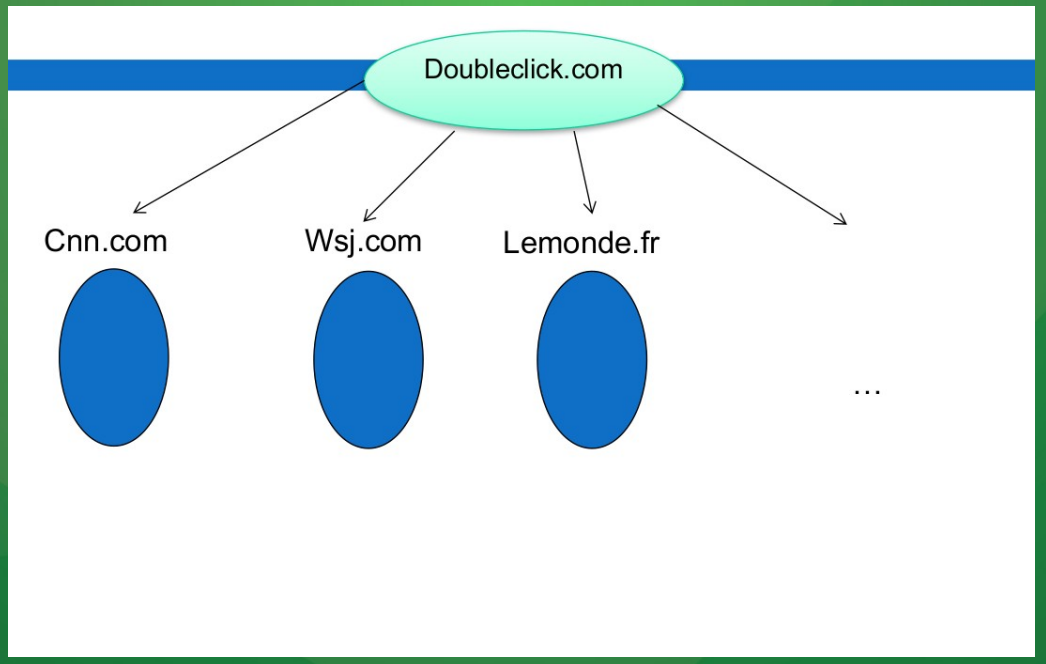
Third-Party Cookies

- Third-party cookies are set on your computer from domains other than the one that you're actually on right now
- Leverage usually a tracking pixel



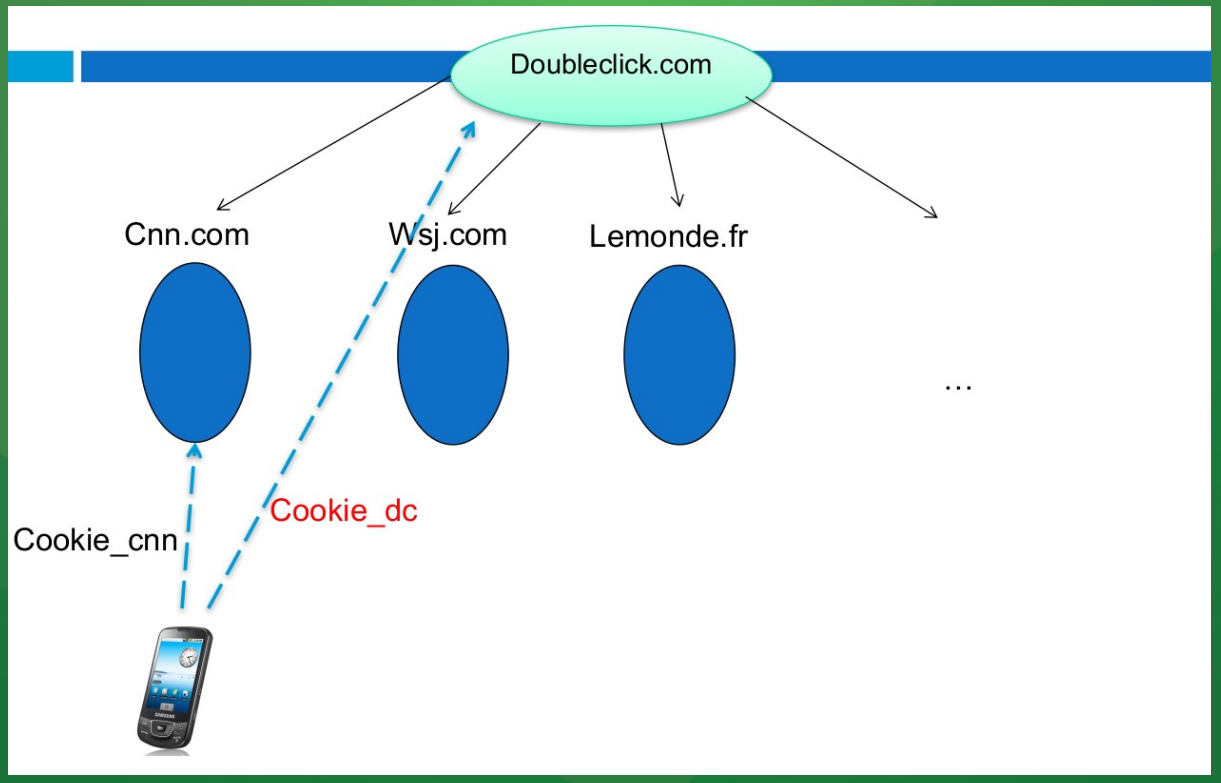
Third-Party Cookies

- Third-party cookies are set on your computer from domains other than the one that you're actually on right now.
- Leverage usually a tracking pixel



