Vie privée, anonymat et données personnelles

Photo: Sean MacEntee/Flickr

guillaume . raschia @univ-nantes.fr

LS2N

Nantes Université

POLYTECH NANTES
Graduate school of engineering of the university of Nantes

dernière modification : 25/01/2024

Vie privée, anonymat et données personnelles

Photo: Sean MacEntee/Flickr

guillaume . raschia @univ-nantes.fr

LS2N

Nantes Université

POLYTECH NANTES
Graduate school of engineering of the university of Nantes

dernière modification : 25/01/2024

Vie privée

- Droit fondamental, affirmé en 1948 dans l'article 12 de la Déclaration universelle des droits de l'homme des Nations unies :

  « Nul ne sera l'objet d'immixtions arbitraires dans sa vie privée, sa famille, son domicile ou sa correspondance, ni d'atteintes à son honneur et à sa réputation. »

- Article 8 de la Convention européenne des droits de l'homme vs. liberté d'expression (article 10)…
- En France, l'article 9 du Code civil protège ce droit depuis la loi du 17 juillet 1970

Vie privée

- Droit fondamental, affirmé en 1948 dans l'article 12 de la Déclaration universelle des droits de l'homme des Nations unies :

  « Nul ne sera l'objet d'immixtions arbitraires dans sa vie privée, sa famille, son domicile ou sa correspondance, ni d'atteintes à son honneur et à sa réputation. »

- Article 8 de la Convention européenne des droits de l'homme vs. liberté d'expression (article 10)…
- En France, l'article 9 du Code civil protège ce droit depuis la loi du 17 juillet 1970

La sphère privée

- protection du domicile
- secret professionnel et médical
- protection de l'image
- protection de l'intimité
- protection de la correspondance : écoutes téléphoniques réglementées

Quel périmètre, quels enjeux, quelles menaces dans la société numérique ?

La sphère privée

- protection du domicile
- secret professionnel et médical
- protection de l'image
- protection de l'intimité
- protection de la correspondance : écoutes téléphoniques réglementées

Quel périmètre, quels enjeux, quelles menaces dans la société numérique ?

Anonymat

- Moyen radical de protéger la vie privée
- Portée pratique limitée : encadrement nécessaire de la propagation des informations personnelles

Quelles dispositions pour préserver l'anonymat ?

Quel cadre légal et réglementaire, quelles bonnes pratiques autour des données personnelles ?

Anonymat

- Moyen radical de protéger la vie privée
- Portée pratique limitée : encadrement nécessaire de la propagation des informations personnelles

Quelles dispositions pour préserver l'anonymat ?

Quel cadre légal et réglementaire, quelles bonnes pratiques autour des données personnelles ?

Au menu 

A. À propos des données personnelles
    1. Profil utilisateur et marché de la donnée
    2. Dispositifs de collecte :
        - Logs, traqueurs, RSx, IoT
    3. Surveillance de masse
    4. Risques et conséquences
B. Cadre législatif et réglementaire

Au menu 

A. À propos des données personnelles
    1. Profil utilisateur et marché de la donnée
    2. Dispositifs de collecte :
        - Logs, traqueurs, RSx, IoT
    3. Surveillance de masse
    4. Risques et conséquences
B. Cadre législatif et réglementaire

5

Profil individuel

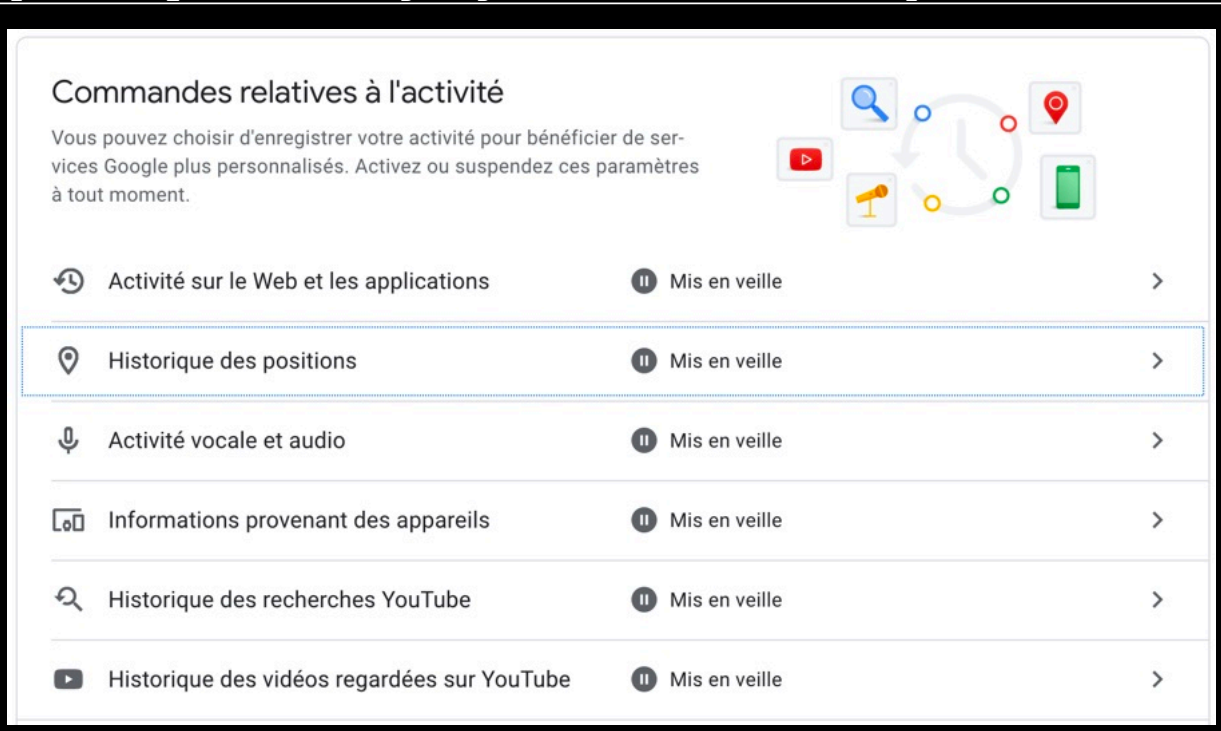
- *Traces* d'activité : préférence, consommation, évaluation, déplacement, communication…
- *Auxiliaires de collecte* : sites marchands, services en ligne, apps, titres de transport, cartes bancaires, cartes de fidélité, objets connectés, vidéosurveillance
- *Technologies de transmission* : Ethernet, wi-fi, puces sans contact (RFID), réseaux mobiles, GPS, Bluetooth

Profil individuel

- *Traces* d'activité : préférence, consommation, évaluation, déplacement, communication…
- *Auxiliaires de collecte* : sites marchands, services en ligne, apps, titres de transport, cartes bancaires, cartes de fidélité, objets connectés, vidéosurveillance
- *Technologies de transmission* : Ethernet, wi-fi, puces sans contact (RFID), réseaux mobiles, GPS, Bluetooth

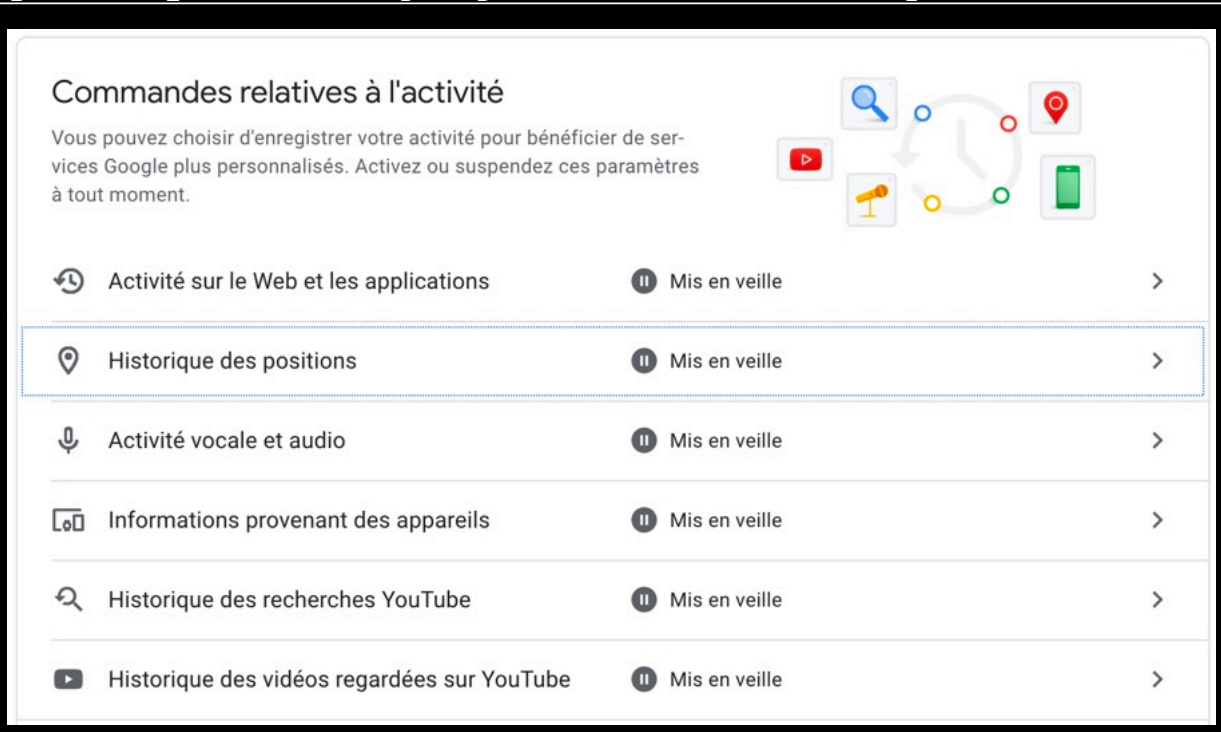
Moi, chez Google

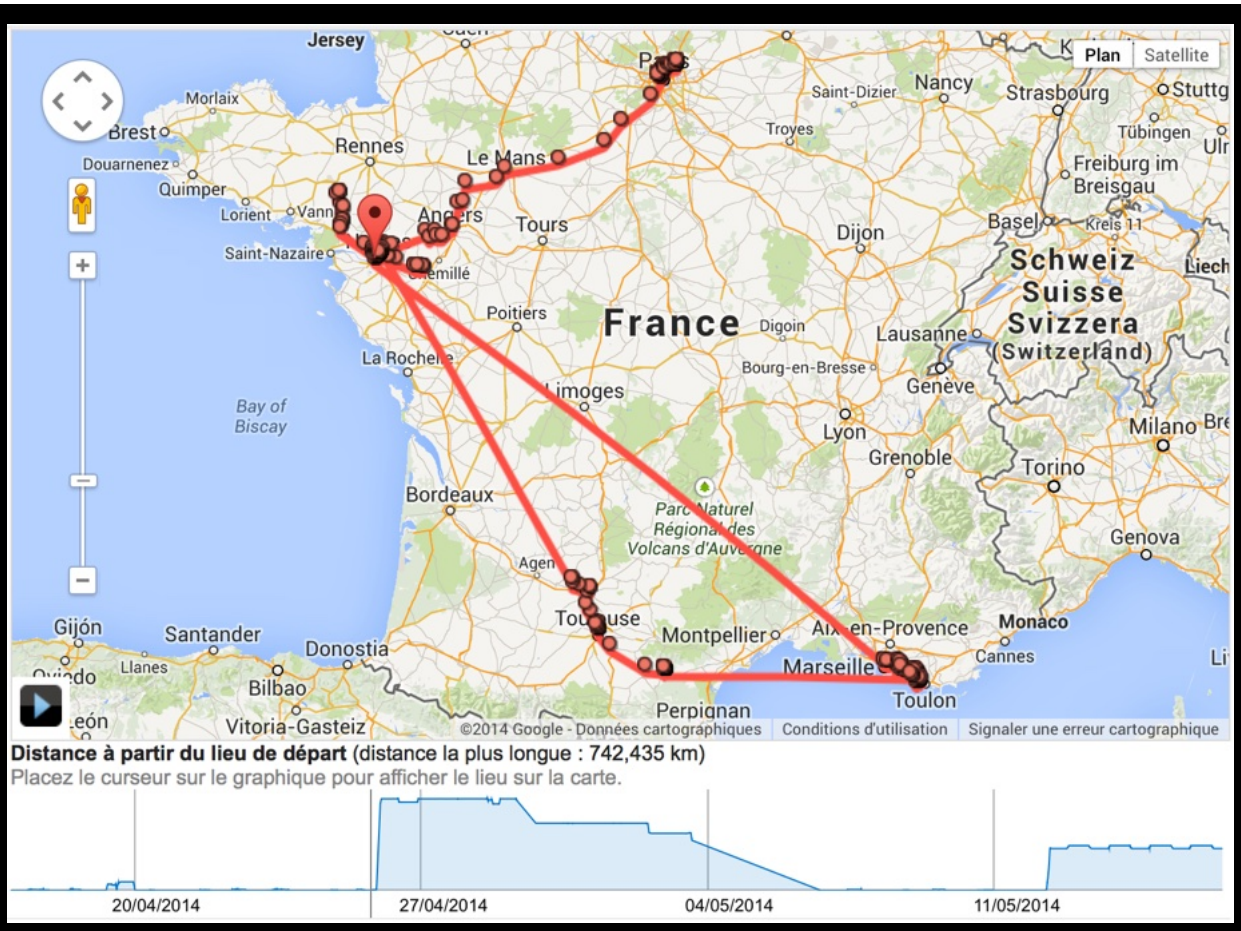
https://myaccount.google.com/data-and-personalization