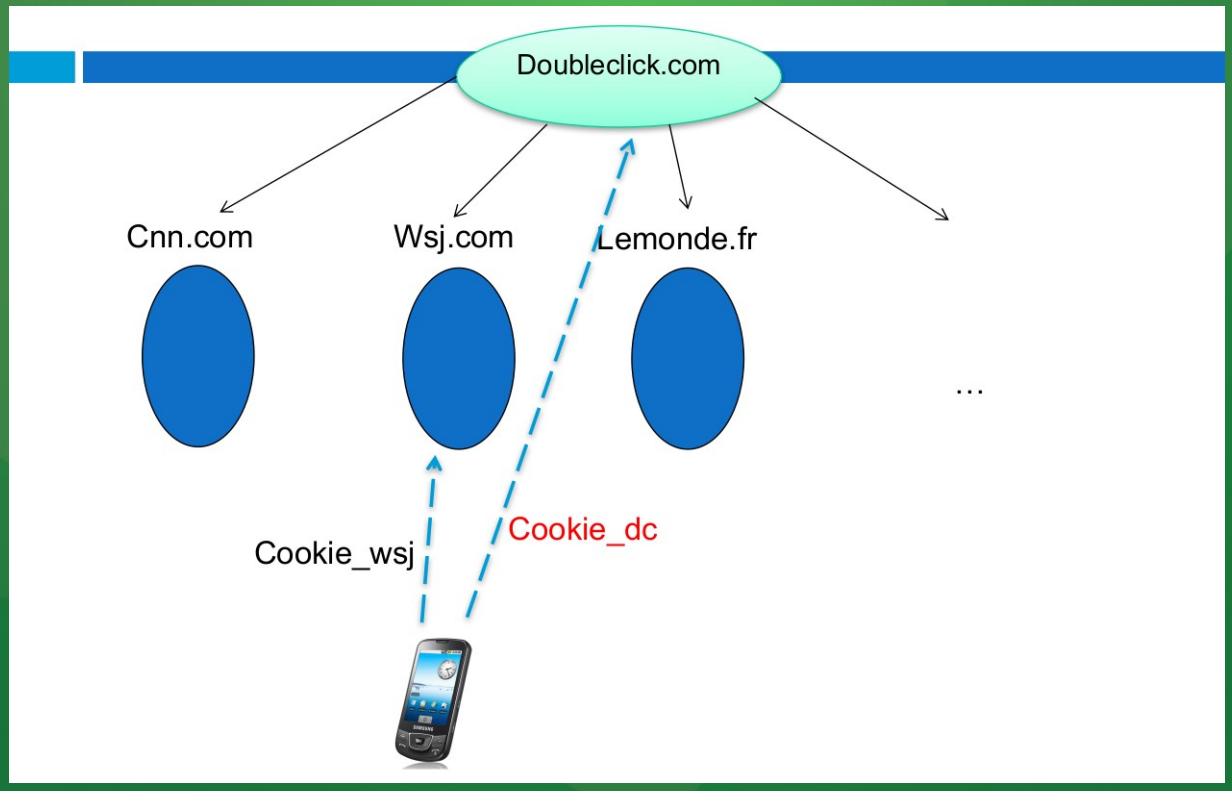
Third-Party Cookies

- Third-party cookies are set on your computer from domains other than the one that you're actually on right now.
- Leverage usually a tracking pixel