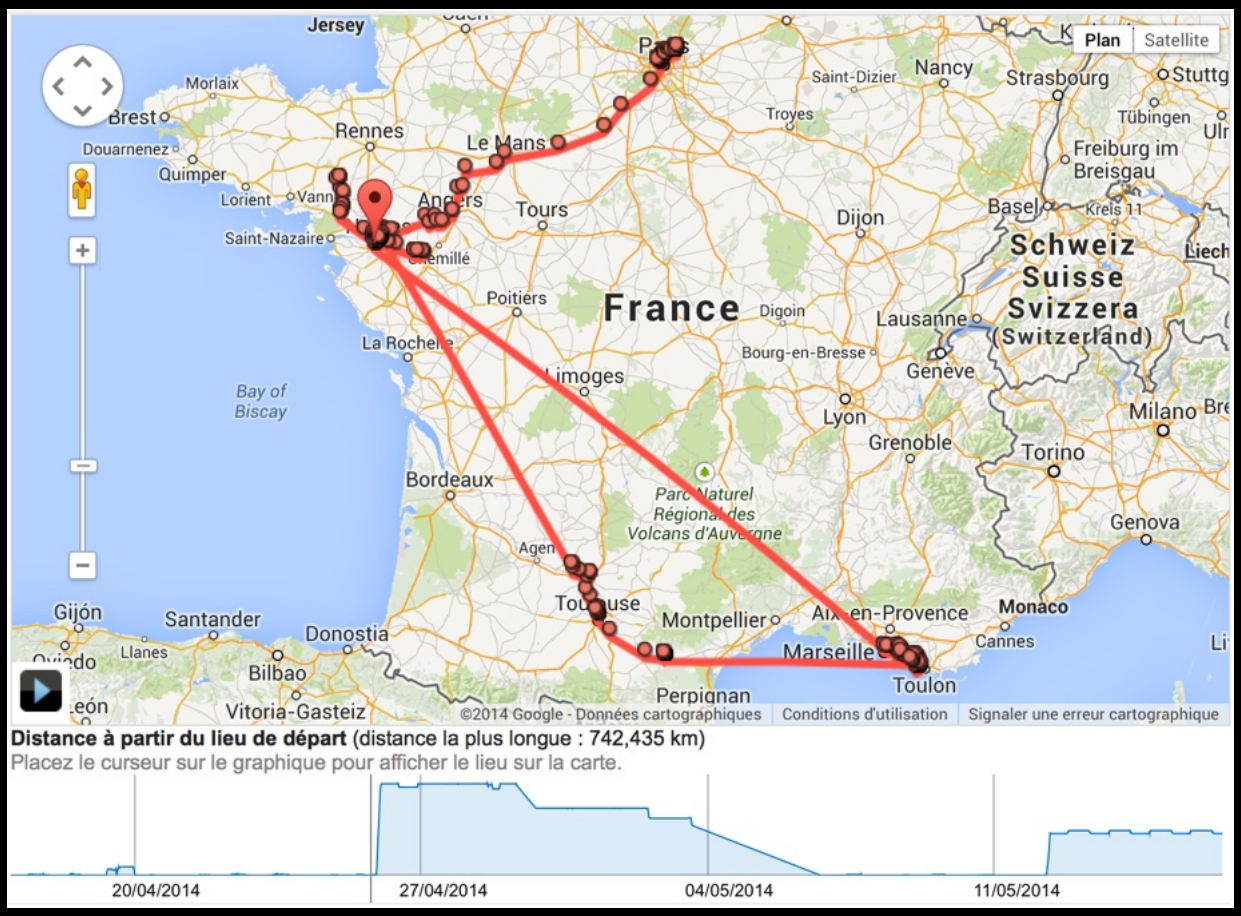
Moi, chez Google

https://myaccount.google.com/data-and-personalization

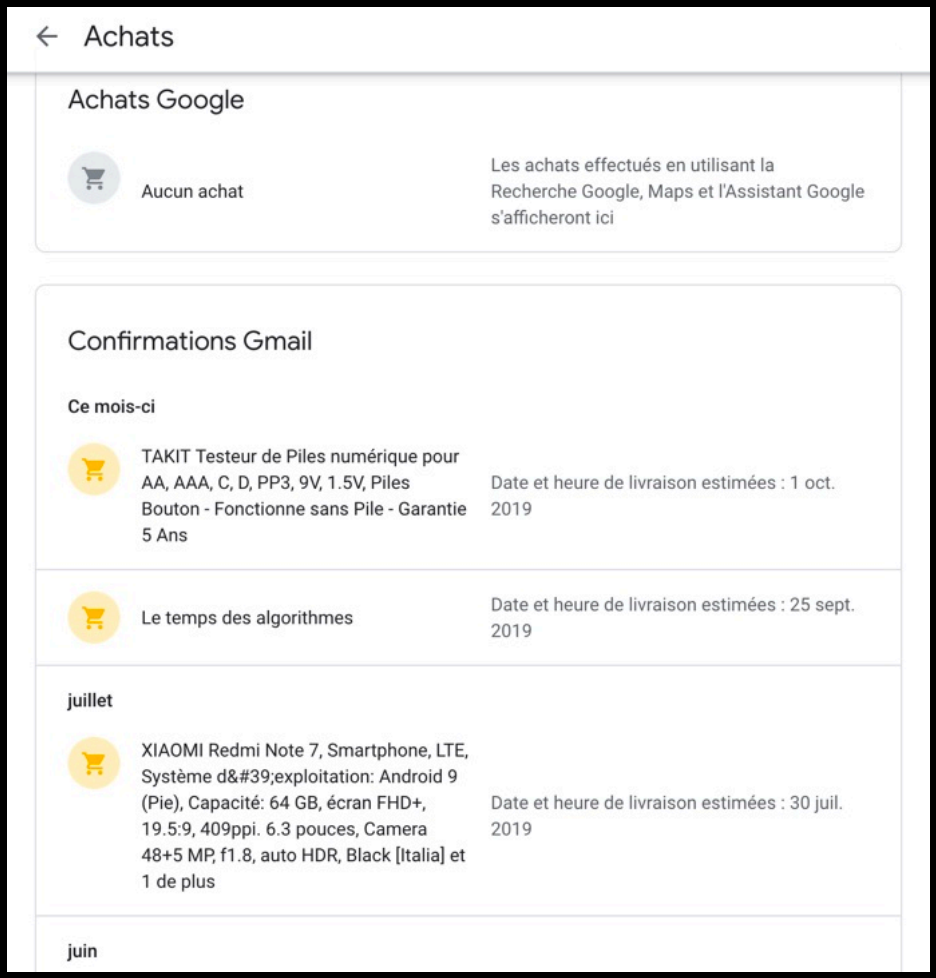




https://www.google.com/maps/timeline



https://www.google.com/maps/timeline



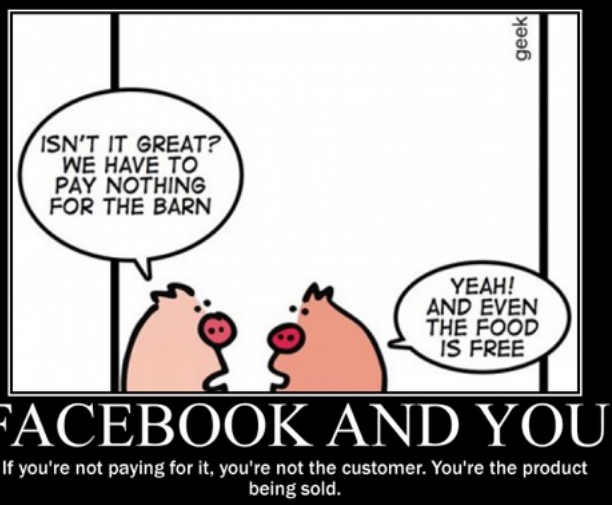
Achats
Achats Google
Aucun achat
Les achats effectués en utilisant la Recherche Google, Maps et l'Assistant Google s'afficheront ici
Confirmations Gmail
Ce mois-ci
TAKIT Testeur de Piles numérique pour AA, AAA, C, D, PP3, 9V, 1.5V, Piles Bouton - Fonctionne sans Pile - Garantie 5 Ans
Date et heure de livraison estimées : 1 oct. 2019
Le temps des algorithmes
Date et heure de livraison estimées : 25 sept. 2019
juillet
XIAOMI Redmi Note 7, Smartphone, LTE, Système d'exploitation: Android 9 (Pie), Capacité: 64 GB, écran FHD+, 19.5:9, 409ppi. 6.3 pouces, Camera 48+5 MP, f1.8, auto HDR, Black [Italia] et 1 de plus
Date et heure de livraison estimées : 30 juil. 2019
juin

9

Échelle mondiale

- Cloud : données et services externalisés
- Maxime geek : « *Si c'est grat8, c'est moi le prod8 ! »*

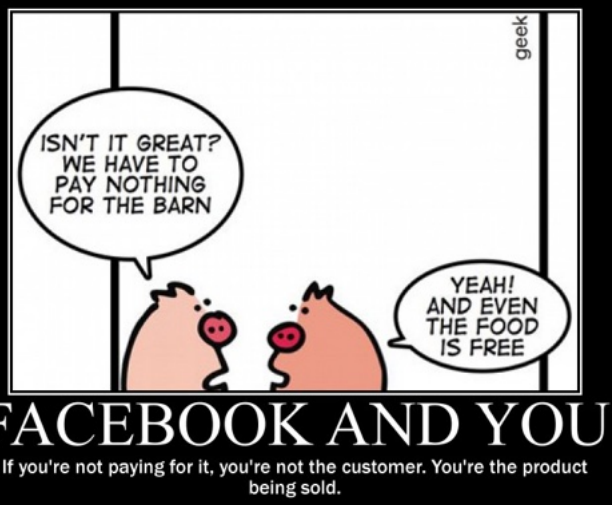
The Big Five: GAFAM



Échelle mondiale

- Cloud : données et services externalisés
- Maxime geek : « *Si c'est grat8, c'est moi le prod8 ! »*

The Big Five: GAFAM

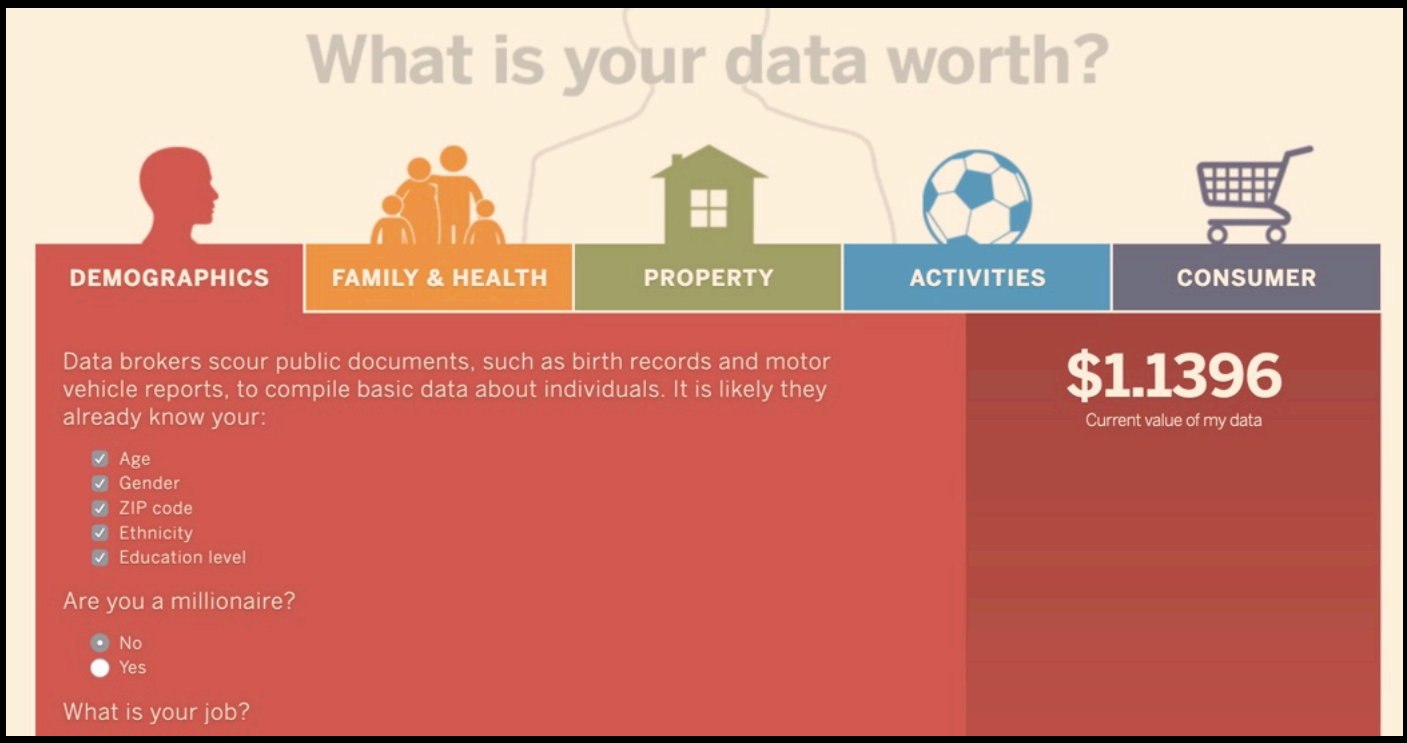


Marché des données

- Data Act (27 juin 2023) : Marché unique Européen
- ~200B€, 1.7Mo/sec/pers en 2020 + GPS + biométrie
- Entreprises spécialisées : Acxiom
- Places de marché : https://www.dawex.com/
- Métier : courtier en données (*data broker*)
- WeWard 0.2€ / 8km : nouveau modèle
  - vendre ou louer soi-même ses données !
  - Embleema, Swagbucks

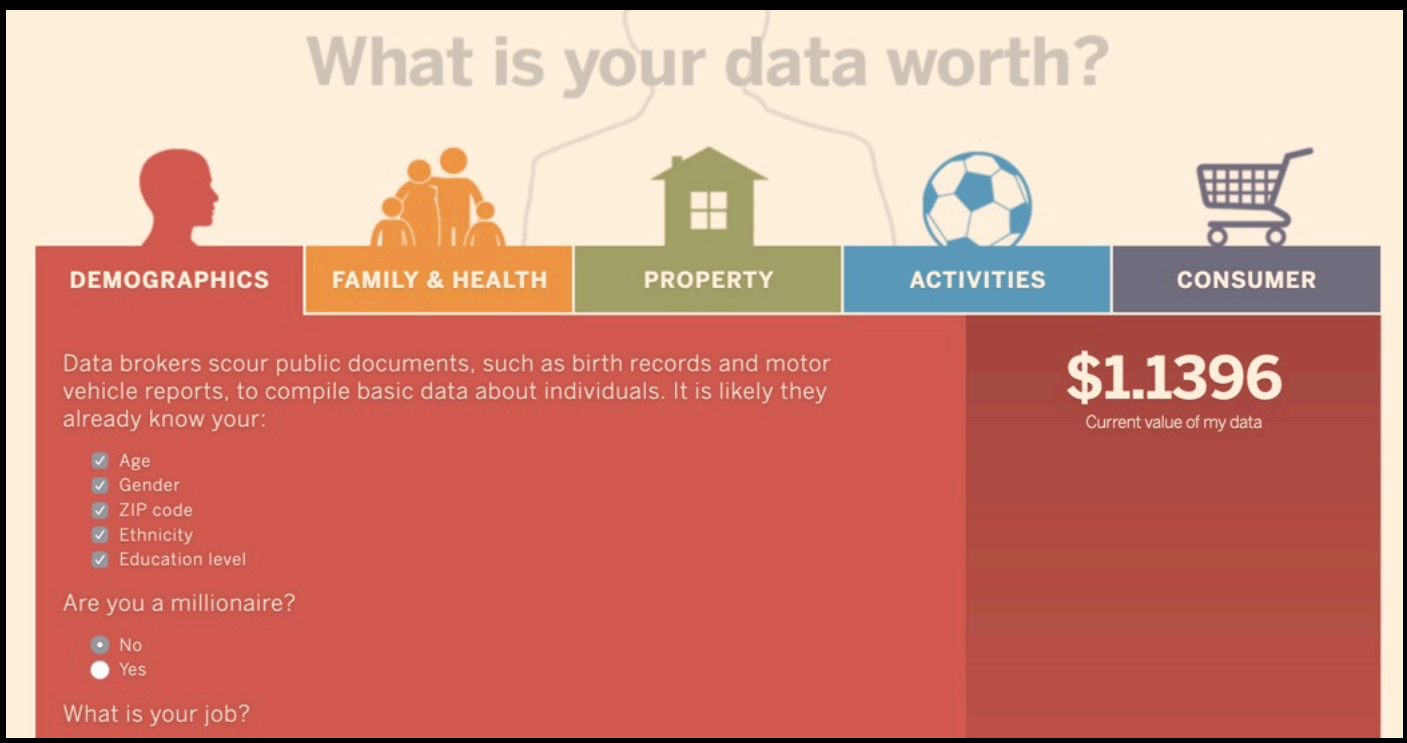
Marché des données

- Data Act (27 juin 2023) : Marché unique Européen
- ~200B€, 1.7Mo/sec/pers en 2020 + GPS + biométrie
- Entreprises spécialisées : Acxiom
- Places de marché : https://www.dawex.com/
- Métier : courtier en données (*data broker*)
- WeWard 0.2€ / 8km : nouveau modèle
  - vendre ou louer soi-même ses données !
  - Embleema, Swagbucks



Source : Financial Times (2017)

https://ig.ft.com/how-much-is-your-personal-data-worth/



Source : Financial Times (2017)

https://ig.ft.com/how-much-is-your-personal-data-worth/

Les raisons de la collecte

- Gain économique
- Sécurité
  - vidéosurveillance
- Exposition de soi
  - réseaux sociaux
- Obligation légale et réglementaire
  - état civil, dossier médical, impôts, recensement, *etc.*
- Contribution active à dessein
  - épidémie de grippe H1N1

Les raisons de la collecte

- Gain économique
- Sécurité
  - vidéosurveillance
- Exposition de soi
  - réseaux sociaux
- Obligation légale et réglementaire
  - état civil, dossier médical, impôts, recensement, *etc.*
- Contribution active à dessein
  - épidémie de grippe H1N1

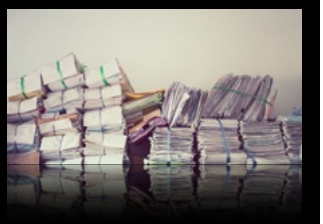
Où en sommes-nous ?