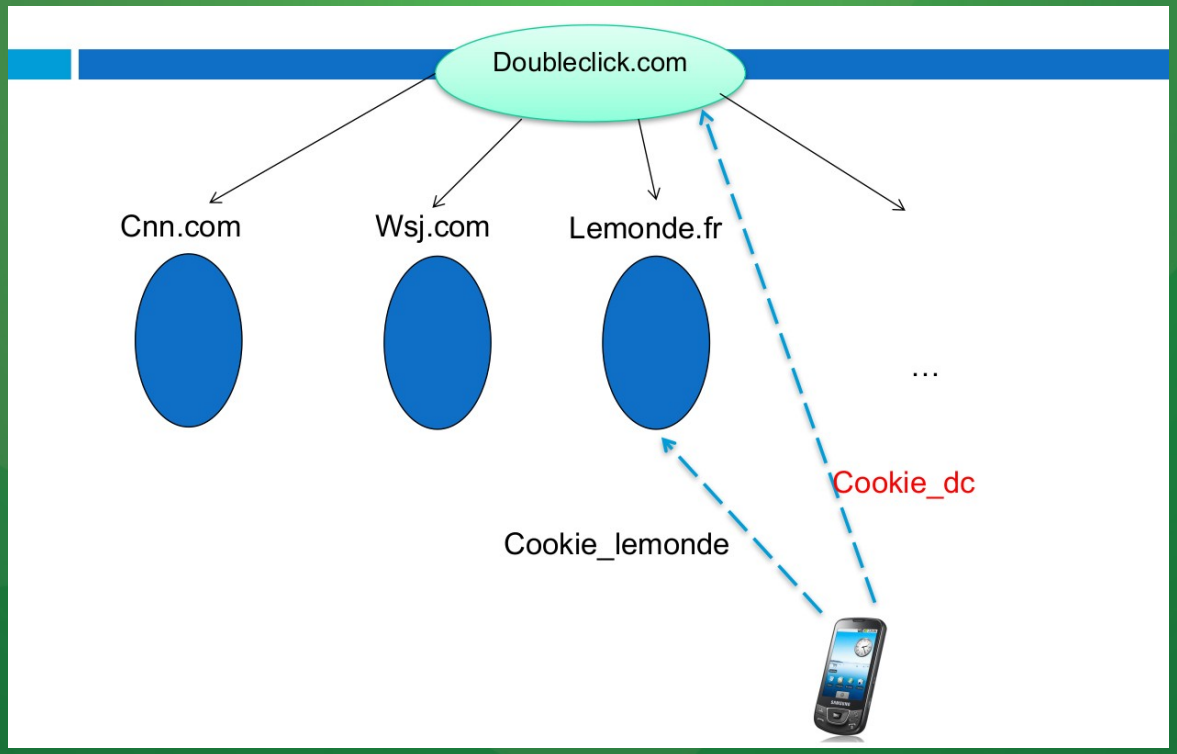
Third-Party Cookies

- Third-party cookies are set on your computer from domains other than the one that you're actually on right now.
- Leverage usually a tracking pixel



Third-Party Cookies

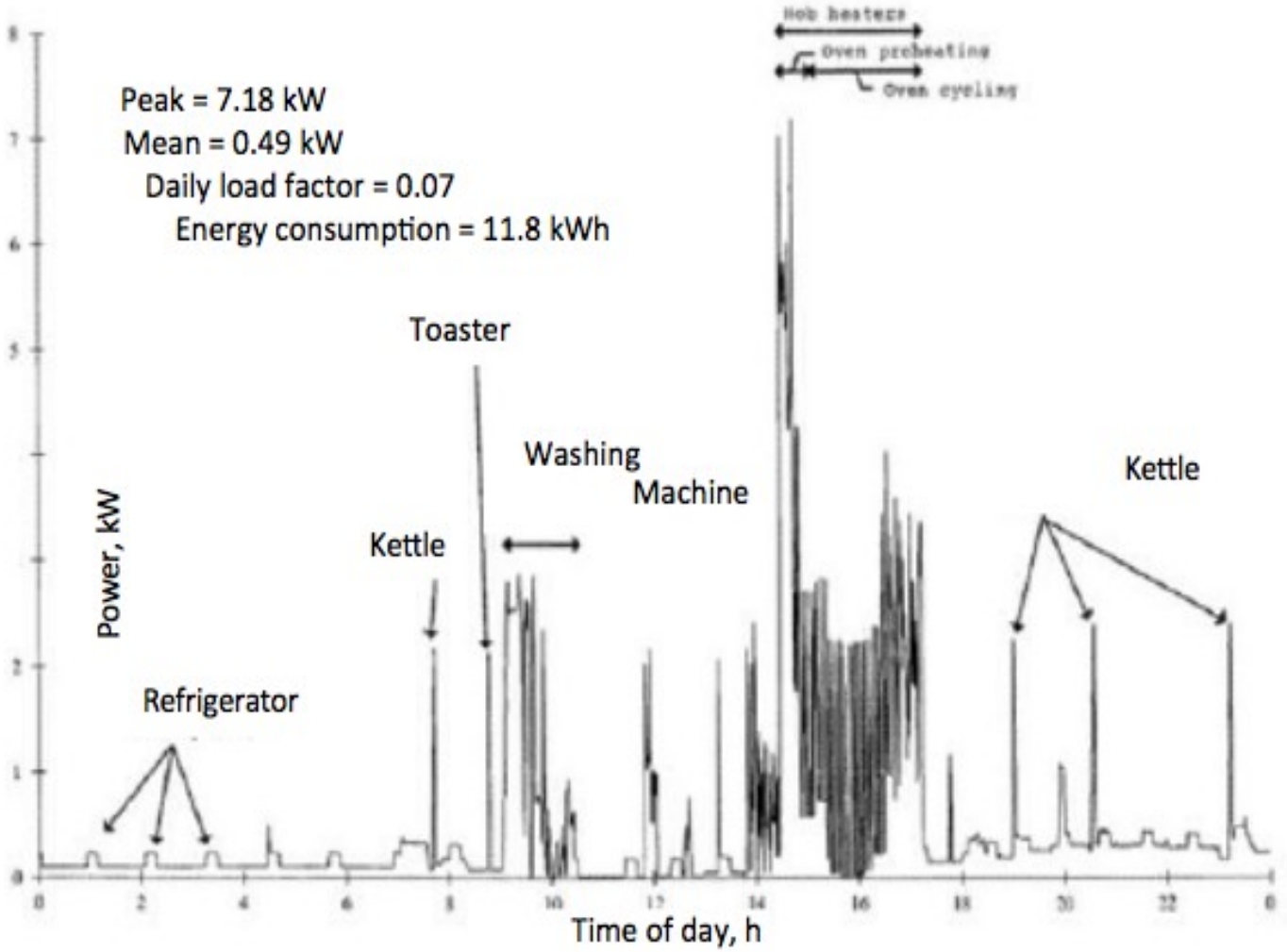
- Third-party cookies are set on your computer from domains other than the one that you're actually on right now.
- Leverage usually a tracking pixel