A. À propos des données personnelles
    1. Profil utilisateur et marché de la donnée
    2. Dispositifs de collecte :
        - ⋆ Logs, traqueurs, RSx, IoT
    3. Surveillance de masse
    4. Risques et conséquences
B. Cadre législatif et réglementaire

Où en sommes-nous ?

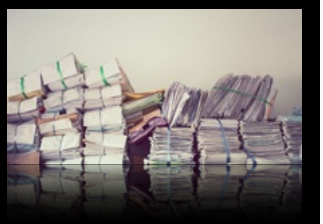
A. À propos des données personnelles
    1. Profil utilisateur et marché de la donnée
    2. Dispositifs de collecte :
        - ⋆ Logs, traqueurs, RSx, IoT
    3. Surveillance de masse
    4. Risques et conséquences
B. Cadre législatif et réglementaire

Logs de connexion



- Pourquoi ? Besoin de sécurité
- Enregistrement de l'activité dans un journal
  - à minima l'accès (id et horodatage in/out), parfois l'opération et la donnée consultée
- Activité (mal) encadrée par la loi
- Cadre juridique de la conservation
  - la CNIL recommande de 6 mois à 1 an

Logs de connexion



- Pourquoi ? Besoin de sécurité
- Enregistrement de l'activité dans un journal
  - à minima l'accès (id et horodatage in/out), parfois l'opération et la donnée consultée
- Activité (mal) encadrée par la loi
- Cadre juridique de la conservation
  - la CNIL recommande de 6 mois à 1 an

FAI, hébergeurs et opérateurs Telecom

- Article L34-1 du Code des Postes et des Télécommunications
  - effacer ou rendre anonyme les données de com.
  - besoin de sécurité : 1 an de conservation !
  - MAIS, la CJUE (arrêt du 06 octobre 2020) interdit la conservation généralisée et indifférenciée
  - MAIS décret du 20 octobre 2021 !
    - « menace grave et actuelle contre la sécurité nationale »

FAI, hébergeurs et opérateurs Telecom

- Article L34-1 du Code des Postes et des Télécommunications
  - effacer ou rendre anonyme les données de com.
  - besoin de sécurité : 1 an de conservation !
  - MAIS, la CJUE (arrêt du 06 octobre 2020) interdit la conservation généralisée et indifférenciée
  - MAIS décret du 20 octobre 2021 !
    - « menace grave et actuelle contre la sécurité nationale »

Logs de requêtes

- Enregistrement dans un registre :

  (date, @ip, termes de la requête[, id-compte])

- Durée de rétention
  - Depuis 2010 : 9 mois pour Google, 6 mois pour Microsoft et 3 mois pour Yahoo
  - la CNIL et le G29 recommandent <6 mois

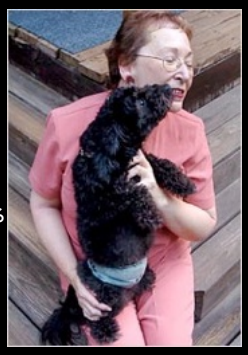
Logs de requêtes

- Enregistrement dans un registre :

  (date, @ip, termes de la requête[, id-compte])

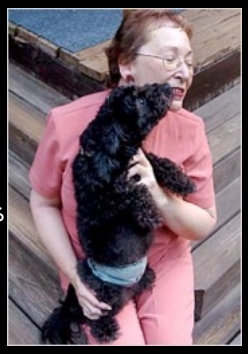
- Durée de rétention
  - Depuis 2010 : 9 mois pour Google, 6 mois pour Microsoft et 3 mois pour Yahoo
  - la CNIL et le G29 recommandent <6 mois

Révélation d'identité



- Dis-moi ce que tu recherches…
- Concours AOL (2006) Source : TNYT. *A Face Is Exposed for AOL Searcher No. 4417749*, 2006
  - 20M requêtes de 658.000 internautes, @IP hachée
  - 3 mois de requêtes pour l'individu #4.417.749 :
    - "chien qui fait pipi partout"
    - "taxe foncière de Harrisburg, Virginie"
    - "homme célibataire de 60 ans"
  - #4.417.749 = Thelma Arnold, veuve de 62 ans, Lilburn, Georgia
  - Fichier et analyses disponibles ici : https://searchids.com

Révélation d'identité



- Dis-moi ce que tu recherches…
- Concours AOL (2006) Source : TNYT. *A Face Is Exposed for AOL Searcher No. 4417749*, 2006
  - 20M requêtes de 658.000 internautes, @IP hachée
  - 3 mois de requêtes pour l'individu #4.417.749 :
    - "chien qui fait pipi partout"
    - "taxe foncière de Harrisburg, Virginie"
    - "homme célibataire de 60 ans"
  - #4.417.749 = Thelma Arnold, veuve de 62 ans, Lilburn, Georgia
  - Fichier et analyses disponibles ici : https://searchids.com

Les bons élèves

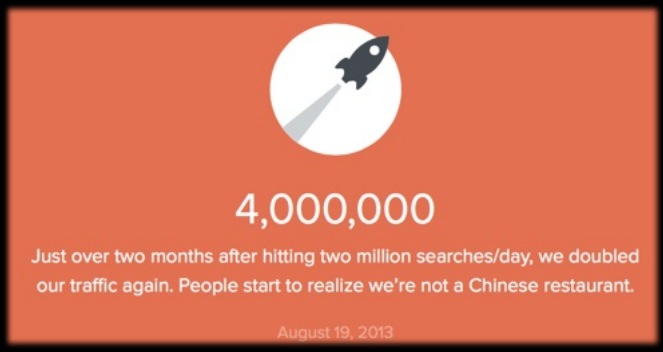
www.qwant.com

2.7B requêtes en 2020

repère : Google = 3.7B requêtes / jour (90%)…

duckduckgo.com

3B requêtes/mois en 2023



Les bons élèves

www.qwant.com

2.7B requêtes en 2020

repère : Google = 3.7B requêtes / jour (90%)…

DuckDuckGo

duckduckgo.com

3B requêtes/mois en 2023

4,000,000
Just over two months after hitting two million searches/day, we doubled our traffic again. People start to realize we're not a Chinese restaurant.
August 19, 2013

Bonnes pratiques du Web

1. multiplier les identités numériques, utiliser des mails jetables (*temp-mail.org*, *yopmail.com*,…)
2. paramétrer le navigateur : navigation privée
3. limiter l'utilisation des cookies
4. utiliser un moteur de recherche anonyme
5. souscrire un service mail payant et sécurisé (*mailfence*, *tuta*, *protonmail*)
6. utiliser un service VPN (pour Réseau Privé Virtuel)

Pour plus d'infos, se référer aux recommandations de la CNIL

Bonnes pratiques du Web

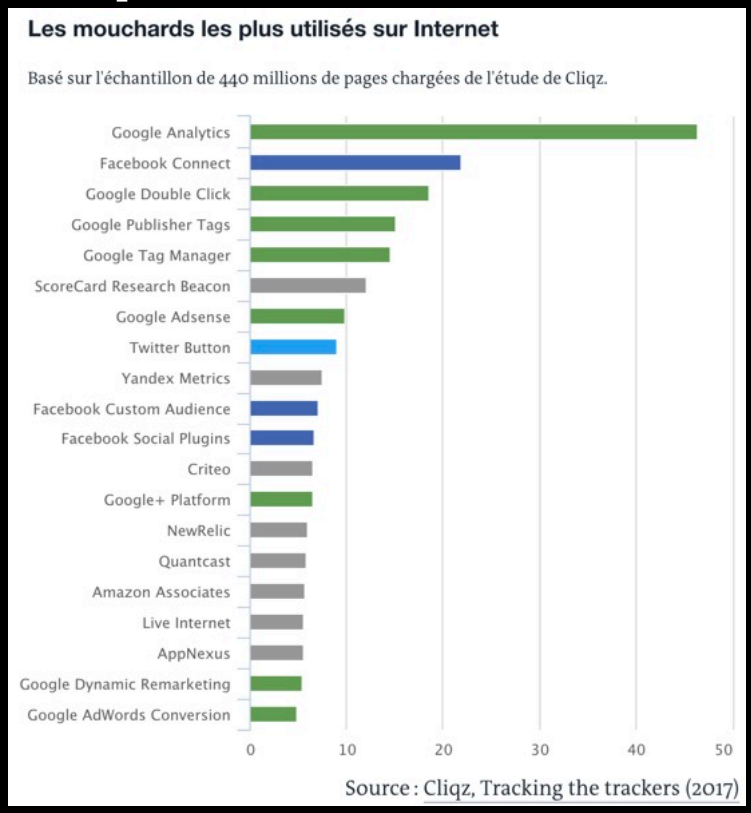
1. multiplier les identités numériques, utiliser des mails jetables (*temp-mail.org*, *yopmail.com*,…)
2. paramétrer le navigateur : navigation privée
3. limiter l'utilisation des cookies
4. utiliser un moteur de recherche anonyme
5. souscrire un service mail payant et sécurisé (*mailfence*, *tuta*, *protonmail*)
6. utiliser un service VPN (pour Réseau Privé Virtuel)

Pour plus d'infos, se référer aux recommandations de la CNIL

Les traqueurs

Étude menée par Cliqz en 2017 :

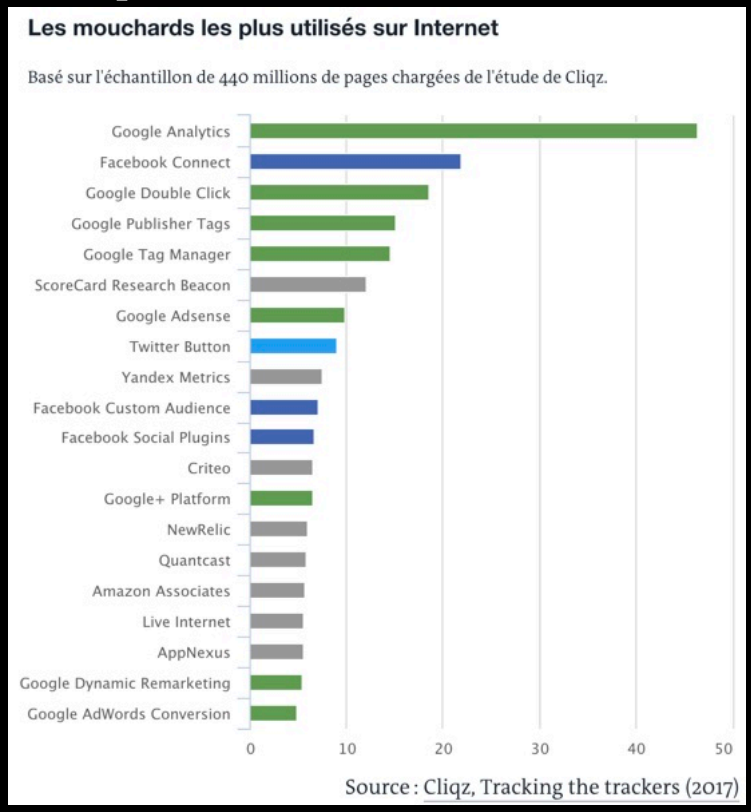
- 440M pages
- 850K visiteurs
- 77% contiennent des mouchards
- 1/4 des 23% restant proviennent de Facebook et Google, indétectables par Ghostery…



Les traqueurs

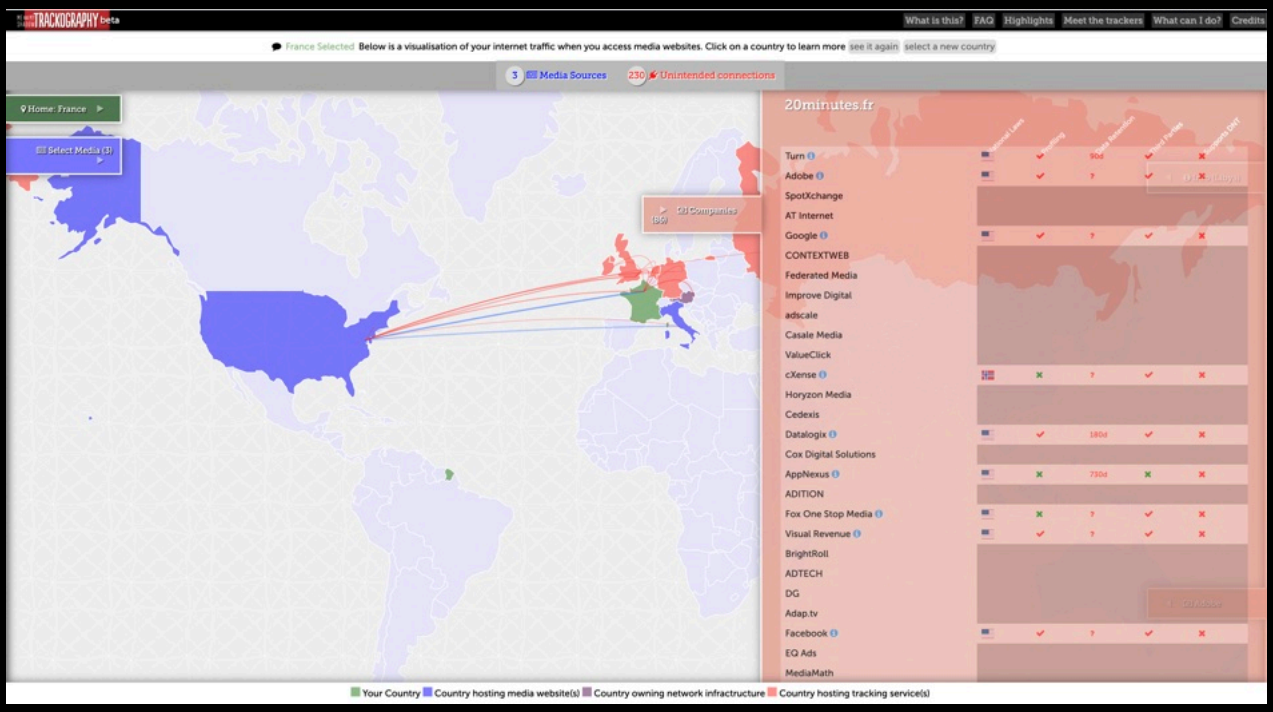
Étude menée par Cliqz en 2017 :