Profiling: consumption habits

- Target has identified 25 products that, when grouped together, may be associated with the likelihood of pregnancy
- Also estimate the due date!
- Sending appropriate shopping coupons for each phase of pregnancy

Profiling: smart meters



(TS//SI//NF) PRISM Collection Details

Current Providers

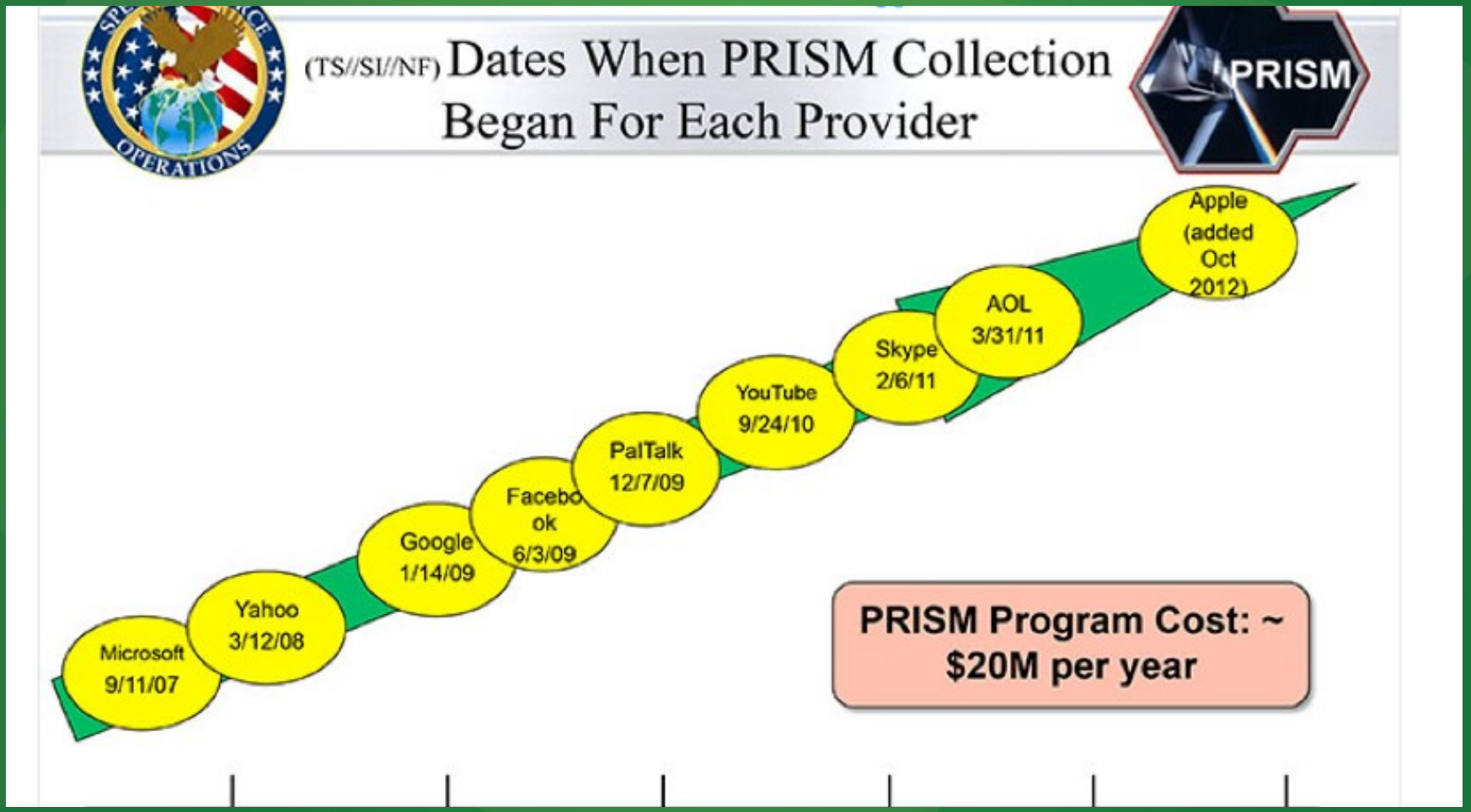
- Microsoft (Hotmail, etc.)
- Google
- Yahoo!
- Facebook
- PalTalk
- YouTube
- Skype
- AOL
- Apple

What Will You Receive in Collection (Surveillance and Stored Comms)?
It varies by provider. In general:

- E-mail
- Chat – video, voice
- Videos
- Photos
- Stored data
- VoIP
- File transfers
- Video Conferencing
- Notifications of target activity – logins, etc.
- Online Social Networking details
- **Special Requests**

Complete list and details on PRISM web page:
Go PRISMFAA

Gouvernemental Surveillance



(TS//SI//NF) FAA702 Operations

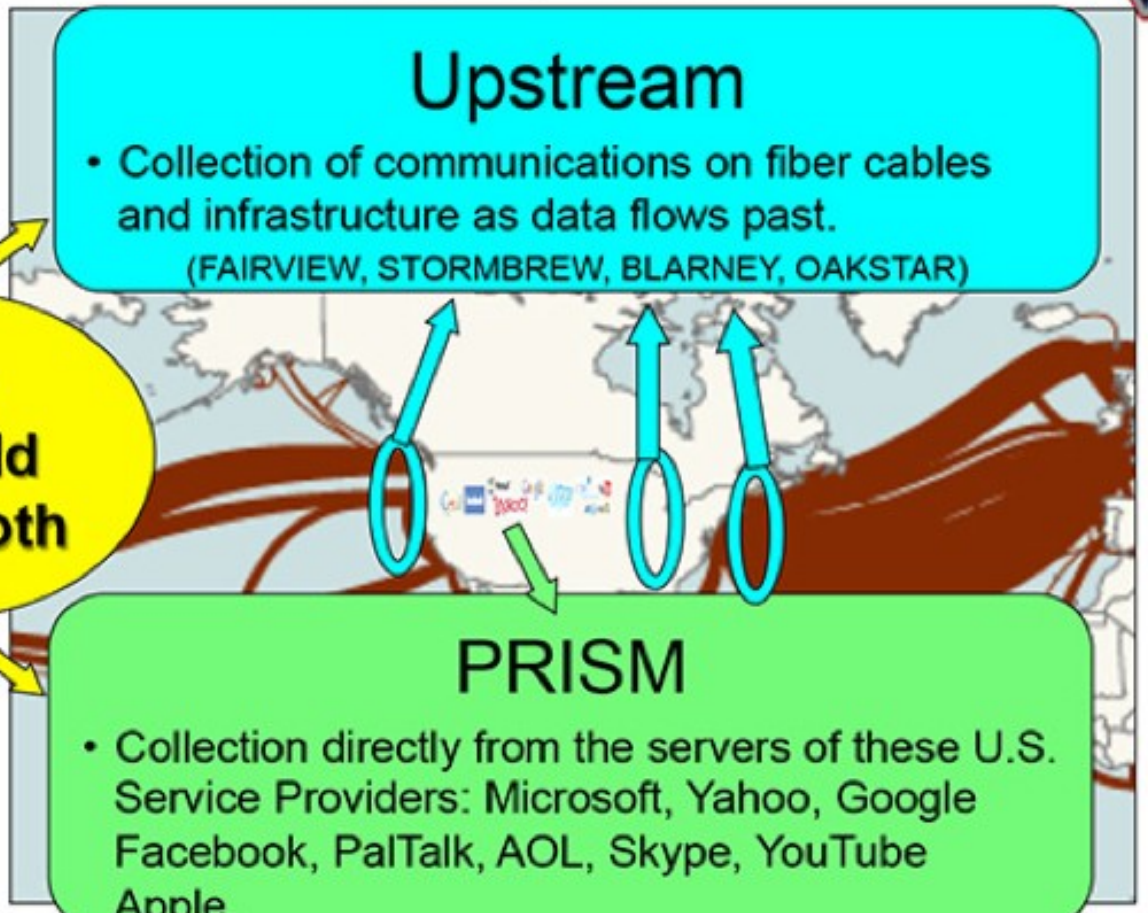
Two Types of Collection

Upstream

- Collection of communications on fiber cables and infrastructure as data flows past.

(FAIRVIEW, STORMBREW, BLARNEY, OAKSTAR)