- 440M pages
- 850K visiteurs
- 77% contiennent des mouchards
- 1/4 des 23% restant proviennent de Facebook et Google, indétectables par Ghostery…



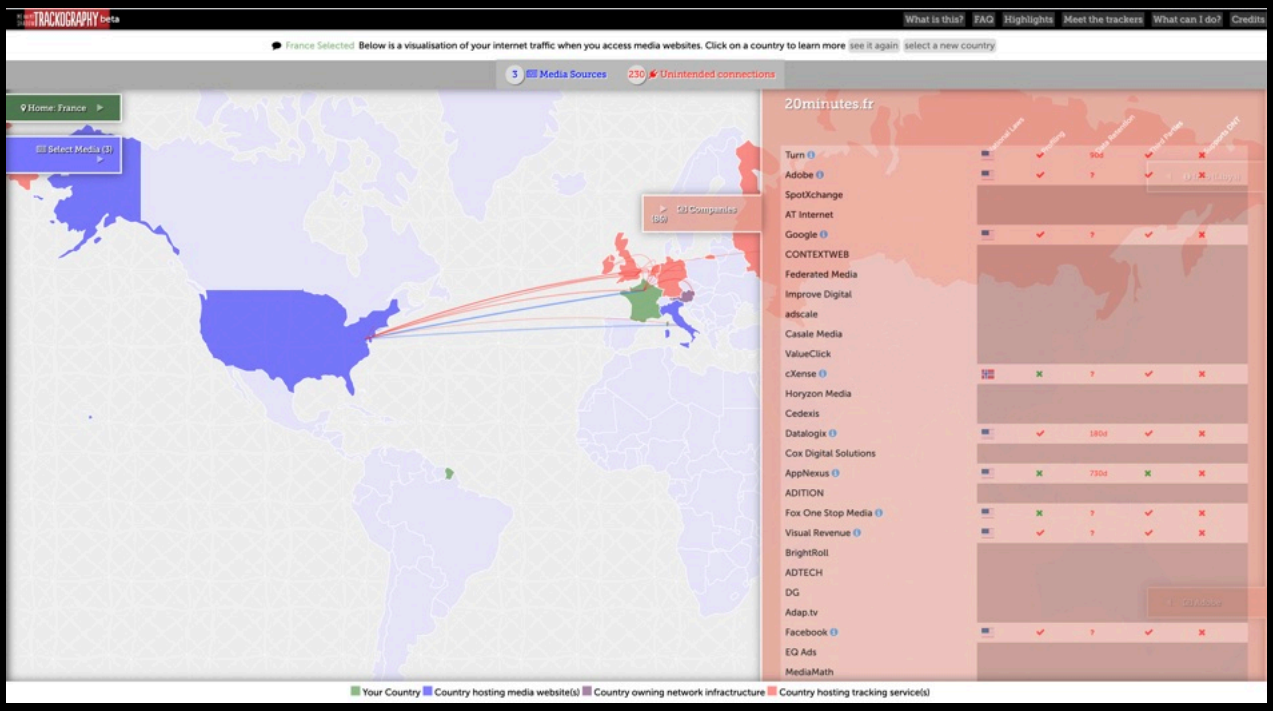
Tracking the Trackers

20minutes, L'Équipe et Le Monde génèrent 230 connexions vers 86 entreprises

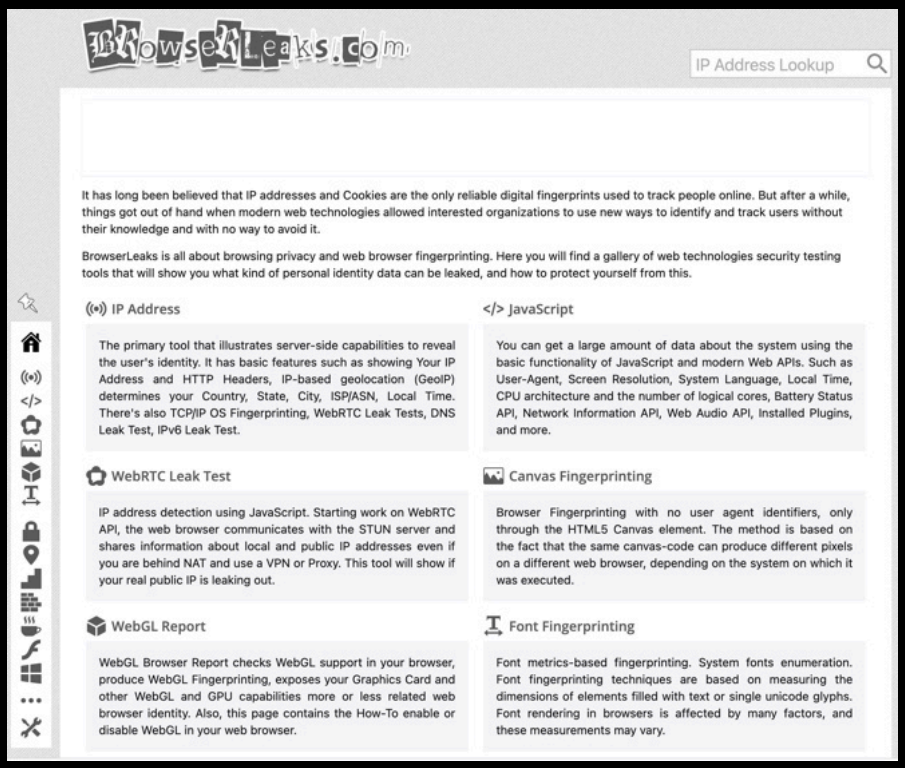


Tracking the Trackers

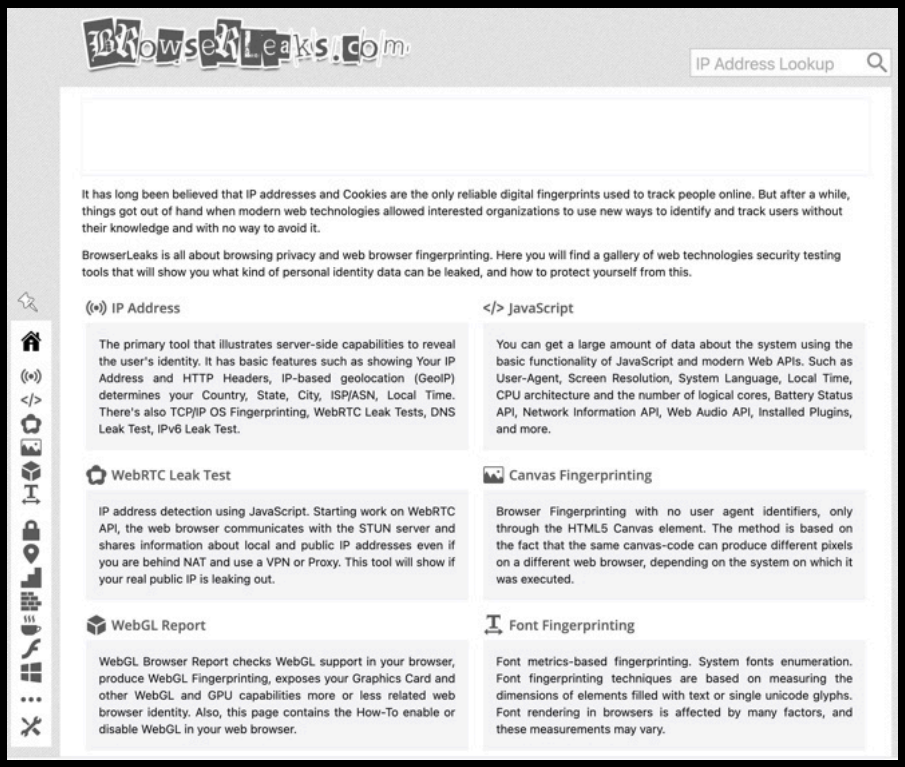
20minutes, L'Équipe et Le Monde génèrent 230 connexions vers 86 entreprises



IP, Cookies et plus…



IP, Cookies et plus…



Réseaux sociaux

Source : gendermatters.in

https://www.blogdumoderateur.com/chiffres-cles-reseaux-sociaux-internet-mobile-2021/

Facebook compte 2,91 milliards d'utilisateurs actifs mensuels en 2022

- Exposition de soi
- Brouillage Espace privé / Espace public
- Extension aux données relationnelles

Réseaux sociaux

Source : gendermatters.in

https://www.blogdumoderateur.com/chiffres-cles-reseaux-sociaux-internet-mobile-2021/

Facebook compte 2,91 milliards d'utilisateurs actifs mensuels en 2022

- Exposition de soi
- Brouillage Espace privé / Espace public
- Extension aux données relationnelles

Source : Infographie Trend Micro (2012)

Vous avez le droit de garder l'anonymat. Tout ce que vous publierez pourra être retenu contre vous
TMI : Trop d'info tue... la confidentialité en ligne
Quelles informations pouvez-vous vous permettre de publier sur vos réseaux sociaux préférés ?
16 % des utilisateurs indiquent où ils se trouvent
16 % des comptes Pinterest sont connectés à Twitter
En moyenne 229 suiveurs
1/4 des utilisateurs précisent où ils se trouvent chaque mois
Plus de 20 M d'internautes américains incluent leur date de naissance dans leur profil
Un utilisateur a en moyenne 229 amis Facebook
TYPE DE DONNÉES
PINTEREST
FACEBOOK
TWITTER
LINKEDIN
TUMBLR
GOOGLE+
SINA WEIBO
MIXI
Pour recruter, 89% des employeurs utilisent LinkedIn
Plus de 10 % des utilisateurs de Weibo ne sont pas chinois
Un utilisateur a en moyenne 287 suiveurs
Seuls 12 % des utilisateurs publient leur vrai nom
Un utilisateur a en moyenne 46 amis
LÉGENDE :
PARAMÈTRES DE CONFIDENTIALITÉ
CONTACTS
COORDONNÉES
DATE DE NAISSANCE
EMPLOI
LIEU
NOM


VOS DONNÉES SONT DES INDICES
Les informations que vous partagez permettent souvent de répondre à des questions de sécurité.
Quelles sont les données les plus partagées ?
63 % anniversaires
61 % écoles
51 % liens familiaux
48 % lieux de résidence
44 % émissions favorites
38 % artistes préférés
33 % livres préférés
26 % destinations de vacances
23 % noms d'animaux de compagnie
Vérifié mais non modifié
50 % des personnes disent vérifier leurs paramètres de confidentialité Facebook au moins tous les 2 ou 3 mois, sans pour autant les modifier aussi souvent.

Source : Infographie Trend Micro (2012)

25

Biopic façon puzzle

- *Portrait Google de Marc L****

  Revue Le Tigre, vol.28 (Nov. 2008)

  - Flickr : "voyages" 17.000 photos + Facebook + Youtube + ...
  - profession, collègues, soirées, événements, habitudes, rencontres, vie affective, goûts musicaux, etc.

"Elle a habité successivement Angers puis Metz, son chat s'appelle Lula, et, physiquement, elle a un peu le même genre que Claudia. À l'été 2006, vous êtes partis dans un camping à Pornic, dans une Golf blanche. La côte Atlantique, puis la Bretagne intérieure. Tu avais les cheveux courts, à l'époque, ça t'allait moins bien."

Biopic façon puzzle

- *Portrait Google de Marc L****

  Revue Le Tigre, vol.28 (Nov. 2008)

  - Flickr : "voyages" 17.000 photos + Facebook + Youtube + ...
  - profession, collègues, soirées, événements, habitudes, rencontres, vie affective, goûts musicaux, etc.

"Elle a habité successivement Angers puis Metz, son chat s'appelle Lula, et, physiquement, elle a un peu le même genre que Claudia. À l'été 2006, vous êtes partis dans un camping à Pornic, dans une Golf blanche. La côte Atlantique, puis la Bretagne intérieure. Tu avais les cheveux courts, à l'époque, ça t'allait moins bien."

Potentiel vertigineux…

- Élaboration de profils individuels détaillés
  - Facebook gagne 34€/an par profil
  - De l'anodin au projet de société : publicité ciblée, fabrication et diffusion de fausses nouvelles, manipulation d'opinion
- Inférence et prédiction
  - Graphe relationnel : projet Gaydar (2007), MIT
  - Variable : opinion politique, genre, race du chien, etc.

Potentiel vertigineux…

- Élaboration de profils individuels détaillés
  - Facebook gagne 34€/an par profil
  - De l'anodin au projet de société : publicité ciblée, fabrication et diffusion de fausses nouvelles, manipulation d'opinion
- Inférence et prédiction
  - Graphe relationnel : projet Gaydar (2007), MIT
  - Variable : opinion politique, genre, race du chien, etc.

Loi de finance 2020

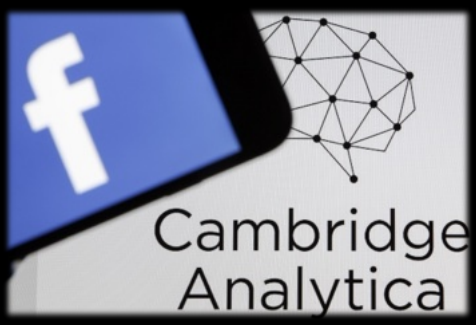
- Article 154 : aspirer/analyser les données personnelles sur Facebook, Twitter, Instagram, Leboncoin
- XP pendant 3 ans (fév. 2024), validée par le C.Const.
- Bercy a investi 20M€ dans les technos de fouille de données
- *Ère artisanale* : enquêteurs du Fisc (comme les compagnies d'assurance et l'assurance maladie)
- *Ère industrielle* :
  1. phase initiale : 20 ans d'historique (1M) de contrôles du fisc pour « apprendre les traits de la fraude » sur RSx
  2. phase de production : détection/prédiction de fraude

Loi de finance 2020

- Article 154 : aspirer/analyser les données personnelles sur Facebook, Twitter, Instagram, Leboncoin
- XP pendant 3 ans (fév. 2024), validée par le C.Const.
- Bercy a investi 20M€ dans les technos de fouille de données
- *Ère artisanale* : enquêteurs du Fisc (comme les compagnies d'assurance et l'assurance maladie)
- *Ère industrielle* :
  1. phase initiale : 20 ans d'historique (1M) de contrôles du fisc pour « apprendre les traits de la fraude » sur RSx
  2. phase de production : détection/prédiction de fraude

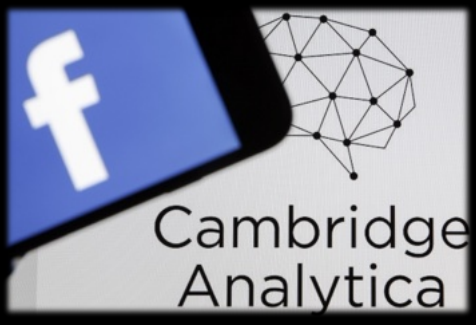
28

. . . et usages frauduleux



- mars 2018 : 87M de profils Facebook siphonnés
  - campagne présidentielle de D. Trump aux USA (2016)
  - campagne pro-Brexit : impact limité voire nul, selon un rapport du 17 oct. 2020 du *Information Commissionner Officer*
- août 2019 : Instagram et HYP3R

. . . et usages frauduleux



- mars 2018 : 87M de profils Facebook siphonnés
  - campagne présidentielle de D. Trump aux USA (2016)
  - campagne pro-Brexit : impact limité voire nul, selon un rapport du 17 oct. 2020 du *Information Commissionner Officer*
- août 2019 : Instagram et HYP3R

Assistants personnels

micro + enceinte + assistant vocal + IA externalisée

Microsoft Cortana Google Home
Apple HomePod/Siri Amazon Echo/Alexa

Source : Ruthe.de

Assistants personnels

micro + enceinte + assistant vocal + IA externalisée

Microsoft Cortana Google Home
Apple HomePod/Siri Amazon Echo/Alexa

Source : Ruthe.de

Assistants personnels

« Entre le 1er janvier et le 30 juin 2017, la société a fait droit à 1 200 requêtes, 189 mandats de recherche et 76 autres ordres émanant de cours de justice. Dans 42% des cas, Amazon disposait de toutes les informations demandées par les autorités. » L'ADN, 19 jan. 2018

Parfait mouchard !