You Should Use Both



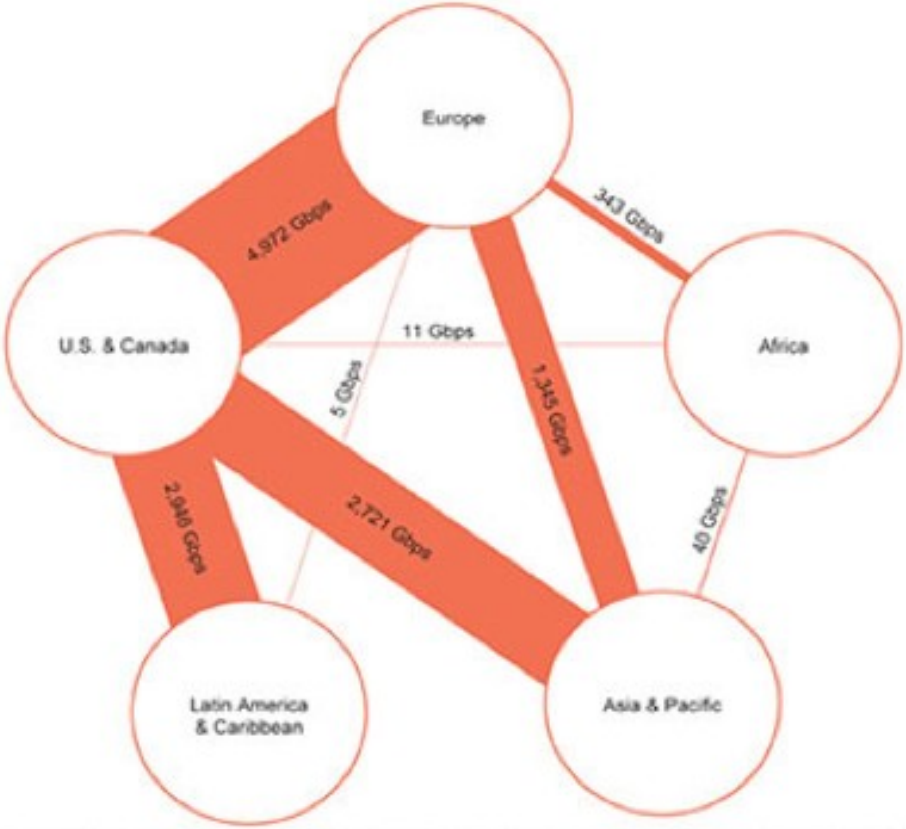
PRISM

- Collection directly from the servers of these U.S. Service Providers: Microsoft, Yahoo, Google Facebook, PalTalk, AOL, Skype, YouTube Apple.

(TS//SI//NF) Introduction

U.S. as World's Telecommunications Backbone

- Much of the world's communications flow through the U.S.
- A target's phone call, e-mail or chat will take the **cheapest** path, **not the physically most direct** path – you can't always predict the path.
- Your target's communications could easily be flowing into and through the U.S.



International Internet Regional Bandwidth Capacity in 2011

Source: Telegeography Research

Nothing to hide?

- **Eric Schmidt, PDG de Google (2009) :** «If you have something that you don’t want anyone to know, maybe you shouldn’t be doing it in the first place. »
- **Cardinal de Richelieu :** « Donnez-moi deux lignes de la main d’un homme, et j’y trouverai de quoi suffire à sa condamnation. »
- **Aleksandr Solzhenitsyn :** « Everyone is guilty of something or has something to conceal. All one has to do is look hard enough to find what it is. »

Nothing to hide?

- **Edward Snowden :** «Arguing that you don't care about the right to privacy because you have nothing to hide is no different than saying you don't care about free speech because you have nothing to say. »

Nothing to hide?

- Who is ready to **reveal everything to anyone**, under any circumstances?
  - Habits, behaviors (in public, at work, at home, etc.), interests, beliefs, preferences (political, personal, sexual, etc.), travel, health (condition, predispositions, etc.), social networks (dating, meetings, etc.)
- **Total transparency** (“life logging”)?

Fallacious conception of the notion of private life

- **Social value of private life is not reduced to the possibility of hiding reprehensible or shameful things:** the possession of an intimate, private area is necessary for the development of the personality, for the emancipation of each person.

Fallacious conception of the notion of private life

- **Social value of privacy:**
  - It does not only protect the individual (surroundings, domino effect, society)
  - Condition of diversity, existence of alternative lifestyles
- **Impact on democratic life itself:** formation and expression of opinions, self-censorship, conformism, tyranny of the majority, manipulation

Experimental study of self-censorship

- Experience using Wikipedia in the months before and after the revelations of Edward Snowden
- Choice of 48 articles related to terrorism and national security (bomb, suicide attack, nuclear, etc.)
- Comparison with a control group
- **Significant (20%) and lasting drop (inversion of the curve) in consultations after June 2013**
- *Chilling effects: Online surveillance and Wikipedia Use, J. W. Penney, Berkeley Technology Law Journal*

Dystopias

PIXELS • FACEBOOK

Bras de fer entre Facebook et NYU sur un projet de recherche sur le ciblage publicitaire politique

Le réseau social demande à l'université le retrait de son extension permettant de copier dans une base de données publique les publicités vues sur la plate-forme.

Le Monde avec AFP • Publié hier à 23h05, mis à jour à 00h27

Lecture 2 min.

Facebook a demandé à l'université de New York (NYU) de mettre fin à un projet de recherche sur ses pratiques en termes de ciblage des publicités politiques. Le réseau social estime que ce dernier enfreint son règlement en collectant les données des utilisateurs de la plate-forme.

Beyond collection: the use of data

- Multiple and often opaque uses:
  - **Personalization**: advertising, information, recommendations, insurance
  - **Facial recognition**: border crossing, access control to premises, surveillance of public space
  - **Generalized scoring**: banks, recruitment, teaching, fight against tax fraud, Chinese social score, etc.
  - **Predictive policing**: Predpol, Paved, Predvol, etc.
  - **Predictive justice**: COMPAS, ...

Privacy protection

- Why should you care?
- **What legal means?**
- What technical means?
- How to assess and reduce risks?

Protection de la vie privée

- Multiples facettes
  - En tension avec d’autres droits (sécurité nationale, sûreté publique, information, ...)
- Perceptions variées selon
  - les époques (évolutions des techniques, des comportements, ...)
  - les cultures (libertés/dignité, état/citoyen, droits fondamentaux/libre marché, ...)
  - les personnes (générations, milieux sociaux, etc.)

Règlement Général sur la Protection des Données (RGPD)

- Effectif depuis mai 2018
- A conduit à une révision de la loi Informatique et Libertés de 1978

France: loi Informatique et Libertés

- Loi du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés (modifiée en août 2004 et en mai 2018):
  - Définitions: données personnelles, traitements, responsables de traitements, ...
  - Principes: finalités, proportionnalité, durée de conservation, sécurité, droit d'information, d'accès, ...
  - Mise en oeuvre: Commission nationale de l'informatique et des libertés (CNIL)
  - Principes qui restent valables, renforcés par le RGPD

RGPD

- Champ d’application très large (données personnelles, activités, ...)
- Evolution d’un processus de nature administrative (obligations a priori) vers une démarche de responsabilisation (accountability) et d’analyse des risques
- Protection de la vie privée par construction (“privacy by design”)
- Importance accordée au contrôle des personnes sur leurs données
- Meilleure coordination et rôle accru des autorités de protection (certifications, anonymisation, sanctions, ...)
- Sanctions plus élevées : jusquà 20 M € ou 4% du chiffre d’affaires mondial
- Actions collectives : association à but non lucratif, avec ou sans mandat, y compris pour demander réparation

RGPD – données personnelles

Définitions

Aux fins du présent règlement, on entend par:

1) «données à caractère personnel», toute information se rapportant à une personne physique identifiée ou identifiable (ci-après dénommée «personne concernée»); est réputée être une «personne physique identifiable» une personne physique qui peut être identifiée, directement ou indirectement, notamment par référence à un identifiant, tel qu'un nom, un numéro d'identification, des données de localisation, un identifiant en ligne, ou à un ou plusieurs éléments spécifiques propres à son identité physique, physiologique, génétique, psychique, économique, culturelle ou sociale;

RGPD - pseudonymisation

Il y a lieu d'appliquer les principes de protection des données à toute information concernant une personne physique identifiée ou identifiable. Les données qui ont fait l'objet d'une pseudonymisation et qui pourraient être attribuées à une personne physique par le recours à des informations supplémentaires devraient être considérées comme des informations concernant une personne physique identifiable. Pour déterminer si une personne est

RGPD – territoire de l'EU

1. Le présent règlement s'applique au traitement des données à caractère personnel effectué dans le cadre des activités d'un établissement d'un responsable du traitement ou d'un sous-traitant sur le territoire de l'Union, que le traitement ait lieu ou non dans l'Union.

2. Le présent règlement s'applique au traitement des données à caractère personnel relatives à des personnes concernées qui se trouvent sur le territoire de l'Union par un responsable du traitement ou un sous-traitant qui n'est pas établi dans l'Union, lorsque les activités de traitement sont liées:

   a) à l'offre de biens ou de services à ces personnes dans l'Union, qu'un paiement soit exigé ou non desdites personnes; ou

   b) à l'observation du comportement de ces personnes, dans la mesure où il s'agit d'un comportement de l'Union européenne.

RGPD – responsabilité / concentement

2. Le responsable du traitement est responsable du respect des dispositions figurant au paragraphe 1 et est en mesure de démontrer que ces dispositions sont respectées (responsabilité).

1. Dans les cas où le traitement est fondé sur un consentement, le responsable du traitement est en mesure de démontrer que la personne concernée a donné son consentement au traitement de données à caractère personnel la concernant.

RGPD – privacy by design

Protection des données dès la conception et protection des données par défaut

Compte tenu de l'état des connaissances et des coûts de mise en œuvre et prenant en considération la nature, la portée, le contexte et les finalités du traitement ainsi que les risques, dont le degré de probabilité et la gravité varie, que présente le traitement pour les droits et libertés des personnes physiques, le responsable du traitement met en œuvre, tant au moment de la détermination des moyens du traitement qu'au moment du traitement lui-même, des mesures techniques et organisationnelles appropriées, telles que la pseudonymisation, qui sont destinées à donner effet aux principes de la protection des données, par exemple la minimisation des données, de façon effective et de manière à ce que le traitement comporte les garanties nécessaires, afin de répondre aux exigences du présent règlement et de protéger les droits de la personne concernée.

RGPD – analyse d’impact

Analyse d'impact relative à la protection des données

Lorsqu'un type de traitement, en particulier par le recours à de nouvelles technologies, et compte tenu de la nature, de la portée, du contexte et des finalités du traitement, est susceptible d'engendrer un risque élevé pour les droits et libertés des personnes physiques, le responsable du traitement effectue avant le traitement une analyse de l'impact des opérations de traitement envisagées sur la protection des données à caractère personnel. Une même analyse peut porter sur un ensemble d'opérations de traitement similaires qui présentent des risques élevés similaires.