- Recommandation CNIL : couper le micro…
- Métier : transcripteur ou dresseur d'IA (La Quadrature du Net, 18 mai 2018)

Assistants personnels

« Entre le 1er janvier et le 30 juin 2017, la société a fait droit à 1 200 requêtes, 189 mandats de recherche et 76 autres ordres émanant de cours de justice. Dans 42% des cas, Amazon disposait de toutes les informations demandées par les autorités. » L'ADN, 19 jan. 2018

Parfait mouchard !

- Recommandation CNIL : couper le micro…
- Métier : transcripteur ou dresseur d'IA (La Quadrature du Net, 18 mai 2018)

Paranoïa ?

Paranoïa ?

Vizio

Février 2017 : collecte de données à partir de la Smart TV de 11 millions de clients
amende = $2,2 millions

Vizio

Février 2017 : collecte de données à partir de la Smart TV de 11 millions de clients
amende = $2,2 millions

Police de Kyiv

fin fév. 2021 : 30 caméras embarquées (avec micro) sans mot de passe

https://reflets.info/articles/reflets-s-invite-par-hasard-dans-les-voitures-de-police-ukrainiennes

Police de Kyiv

fin fév. 2021 : 30 caméras embarquées (avec micro) sans mot de passe

https://reflets.info/articles/reflets-s-invite-par-hasard-dans-les-voitures-de-police-ukrainiennes

Santé

- objets connectés (« *quantified self* ») : montres, bracelets balances, etc.
- collecte de : nom, prénom, taille, âge, genre, poids, pression artérielle, fréquence cardiaque, glycémie, masse graisseuse, qualité du sommeil, …
- … et des fuites :
  - 16,71 Go de données concernant 61M d'enregistrements
    base de données de GetHealth (agrège FitBit et Apple HealthKit)
- 8000 failles recensées sur les pacemakers !

http://blog.whitescope.io/2017/05/understanding-pacemaker-systems.html

Santé

- objets connectés (« *quantified self* ») : montres, bracelets balances, etc.
- collecte de : nom, prénom, taille, âge, genre, poids, pression artérielle, fréquence cardiaque, glycémie, masse graisseuse, qualité du sommeil, …
- … et des fuites :
  - 16,71 Go de données concernant 61M d'enregistrements
    base de données de GetHealth (agrège FitBit et Apple HealthKit)
- 8000 failles recensées sur les pacemakers !

http://blog.whitescope.io/2017/05/understanding-pacemaker-systems.html

Standard Innovation

mars 2017 : collecte à partir de We-Vibe
amende = $3 millions

Standard Innovation

mars 2017 : collecte à partir de We-Vibe
amende = $3 millions

Où en sommes-nous ?

A. À propos des données personnelles
    1. Profil utilisateur et marché de la donnée
    2. Dispositifs de collecte :
        - ⋆ logs, traqueurs, RSx, IoT
    3. Surveillance de masse
    4. Risques et conséquences
B. Cadre législatif et réglementaire

Où en sommes-nous ?

A. À propos des données personnelles
    1. Profil utilisateur et marché de la donnée
    2. Dispositifs de collecte :
        - ⋆ logs, traqueurs, RSx, IoT
    3. Surveillance de masse
    4. Risques et conséquences
B. Cadre législatif et réglementaire

Pratique des États

- Renseignement, police, anti-terrorisme
  - Aux USA : programme PRISM de la NSA
    - révélations de Edward Snowden (juin 2013—?)
    - exilé en Russie, droit d'asile refusé par la France le 19/09/2019
    - Capacité de traitement des appels téléphoniques : « *three hops* »
  - En Europe : directive 2006/24/CE
    - conservation des données de connexion de 6 mois à 2 ans
    - invalidée par la Cour de Justice de l'UE le 08 avril 2014
  - En France : Loi Renseignement (19 mars 2015)

Pratique des États

- Renseignement, police, anti-terrorisme
  - Aux USA : programme PRISM de la NSA
    - révélations de Edward Snowden (juin 2013—?)
    - exilé en Russie, droit d'asile refusé par la France le 19/09/2019
    - Capacité de traitement des appels téléphoniques : « *three hops* »
  - En Europe : directive 2006/24/CE
    - conservation des données de connexion de 6 mois à 2 ans
    - invalidée par la Cour de Justice de l'UE le 08 avril 2014
  - En France : Loi Renseignement (19 mars 2015)

Tableau 3.1: Principaux programmes de surveillance

Nom du programme	Description des programmes allégués
Prism	Fournit à la NSA un accès direct aux serveurs centraux des neuf principales sociétés du numérique aux États-Unis, lui permettant de collecter des données sur leurs clients, ainsi que d'examiner leur historique, le contenu des courriels, les transferts de fichiers et les discussions en ligne.
Xkeyscore	Permet aux analystes de la NSA d'examiner, sans y être préalablement autorisés, de vastes bases de données contenant des courriels, des discussions en ligne et l'historique de navigation de millions d'utilisateurs d'internet ainsi que leurs métadonnées.
Upstream	Programmes de collecte gérés par la NSA qui consistent à mettre sur écoute sans autorisation des connexions internet câblées.
Bullrun	Programme de décryptage géré par la NSA pour tenter de pénétrer des technologies de cryptage largement utilisées, lui permettant de contourner le cryptage de données utilisé par des millions de personnes dans leurs transactions en ligne et leurs courriels.
Muscular	Programme conjoint utilisé par la NSA et le GCHQ visant à intercepter, à partir de liens privés, la circulation de données entre les principales plateformes telles que Yahoo, Google, Microsoft Hotmail et Windows Live Messenger.
Tempora	Activité de surveillance en amont permettant au GCHQ d'avoir accès aux câbles de fibre optique transportant d'énormes quantités de communications privées entre utilisateurs d'internet et de les partager avec la NSA.
Edgehill	Programme de décryptage géré par le GCHQ visant à décoder le trafic crypté utilisé par les sociétés pour fournir un accès à distance à leur système.

Sources: Moraes, C. (2013), document de travail n° 1, sur les programmes de surveillance US/UE et leur impact sur les droits fondamentaux des citoyens européens, *PE524.799v01-00, Bruxelles, 11 décembre 2013; Bowden, C. (2013),* Les programmes américains de surveillance et leur impact sur les droits fondamentaux des citoyens européens, *Étude réalisée pour le Parlement européen, PE 474.405, Bruxelles, septembre 2013*

FRA EUROPEAN UNION AGENCY FOR FUNDAMENTAL RIGHTS

Rapport annuel, Chap. 3, (2013)

Tableau 3.1: Principaux programmes de surveillance

Nom du programme	Description des programmes allégués
Prism	Fournit à la NSA un accès direct aux serveurs centraux des neuf principales sociétés du numérique aux États-Unis, lui permettant de collecter des données sur leurs clients, ainsi que d'examiner leur historique, le contenu des courriels, les transferts de fichiers et les discussions en ligne.
Xkeyscore	Permet aux analystes de la NSA d'examiner, sans y être préalablement autorisés, de vastes bases de données contenant des courriels, des discussions en ligne et l'historique de navigation de millions d'utilisateurs d'internet ainsi que leurs métadonnées.
Upstream	Programmes de collecte gérés par la NSA qui consistent à mettre sur écoute sans autorisation des connexions internet câblées.
Bullrun	Programme de décryptage géré par la NSA pour tenter de pénétrer des technologies de cryptage largement utilisées, lui permettant de contourner le cryptage de données utilisé par des millions de personnes dans leurs transactions en ligne et leurs courriels.
Muscular	Programme conjoint utilisé par la NSA et le GCHQ visant à intercepter, à partir de liens privés, la circulation de données entre les principales plateformes telles que Yahoo, Google, Microsoft Hotmail et Windows Live Messenger.
Tempora	Activité de surveillance en amont permettant au GCHQ d'avoir accès aux câbles de fibre optique transportant d'énormes quantités de communications privées entre utilisateurs d'internet et de les partager avec la NSA.
Edgehill	Programme de décryptage géré par le GCHQ visant à décoder le trafic crypté utilisé par les sociétés pour fournir un accès à distance à leur système.

Sources: Moraes, C. (2013), document de travail n° 1, sur les programmes de surveillance US/UE et leur impact sur les droits fondamentaux des citoyens européens, *PE524.799v01-00, Bruxelles, 11 décembre 2013; Bowden, C. (2013),* Les programmes américains de surveillance et leur impact sur les droits fondamentaux des citoyens européens, *Étude réalisée pour le Parlement européen, PE 474.405, Bruxelles, septembre 2013*

FRA EUROPEAN UNION AGENCY FOR FUNDAMENTAL RIGHTS

Rapport annuel, Chap. 3, (2013)

7e rapport d'activité de la CNCTR (2022)

	2018	2019	2020	2021	2022	Évolution 2021 / 2022
Nombre de personnes surveillées	22 038	22 210	21 952	22 958	20 958	- 8,7 %
au titre de la prévention du terrorisme	8 579	7 736	8 786	7 826	6 478	-17,2%
au titre de la prévention de la criminalité et de la délinquance organisées	5 416	5 693	5 021	5 932	5 471	- 7,8 %
au titre de la finalité prévue au 5° de l'article L. 811-3 du code de la sécurité intérieure[2]	2 116	3 021	3 238	3 466	2 692	- 22,3 %

7e rapport d'activité de la CNCTR (2022)

	2018	2019	2020	2021	2022	Évolution 2021 / 2022
Nombre de personnes surveillées	22 038	22 210	21 952	22 958	20 958	- 8,7 %
au titre de la prévention du terrorisme	8 579	7 736	8 786	7 826	6 478	-17,2%
au titre de la prévention de la criminalité et de la délinquance organisées	5 416	5 693	5 021	5 932	5 471	- 7,8 %
au titre de la finalité prévue au 5° de l'article L. 811-3 du code de la sécurité intérieure[2]	2 116	3 021	3 238	3 466	2 692	- 22,3 %

40

	2018	2019	2020	2021	2022	Évolution 2021 / 2022	Évolution 2018 / 2022
Accès aux données de connexion en temps différé *(Identification d'abonnés ou recensement de numéros d'abonnement) (article L.851-1 du code de la sécurité intérieure)*	28 741						
Accès aux données de connexion en temps différé *(Autres demandes, dont celles de « factures détaillées ») (article L. 851-1)*	17 443						
Accès aux données de connexion en temps réel *(article L. 851-2)*	278						
Géolocalisations en temps réel *(article L. 851-4)*	5 191						
Interceptions de sécurité via le GIC *(I de l'article L. 852-1)*	10 562						
Interceptions des communications par *IMSI catcher* *(II de l'article L. 852-1)*	0						
Interceptions de sécurité sur les réseaux exclusivement hertziens *(article L. 852-2)*	3						
Interceptions de correspondances émises ou reçues par la voie satellitaire *(article L. 852-3)*		-	-	-	-	0	-	-
Localisations des personnes ou des objets *(« Balisages ») (article L. 851 5)*		1 510	1 793	1 598	2 006	1 951	- 2,7 %	+ 29,2 %
Recueils de données de connexion par IMSI catcher *(article L. 851-6)*		272	288	311	583	641	+ 9,9 %	+ 135,7 %
Captations de paroles prononcées à titre privé *(article L. 853-1)*		566	655	629	812	914	+ 12,6 %	+ 61,5 %
Captations d'images dans un lieu privé *(article L. 853-1)*		2 437	2 627	935	1 326	2 400	+ 81 %	- 1,5 %
Recueils et captations de données informatiques *(article L. 853-2)*		3 082	3 591	2 418	3 758	4 260	+ 13,4 %	+ 38,2 %
Introductions dans des lieux privés *(article L. 853-3)*		3 206	3 599	2 021	2 682	3 767	+ 40,5 %[4]	+ 17,5 %
Ensemble des techniques de renseignement		73 291	73 534	79 605	87 588	89 502	+ 2,2 %	+ 22,1 %

	2018	2019	2020	2021	2022	Évolution 2021 / 2022	Évolution 2018 / 2022
Accès aux données de connexion en temps différé *(Identification d'abonnés ou recensement de numéros d'abonnement) (article L.851-1 du code de la sécurité intérieure)*	28 741						
Accès aux données de connexion en temps différé *(Autres demandes, dont celles de « factures détaillées ») (article L. 851-1)*	17 443						
Accès aux données de connexion en temps réel *(article L. 851-2)*	278						
Géolocalisations en temps réel *(article L. 851-4)*	5 191						
Interceptions de sécurité via le GIC *(I de l'article L. 852-1)*	10 562						
Interceptions des communications par *IMSI catcher* *(II de l'article L. 852-1)*	0						
Interceptions de sécurité sur les réseaux exclusivement hertziens *(article L. 852-2)*	3						
Interceptions de correspondances émises ou reçues par la voie satellitaire *(article L. 852-3)*		-	-	-	-	0	-	-
Localisations des personnes ou des objets *(« Balisages ») (article L. 851 5)*		1 510	1 793	1 598	2 006	1 951	- 2,7 %	+ 29,2 %
Recueils de données de connexion par IMSI catcher *(article L. 851-6)*		272	288	311	583	641	+ 9,9 %	+ 135,7 %
Captations de paroles prononcées à titre privé *(article L. 853-1)*		566	655	629	812	914	+ 12,6 %	+ 61,5 %
Captations d'images dans un lieu privé *(article L. 853-1)*		2 437	2 627	935	1 326	2 400	+ 81 %	- 1,5 %
Recueils et captations de données informatiques *(article L. 853-2)*		3 082	3 591	2 418	3 758	4 260	+ 13,4 %	+ 38,2 %
Introductions dans des lieux privés *(article L. 853-3)*		3 206	3 599	2 021	2 682	3 767	+ 40,5 %[4]	+ 17,5 %
Ensemble des techniques de renseignement		73 291	73 534	79 605	87 588	89 502	+ 2,2 %	+ 22,1 %

Loi Renseignement

- 3 « boîtes noires » actives en France depuis 2018 :
- Motif :
  - contre-terrorisme
  - extension à l'ingérence étrangère (cyber-guerre) en cours

« Ce mécanisme autorise les services du renseignement à aspirer un volume de données de connexion (les « *informations et documents* ») auprès de n'importe quel intermédiaire en ligne (hébergeurs, FAI, opérateurs, services en ligne…) pour les faire analyser par un algorithme classé secret-défense. »

- extension en cours aux URLs complètes (problème avec *https*)

Loi Renseignement

- 3 « boîtes noires » actives en France depuis 2018 :
- Motif :
  - contre-terrorisme
  - extension à l'ingérence étrangère (cyber-guerre) en cours

« Ce mécanisme autorise les services du renseignement à aspirer un volume de données de connexion (les « *informations et documents* ») auprès de n'importe quel intermédiaire en ligne (hébergeurs, FAI, opérateurs, services en ligne…) pour les faire analyser par un algorithme classé secret-défense. »

- extension en cours aux URLs complètes (problème avec *https*)

Fichiers de police

Rapport d'information de l'AN sur les fichiers mis à la disposition des forces de sécurité, D. Paris et P. Morel-À-L'Huissier (2018)

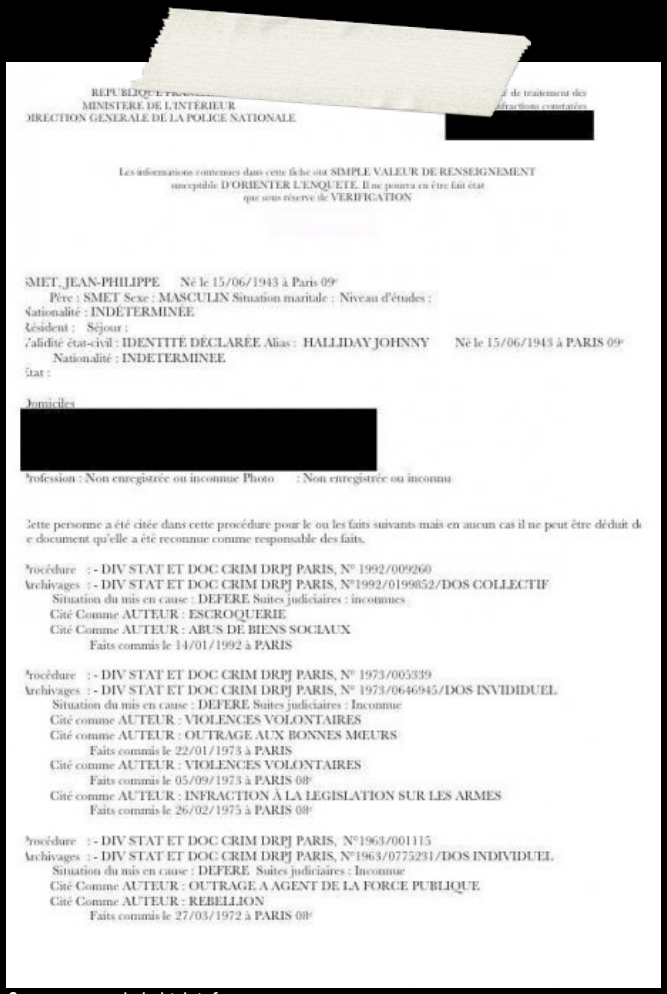
- 106 fichiers en France en 2018 (58 en 2009!)
  - FAED : empreintes digitales. 6,2M fiches en 2018, 220 non résolues
  - FNAEG : empreintes génétiques. 2,9M fiches en 2018, 480K non identifiées
  - Traitement d'Antécédents Judiciaires (TAJ) : auteurs, victimes, témoins
    - 87M affaires, 18,8M fiches de personnes mises en cause
    - fusionne STIC et JUDEX
  - Fichier des Personnes Recherchées (FPR):
    - 642K fiches pour 580K personnes en 2019
    - catégories : S (sûreté d'e l'État), T (débiteurs envers le Trésor), M (mineurs fugueurs), etc.
  - CRISTINA, FSPRT, GESTEREXT, FNIS, EASP, PASP, FIJAIS, …

Fichiers de police

Rapport d'information de l'AN sur les fichiers mis à la disposition des forces de sécurité, D. Paris et P. Morel-À-L'Huissier (2018)

- 106 fichiers en France en 2018 (58 en 2009!)
  - FAED : empreintes digitales. 6,2M fiches en 2018, 220 non résolues
  - FNAEG : empreintes génétiques. 2,9M fiches en 2018, 480K non identifiées
  - Traitement d'Antécédents Judiciaires (TAJ) : auteurs, victimes, témoins
    - 87M affaires, 18,8M fiches de personnes mises en cause
    - fusionne STIC et JUDEX
  - Fichier des Personnes Recherchées (FPR):
    - 642K fiches pour 580K personnes en 2019
    - catégories : S (sûreté d'e l'État), T (débiteurs envers le Trésor), M (mineurs fugueurs), etc.
  - CRISTINA, FSPRT, GESTEREXT, FNIS, EASP, PASP, FIJAIS, …

43



MINISTERE DE L'INTERIEUR
DIRECTION GENERALE DE LA POLICE NATIONALE

Les informations contenues dans cette fiche ont SIMPLE VALEUR DE RENSEIGNEMENT susceptible D'ORIENTER L'ENQUETE. Il ne pourra en être fait état que sous réserve de VERIFICATION

SMET, JEAN-PHILIPPE Né le 15/06/1943 à Paris 09e
Père : SMET Sexe : MASCULIN Situation maritale : Niveau d'études :
Nationalité : INDETERMINEE
Résident : Séjour :
Validité état-civil : IDENTITE DÉCLARÉE Alias : HALLIDAY JOHNNY Né le 15/06/1943 à PARIS 09e
Nationalité : INDETERMINEE
Etat :

Domiciles
[redacted]
Profession : Non enregistrée ou inconnue Photo : Non enregistrée ou inconnu

Cette personne a été citée dans cette procédure pour le ou les faits suivants mais en aucun cas il ne peut être déduit de ce document qu'elle a été reconnue comme responsable des faits.

Procédure : - DIV STAT ET DOC CRIM DRPJ PARIS, N° 1992/009260
Archivages : - DIV STAT ET DOC CRIM DRPJ PARIS, N°1992/0199852/DOS COLLECTIF
Situation du mis en cause : DEFERE Suites judiciaires : inconnues
Cité Comme AUTEUR : ESCROQUERIE
Cité Comme AUTEUR : ABUS DE BIENS SOCIAUX
Faits commis le 14/01/1992 à PARIS

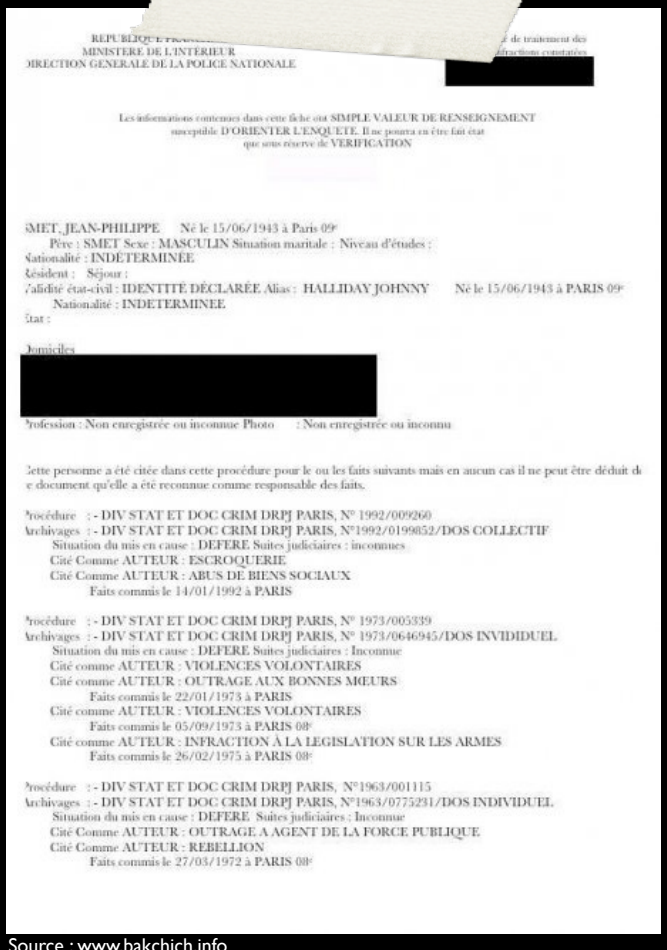
Procédure : - DIV STAT ET DOC CRIM DRPJ PARIS, N° 1973/005339
Archivages : - DIV STAT ET DOC CRIM DRPJ PARIS, N° 1973/0646945/DOS INVIDIDUEL
Situation du mis en cause : DEFERE Suites judiciaires : Inconnue
Cité comme AUTEUR : VIOLENCES VOLONTAIRES
Cité comme AUTEUR : OUTRAGE AUX BONNES MEURS
Faits commis le 22/01/1973 à PARIS
Cité comme AUTEUR : VIOLENCES VOLONTAIRES
Faits commis le 05/09/1973 à PARIS 08e
Cité comme AUTEUR : INFRACTION À LA LEGISLATION SUR LES ARMES
Faits commis le 26/02/1975 à PARIS 08e

Procédure : - DIV STAT ET DOC CRIM DRPJ PARIS, N°1963/001115
Archivages : - DIV STAT ET DOC CRIM DRPJ PARIS, N°1963/0775231/DOS INDIVIDUEL
Situation du mis en cause : DEFERE Suites judiciaires : Inconnue
Cité Comme AUTEUR : OUTRAGE A AGENT DE LA FORCE PUBLIQUE
Cité Comme AUTEUR : REBELLION
Faits commis le 27/03/1972 à PARIS 08e

Source : www.bakchich.info

STIC de J. H.

- Profession : *non enregistrée ou inconnue*
- Nationalité : *indéterminée*
- 26.10.67, Paris XVIe : auteur de *violences volontaires* [Déféré]
- [suites judiciaires inconnues] pour
  - *violence volontaire* (1973)
  - *outrage aux bonnes moeurs* (1973)
  - *infraction à la législation sur les armes* (1975)
  - *abus de biens sociaux* (1992)
  - *escroquerie* (1992)
- absence de mises à jour
- conservation illégale des données (15/40 ans)
- revente d'infos ("tricoche")



MINISTERE DE L'INTERIEUR
DIRECTION GENERALE DE LA POLICE NATIONALE

Les informations contenues dans cette fiche ont SIMPLE VALEUR DE RENSEIGNEMENT susceptible D'ORIENTER L'ENQUETE. Il ne pourra en être fait état que sous réserve de VERIFICATION

SMET, JEAN-PHILIPPE Né le 15/06/1943 à Paris 09e
Père : SMET Sexe : MASCULIN Situation maritale : Niveau d'études :
Nationalité : INDETERMINEE
Résident : Séjour :
Validité état-civil : IDENTITE DÉCLARÉE Alias : HALLIDAY JOHNNY Né le 15/06/1943 à PARIS 09e
Nationalité : INDETERMINEE
Etat :

Domiciles
[redacted]
Profession : Non enregistrée ou inconnue Photo : Non enregistrée ou inconnu

Cette personne a été citée dans cette procédure pour le ou les faits suivants mais en aucun cas il ne peut être déduit de ce document qu'elle a été reconnue comme responsable des faits.

Procédure : - DIV STAT ET DOC CRIM DRPJ PARIS, N° 1992/009260
Archivages : - DIV STAT ET DOC CRIM DRPJ PARIS, N°1992/0199852/DOS COLLECTIF
Situation du mis en cause : DEFERE Suites judiciaires : inconnues
Cité Comme AUTEUR : ESCROQUERIE
Cité Comme AUTEUR : ABUS DE BIENS SOCIAUX
Faits commis le 14/01/1992 à PARIS

Procédure : - DIV STAT ET DOC CRIM DRPJ PARIS, N° 1973/005339
Archivages : - DIV STAT ET DOC CRIM DRPJ PARIS, N° 1973/0646945/DOS INVIDIDUEL
Situation du mis en cause : DEFERE Suites judiciaires : Inconnue
Cité comme AUTEUR : VIOLENCES VOLONTAIRES
Cité comme AUTEUR : OUTRAGE AUX BONNES MEURS
Faits commis le 22/01/1973 à PARIS
Cité comme AUTEUR : VIOLENCES VOLONTAIRES
Faits commis le 05/09/1973 à PARIS 08e
Cité comme AUTEUR : INFRACTION À LA LEGISLATION SUR LES ARMES
Faits commis le 26/02/1975 à PARIS 08e

Procédure : - DIV STAT ET DOC CRIM DRPJ PARIS, N°1963/001115
Archivages : - DIV STAT ET DOC CRIM DRPJ PARIS, N°1963/0775231/DOS INDIVIDUEL
Situation du mis en cause : DEFERE Suites judiciaires : Inconnue
Cité Comme AUTEUR : OUTRAGE A AGENT DE LA FORCE PUBLIQUE
Cité Comme AUTEUR : REBELLION
Faits commis le 27/03/1972 à PARIS 08e

Source : www.bakchich.info

STIC de J. H.

- Profession : *non enregistrée ou inconnue*
- Nationalité : *indéterminée*
- 26.10.67, Paris XVIe : auteur de *violences volontaires* [Déféré]
- [suites judiciaires inconnues] pour
  - *violence volontaire* (1973)
  - *outrage aux bonnes moeurs* (1973)
  - *infraction à la législation sur les armes* (1975)
  - *abus de biens sociaux* (1992)
  - *escroquerie* (1992)
- absence de mises à jour
- conservation illégale des données (15/40 ans)
- revente d'infos ("tricoche")

TES

Titres Électroniques Sécurisés : « le fichier des gens honnêtes »

- demandes de CNIe et passeports (60M)
  - données d'état civil, avec filiation
  - couleur des yeux, taille, adresse
  - photo du visage, empreinte digitale, signature numérisée
- Rétention = 15 ans (90j en Allemande et en Belgique)
- Interconnexions avec des fichiers français et paneuropéens (INTERPOL)
- Pile de rapports défavorables de la CNIL et l'ANSSI

1974 ~~SAFARI~~ * 2012 ~~CNIE~~ * 2017 TES

TES

Titres Électroniques Sécurisés : « le fichier des gens honnêtes »

- demandes de CNIe et passeports (60M)
  - données d'état civil, avec filiation
  - couleur des yeux, taille, adresse
  - photo du visage, empreinte digitale, signature numérisée
- Rétention = 15 ans (90j en Allemande et en Belgique)
- Interconnexions avec des fichiers français et paneuropéens (INTERPOL)
- Pile de rapports défavorables de la CNIL et l'ANSSI

1974 ~~SAFARI~~ * 2012 ~~CNIE~~ * 2017 TES

TousAntiCovid

- Application de suivi des contacts (*contact tracing*)
- Déploiement à partir du 02 juin 2020 en France
- Développement coordonné par Inria et ANSSI
- Infrastructure souveraine (Dassault Systèmes)
- Protocole centralisé ROBERT (vs. DP-3T, 3e voie DESIRE)
- contact = 15mn à moins d'1 mètre, par Bluetooth
- Controverse à propos de l'efficacité vs. sécurité/vie privée

TousAntiCovid

- Application de suivi des contacts (*contact tracing*)
- Déploiement à partir du 02 juin 2020 en France
- Développement coordonné par Inria et ANSSI
- Infrastructure souveraine (Dassault Systèmes)
- Protocole centralisé ROBERT (vs. DP-3T, 3e voie DESIRE)
- contact = 15mn à moins d'1 mètre, par Bluetooth
- Controverse à propos de l'efficacité vs. sécurité/vie privée

« Sousveillance »
un concept ambivalent

- Lanceurs d'alerte
  - Affaire NSA et Edward Snowden
  - Wikileaks
  - @ délinquants sexuels : familywatchdog.us
- Pilori numérique (« *bashing* »)
  - dénoncer publiquement une personne physique ou morale
  - mouvements structurants : #BalanceTonPorc, #MeeToo

« Sousveillance »
un concept ambivalent

- Lanceurs d'alerte
  - Affaire NSA et Edward Snowden
  - Wikileaks
  - @ délinquants sexuels : familywatchdog.us
- Pilori numérique (« *bashing* »)
  - dénoncer publiquement une personne physique ou morale
  - mouvements structurants : #BalanceTonPorc, #MeeToo

Où en sommes-nous ?