RGPD – risque élevé

Le responsable du traitement consulte l'autorité de contrôle avant le traitement de données à caractère personnel lorsqu'une analyse d'impact relative à la protection des données, telle qu'elle est prévue à l'article 33, indique que le traitement présenterait un risque élevé si le responsable du traitement ne prenait pas de mesures pour atténuer le risque.

RGPD – violation / autorité

Notification à l'autorité de contrôle d'une violation de données à caractère personnel

En cas de violation de données à caractère personnel, le responsable du traitement en adresse notification à l'autorité de contrôle compétente conformément à l'article 51, dans les meilleurs délais et, si possible, 72 heures au plus tard après en avoir pris connaissance, à moins qu'il soit peu probable que la violation en question engendre un risque pour les droits et libertés des personnes physiques. Lorsqu'elle a lieu après ce délai de 72 heures, la notification comporte une motivation.

RGPD - violation / personne

Communication à la personne concernée d'une violation de données à caractère personnel

Lorsqu'une violation de données à caractère personnel est susceptible d'engendrer un risque élevé pour les droits et libertés d'une personne physique, le responsable du traitement communique la violation à la personne concernée dans les meilleurs délais.

RGPD – retirer son consentement

La personne concernée a le droit de retirer son consentement à tout moment. Le retrait du consentement ne compromet pas la licéité du traitement fondé sur le consentement préalablement donné. La personne concernée en est informée avant de donner son consentement. Il est aussi simple de retirer son consentement que de le donner.

RGPD – exporter les données

Les personnes concernées ont le droit de recevoir les données les concernant qu'elles ont communiquées à un responsable du traitement, dans un format structuré, couramment utilisé et lisible par machine, et de les transmettre à un autre responsable du traitement sans que le responsable du traitement auquel les données à caractère personnel ont été communiquées y fasse obstacle, lorsque:

a) le traitement est fondé sur le consentement en application de l'article 6, paragraphe 1, point a), ou de l'article 9, paragraphe 2, point a), ou sur un contrat en application de l'article 6, paragraphe 1, point b); et

b) le traitement est automatisé.

RGPD – traitement automatisé

Décision individuelle automatisée, y compris le profilage

La personne concernée a le droit de ne pas faire l'objet d'une décision résultant exclusivement d'un traitement automatisé, y compris le profilage, produisant des effets juridiques la concernant ou, de façon similaire, l'affectant de manière sensible.

CNIL

- **Rôles multiples**
  - **Règlementation :** autorisations de transferts hors UE, recherche médicale, labellisation
  - **Information, conseil :** gouvernement, citoyens, entreprises
  - **Protection :** plaintes, demandes de droit d’accès indirect, vérifications, notifications de violations de données
  - **Sanctions :** contrôles, mises en demeure, sanctions financières
  - **Prospective :** études, partenariats avec la recherche

CNIL

Reconnaissance faciale : pour un débat à la hauteur des enjeux

La reconnaissance faciale est de plus en plus présente dans le débat public aux niveaux national, européen et mondial. Cette technologie soulève en effet des questions inédites touchant à des choix de société. C'est pourquoi la CNIL a appelé, en 2018[9], à un débat démocratique sur ce sujet, ainsi que plus largement sur les nouveaux usages de la vidéo. En novembre 2019[10], elle a contribué au débat en présentant les éléments techniques, juridiques et éthiques qui doivent, selon elle, être pris en compte dans l'approche de cette question complexe.

CNIL

CNIL.

JE SUIS UN PARTICULIER | PROFESSIONNEL

Protéger les données personnelles, accompagner l'innovation, préserver les libertés individuelles

MES DÉMARCHES | THÉMATIQUES | TECHNOLOGIES | TEXTES OFFICIELS | LA CNIL

> La formation restreinte de la CNIL prononce une sanction de 50 millions d'euros à l'encontre de la société GOOGLE LLC

La formation restreinte de la CNIL prononce une sanction de 50 millions d'euros à l'encontre de la société GOOGLE LLC

21 janvier 2019

Le 21 janvier 2019, la formation restreinte de la CNIL a prononcé une sanction de 50 millions d'euros à l'encontre de la société GOOGLE LLC en application du RGPD pour manque de transparence, information insatisfaisante et absence de consentement valable pour la personnalisation de la publicité.

CNIL

- Autorité de protection très active, misant plus sur les volets conseil, l’accompagnement que les sanctions:
- Sanction maximale prononcée avant 2019: 150 000 €
- Janvier 2019: sanction de 50 M€ à l’encontre de Google (manquement aux obligations de transparence et d’information et manquement à l’obligation de disposer d’une base légale pour les traitements de personnalisation de la publicité liés à l’utilisation d’Android)
- La CNIL est une des autorités les plus importantes et influentes en Europe

Protection de vie privée

- Pourquoi s’en soucier
- Quels moyens juridiques ?
- **Quels moyens techniques ? (→ OT Security & Privacy)**
- Comment évaluer et réduire les risques ?