A. À propos des données personnelles
    1. Profil utilisateur et marché de la donnée
    2. Dispositifs de collecte :
        ⋆ moteurs de recherche, RSx, IoT
    3. Surveillance de masse
    4. Risques et conséquences
B. Cadre législatif et réglementaire

Où en sommes-nous ?

A. À propos des données personnelles
    1. Profil utilisateur et marché de la donnée
    2. Dispositifs de collecte :
        ⋆ moteurs de recherche, RSx, IoT
    3. Surveillance de masse
    4. Risques et conséquences
B. Cadre législatif et réglementaire

Les risques

- Usage détourné
- Incohérence des fichiers
- Fuite de données
  - Exploitation d'une faille de sécurité
  - Malveillance interne
  - Maladresse
- Revente
  - Marché (noir) des données

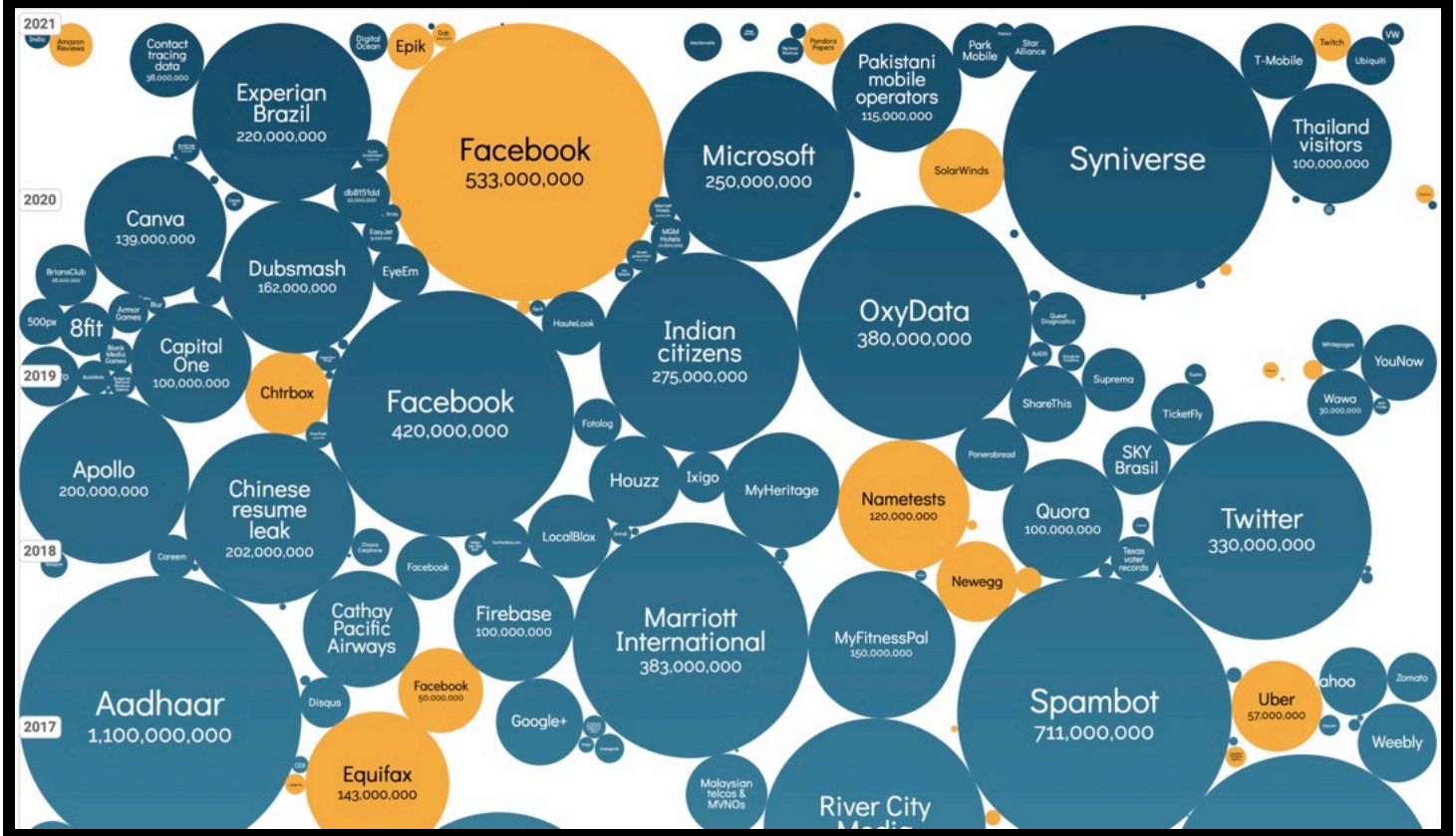
Les risques

- Usage détourné
- Incohérence des fichiers
- Fuite de données
  - Exploitation d'une faille de sécurité
  - Malveillance interne
  - Maladresse
- Revente
  - Marché (noir) des données

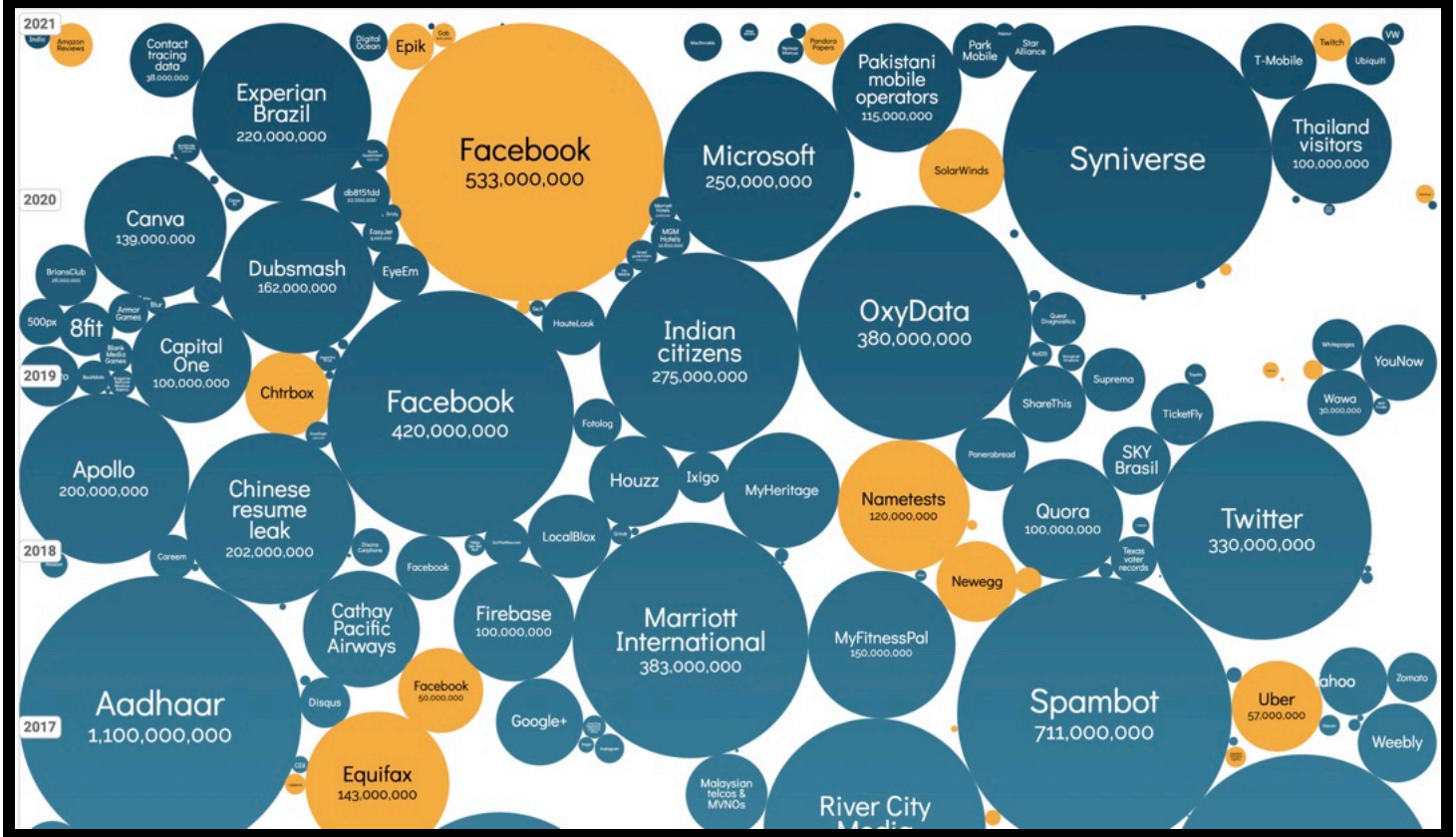
- Banque majeure aux US, ~plus gros émetteur de CB au monde
- vol de données de 106 millions de clients :
  - 140K N°SS, 80K ID comptes, 1M N° assurance sociale, nom, adresse, mail, tel, revenu déclaré, transactions bancaires sur 23j, historique des paiements, limite de crédit, solde courant.
- Selon The Guardian, intrusion de la hackeuse Paige Thompson le 22-23 mars 2019, découverte le 19 juillet 2019
- Des données personnelles auraient été aperçues sur GitHub le 17 juillet 2020 !

- Banque majeure aux US, ~plus gros émetteur de CB au monde
- vol de données de 106 millions de clients :
  - 140K N°SS, 80K ID comptes, 1M N° assurance sociale, nom, adresse, mail, tel, revenu déclaré, transactions bancaires sur 23j, historique des paiements, limite de crédit, solde courant.
- Selon The Guardian, intrusion de la hackeuse Paige Thompson le 22-23 mars 2019, découverte le 19 juillet 2019
- Des données personnelles auraient été aperçues sur GitHub le 17 juillet 2020 !

50



informationisbeautiful.net / données: databreaches.net

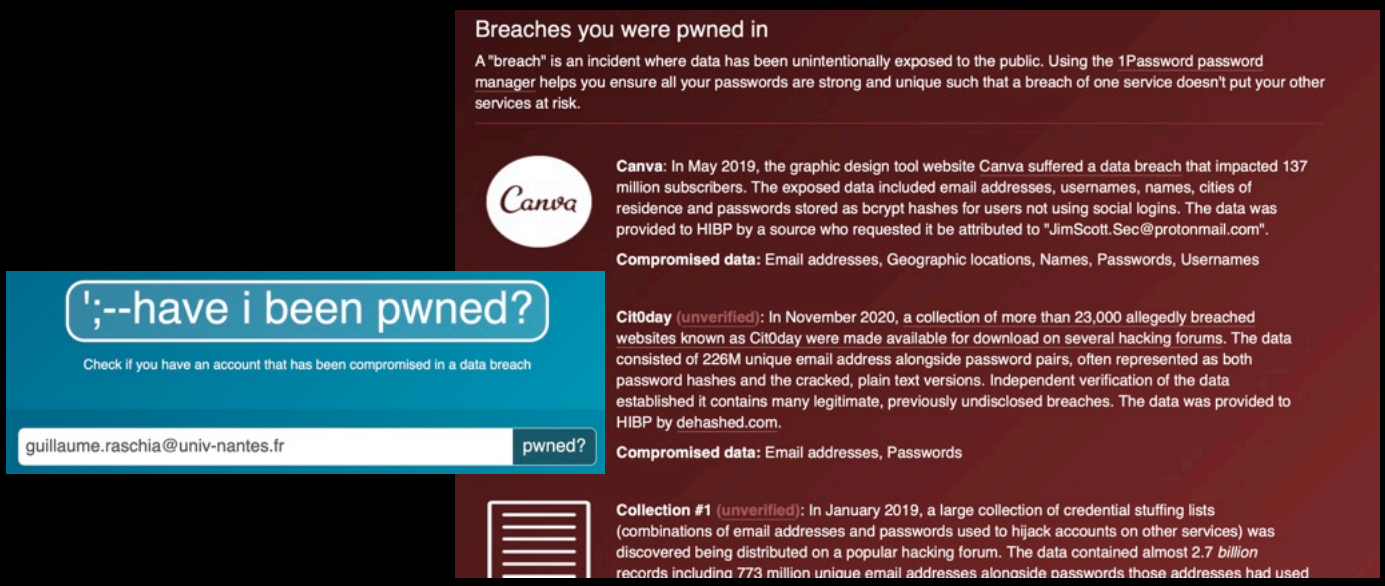


informationisbeautiful.net / données: databreaches.net

Have I Been Pwned?

Projet de Troy Hunt, expert Microsoft en sécurité informatique

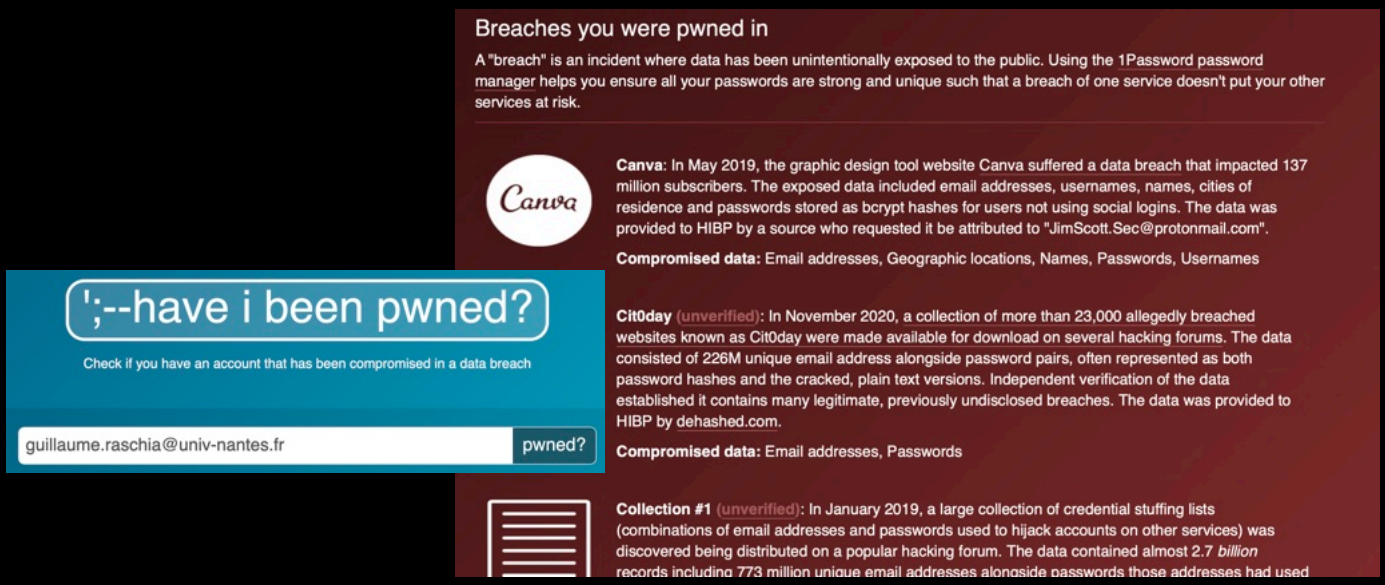
https://haveibeenpwned.com



Have I Been Pwned?

Projet de Troy Hunt, expert Microsoft en sécurité informatique

https://haveibeenpwned.com



52

Sur le Dark Web…

- Marché noir de la donnée volée
  - N°SS aléatoire = 0,05€, N°SS ciblé = 3$
  - Pack « Fullz » (avec N°CB), de 24$ à 100$
  - Données bancaires : de 270$ à 1100$, selon solvabilité !
- Données de santé :
  - 500K dossiers médicaux Fr, achetés 1370 fois (taux de satisfaction client = 95,80% !)
  - 50M dossiers patients US, *etc.*
- Selon Javelin Strategy & Research :
  - 776$ et 20h pour régulariser une usurpation d'identité

Sur le Dark Web…

- Marché noir de la donnée volée
  - N°SS aléatoire = 0,05€, N°SS ciblé = 3$
  - Pack « Fullz » (avec N°CB), de 24$ à 100$
  - Données bancaires : de 270$ à 1100$, selon solvabilité !
- Données de santé :
  - 500K dossiers médicaux Fr, achetés 1370 fois (taux de satisfaction client = 95,80% !)
  - 50M dossiers patients US, *etc.*
- Selon Javelin Strategy & Research :
  - 776$ et 20h pour régulariser une usurpation d'identité

Les conséquences

- Publicité ciblée, manipulation
- Réputation, crédibilité et vie privée
  - 35% recruteurs ont déjà écarté un candidat ayant une mauvaise e-réputation
- Usurpation d'identité, escroquerie
  - 24B id. volés selon Risk Based Security Inc.
- Intégrité physique et morale
  - dissidence politique
  - cambriolage, vendetta, *etc.*
  - cyber-humiliation, cyber-harcèlement, cyber-prédation

Les conséquences

- Publicité ciblée, manipulation
- Réputation, crédibilité et vie privée
  - 35% recruteurs ont déjà écarté un candidat ayant une mauvaise e-réputation
- Usurpation d'identité, escroquerie
  - 24B id. volés selon Risk Based Security Inc.
- Intégrité physique et morale
  - dissidence politique
  - cambriolage, vendetta, *etc.*
  - cyber-humiliation, cyber-harcèlement, cyber-prédation

Où en sommes-nous ?

A. À propos des données personnelles
B. Cadre législatif et réglementaire
    1. Données à caractère personnel
    2. RGPD
    3. Anonymisation

Où en sommes-nous ?

A. À propos des données personnelles
B. Cadre législatif et réglementaire
    1. Données à caractère personnel
    2. RGPD
    3. Anonymisation

La vie privée

- Droit fondamental consacré (D.U.D.H. 1948)
- Une liberté en tension :
  - vs. la liberté d'expression
  - vs. la sécurité *"rien à me reprocher"*
  - vs. des gains économiques
- Une liberté érodée par :
  - le progrès technologique
  - l'exposition et la mesure de soi
- Application délicate du cadre juridique

La vie privée

- Droit fondamental consacré (D.U.D.H. 1948)
- Une liberté en tension :
  - vs. la liberté d'expression
  - vs. la sécurité *"rien à me reprocher"*
  - vs. des gains économiques
- Une liberté érodée par :
  - le progrès technologique
  - l'exposition et la mesure de soi
- Application délicate du cadre juridique

Histoire législative

- Loi informatique et liberté du 6 jan. 1978
  - Création de la Commission Nationale de l'Informatique et des Libertés (CNIL)
- Directive européenne 95/46/CE de oct. 1995
- Transposition en droit français : loi du 6 août 2004 décret n° 2005-1309 du 20 oct. 2005
- Règlement Eu 2016/679 : **25 mai 2018**
- Donnée *à caractère personnel* | traitement | fichiers

Histoire législative

- Loi informatique et liberté du 6 jan. 1978
  - Création de la Commission Nationale de l'Informatique et des Libertés (CNIL)
- Directive européenne 95/46/CE de oct. 1995
- Transposition en droit français : loi du 6 août 2004 décret n° 2005-1309 du 20 oct. 2005
- Règlement Eu 2016/679 : **25 mai 2018**
- Donnée *à caractère personnel* | traitement | fichiers

“à caractère personnel”

- Toute information relative à une personne physique identifiée ou identifiable, directement ou indirectement, telle que :
    - nom, prénom, photographie
    - adresse, numéro de S.S.
    - numéro de tél./carte d’identité/compte bancaire/matricule/dossier/etc.
    - plaque d’immatriculation automobile
    - @mail, cookie (sessionID), *etc.*
- Cas historiquement épineux (et résolu) de l’@IP

“à caractère personnel”

- Toute information relative à une personne physique identifiée ou identifiable, directement ou indirectement, telle que :
    - nom, prénom, photographie
    - adresse, numéro de S.S.
    - numéro de tél./carte d’identité/compte bancaire/matricule/dossier/etc.
    - plaque d’immatriculation automobile
    - @mail, cookie (sessionID), *etc.*
- Cas historiquement épineux (et résolu) de l’@IP

@IP : identification

Received: from localhost (debian [127.0.0.1])
by smtp-tls.univ-nantes.fr (Postfix) with ESMTP id 214E812810D
for ; Tue, 13 Oct 2009 12:02:55 +0200 (CEST)
X-Virus-Scanned: Debian amavisd-new at univ-nantes.fr
Received: from smtp-tls.univ-nantes.fr ([127.0.0.1])
by localhost (SMTP-TLS.univ-nantes.fr [127.0.0.1]) (amavisd-new, port 10024)
with LMTP id YQzhP7Vh6Imy for ;
Tue, 13 Oct 2009 12:02:55 +0200 (CEST)
Received: from [192.168.248.203] (nantes.wifi.univ-nantes.fr [193.52.107.31])
(using TLSv1 with cipher AES128-SHA (128/128 bits))
(No client certificate requested)
by smtp-tls.univ-nantes.fr (Postfix) with ESMTPSA id 0C1221280E4
for ; Tue, 13 Oct 2009 12:02:55 +0200 (CEST)
Message-Id:
From: Guillaume Raschia
To: test@yopmail.com
Content-Type: text/plain
Content-Transfer-Encoding: 7bit
Mime-Version: 1.0 (Apple Message framework v936)
Subject: test : @IP?
Date: Tue, 13 Oct 2009 12:02:38 +0200
X-Mailer: Apple Mail (2.936)

@IP : identification

Received: from localhost (debian [127.0.0.1])
by smtp-tls.univ-nantes.fr (Postfix) with ESMTP id 214E812810D
for ; Tue, 13 Oct 2009 12:02:55 +0200 (CEST)
X-Virus-Scanned: Debian amavisd-new at univ-nantes.fr
Received: from smtp-tls.univ-nantes.fr ([127.0.0.1])
by localhost (SMTP-TLS.univ-nantes.fr [127.0.0.1]) (amavisd-new, port 10024)
with LMTP id YQzhP7Vh6Imy for ;
Tue, 13 Oct 2009 12:02:55 +0200 (CEST)
Received: from [192.168.248.203] (nantes.wifi.univ-nantes.fr [193.52.107.31])
(using TLSv1 with cipher AES128-SHA (128/128 bits))
(No client certificate requested)
by smtp-tls.univ-nantes.fr (Postfix) with ESMTPSA id 0C1221280E4
for ; Tue, 13 Oct 2009 12:02:55 +0200 (CEST)
Message-Id:
From: Guillaume Raschia
To: test@yopmail.com
Content-Type: text/plain
Content-Transfer-Encoding: 7bit
Mime-Version: 1.0 (Apple Message framework v936)
Subject: test : @IP?
Date: Tue, 13 Oct 2009 12:02:38 +0200
X-Mailer: Apple Mail (2.936)

@IP : contenu sensible

@IP : contenu sensible

@IP : géo-localisation

@IP : géo-localisation

@IPs dans le Futur

- IPV6, réseaux ad hoc : informatique ubiquitaire, intelligence ambiante, internet des objets, réseaux de capteurs, convergence 5G, etc.
- Une @IP unique et fixe / “machin”
- Plusieurs “machins” / personne
- Une communication et des échanges permanents

@IPs dans le Futur

- IPV6, réseaux ad hoc : informatique ubiquitaire, intelligence ambiante, internet des objets, réseaux de capteurs, convergence 5G, etc.
- Une @IP unique et fixe / “machin”
- Plusieurs “machins” / personne
- Une communication et des échanges permanents

Les données sensibles

- Des données « très » personnelles
  - origines ethniques
  - opinions politiques, philosophiques, religieuses
  - appartenance syndicale
  - orientation et pratiques sexuelles
  - données de santé
  - données génétiques
  - données biométriques

Les données sensibles

- Des données « très » personnelles
  - origines ethniques
  - opinions politiques, philosophiques, religieuses
  - appartenance syndicale
  - orientation et pratiques sexuelles
  - données de santé
  - données génétiques
  - données biométriques

Cas pratique

A votre avis, la collecte des données suivantes doit-elle être considérée comme un traitement de données personnelles ?

Année de naissance	**Lieu de naissance**	**Profession**	**Nationalité**	**Date de l'interruption volontaire de grossesse**
1978	Paris	Enseignante	FR	10/10/2000
1976	Strasbourg	Agricultrice	FR	11/08/1999
1975	Ceintrey	Avocate	IT	12/01/1995
1974	Lyon	Commerçante	DK	23/03/1996

Réponse de la CNIL : effectivement, il s'agit bien d'un traitement de données indirectement nominatives (la question était posée avant 2004). Pourquoi ? Simplement en raison du risque d'atteinte à la vie privée qui résulterait de l'identification des personnes. Risque suffisant pour dicter à l'autorité administrative la plus grande prudence dans sa décision.

Source : post de blog http://www.donneespersonnelles.fr/adresse-ip-est-elle-donnee-personnelle par T. Devergranne le 18/11/2011

Cas pratique

A votre avis, la collecte des données suivantes doit-elle être considérée comme un traitement de données personnelles ?

Année de naissance	**Lieu de naissance**	**Profession**	**Nationalité**	**Date de l'interruption volontaire de grossesse**
1978	Paris	Enseignante	FR	10/10/2000
1976	Strasbourg	Agricultrice	FR	11/08/1999
1975	Ceintrey	Avocate	IT	12/01/1995
1974	Lyon	Commerçante	DK	23/03/1996

Réponse de la CNIL : effectivement, il s'agit bien d'un traitement de données indirectement nominatives (la question était posée avant 2004). Pourquoi ? Simplement en raison du risque d'atteinte à la vie privée qui résulterait de l'identification des personnes. Risque suffisant pour dicter à l'autorité administrative la plus grande prudence dans sa décision.

Source : post de blog http://www.donneespersonnelles.fr/adresse-ip-est-elle-donnee-personnelle par T. Devergranne le 18/11/2011

Où en sommes-nous ?

A. À propos des données personnelles
B. Cadre législatif et réglementaire
    1. Données à caractère personnel
    2. RGPD
    3. Anonymisation

Où en sommes-nous ?

A. À propos des données personnelles
B. Cadre législatif et réglementaire
    1. Données à caractère personnel
    2. RGPD
    3. Anonymisation

RGPD

Règlement Général sur la Protection des Données

en vigueur depuis mai 2018

- Uniformiser les règles communautaires
- Sanctuariser le droit des personnes
- Renforcer les sanctions
- Clarifier les responsabilités
- Cadrer la conformité

En Chine : *Personal Information Protection Law* (PIPL)

depuis le 1er novembre 2021

RGPD

Règlement Général sur la Protection des Données

en vigueur depuis mai 2018

- Uniformiser les règles communautaires
- Sanctuariser le droit des personnes
- Renforcer les sanctions
- Clarifier les responsabilités
- Cadrer la conformité

En Chine : *Personal Information Protection Law* (PIPL)

depuis le 1er novembre 2021

66

Les acteurs

- La personne (dont les données personnelles sont en jeu)
- Responsable de traitement (*Data Controller*): entité qui détermine les finalités et les moyens (pas de décharge en cas de sous-traitance)
- Sous-traitant (*Data Processor*)
- Destinataire : récipiendaire des données
- Tiers

Les acteurs

- La personne (dont les données personnelles sont en jeu)
- Responsable de traitement (*Data Controller*): entité qui détermine les finalités et les moyens (pas de décharge en cas de sous-traitance)
- Sous-traitant (*Data Processor*)
- Destinataire : récipiendaire des données
- Tiers

Les droits CNIL

- Droit d'information
- Consentement (*Opt-in*), notification
- Droit d'opposition légitimée
- Droit d'accès direct ou indirect
- Droit de rectification
- Durée de conservation limitée au service

Les droits CNIL

- Droit d'information
- Consentement (*Opt-in*), notification
- Droit d'opposition légitimée
- Droit d'accès direct ou indirect
- Droit de rectification
- Durée de conservation limitée au service

Complément RGPD

- Droit à l'oubli (ou effacement)
- Portabilité
- Réclamation, recours et réparation
- Autorisation parentale pour les mineurs

Complément RGPD

- Droit à l'oubli (ou effacement)
- Portabilité
- Réclamation, recours et réparation
- Autorisation parentale pour les mineurs

Nouveaux paradigmes

- Privacy by Design
- Privacy by Default

Nouveaux paradigmes

- Privacy by Design
- Privacy by Default

Obligations de conformité

- Tenue d'un registre des traitements
- Délégué à la protection des données
- Politique de sécurité/intégrité
- Clauses contractuelles type ou règles contraignantes pour la sous-traitance
- Notification des violations de données sous 72h

Obligations de conformité

- Tenue d'un registre des traitements
- Délégué à la protection des données
- Politique de sécurité/intégrité
- Clauses contractuelles type ou règles contraignantes pour la sous-traitance
- Notification des violations de données sous 72h

Obligations de conformité

- Changement de régime avant/après RGPD :
  - pas de déclaration CNIL préalable
- Analyse d'impact (*Data Protection Impact Assessment)*
  - pour les traitements présentant un « risque élevé » pour les droits et libertés
  - à soumettre à l'autorité de contrôle (CNIL)

Obligations de conformité

- Changement de régime avant/après RGPD :
  - pas de déclaration CNIL préalable
- Analyse d'impact (*Data Protection Impact Assessment)*
  - pour les traitements présentant un « risque élevé » pour les droits et libertés
  - à soumettre à l'autorité de contrôle (CNIL)

Le DPO

- Délégué à la protection des données (*Data Privacy Officer*) à partir de mai 2018
- ex-Correspondant Informatique et Libertés
- référent pour une collectivité territoriale, une entreprise publique ou privée, une association
- « chef d'orchestre » de la conformité
- rôle d'information, contrôle, conseil

Le DPO

- Délégué à la protection des données (*Data Privacy Officer*) à partir de mai 2018
- ex-Correspondant Informatique et Libertés
- référent pour une collectivité territoriale, une entreprise publique ou privée, une association
- « chef d'orchestre » de la conformité
- rôle d'information, contrôle, conseil

Mécanisme de sanctions

- Enquêtes, contrôles, audits
- Avertissement, mise en demeure
- Suspension de flux et/ou certification
- Amende administrative
  - 20 millions € ou 4% du CA annuel mondial
  - Principe de proportionnalité et de dissuasion

Mécanisme de sanctions

- Enquêtes, contrôles, audits
- Avertissement, mise en demeure
- Suspension de flux et/ou certification
- Amende administrative
  - 20 millions € ou 4% du CA annuel mondial
  - Principe de proportionnalité et de dissuasion

Périmètre d'application

- les entreprises privées ou publiques qui :
  - proposent des biens et services dans le marché de l'UE
  - collectent et traitent des données personnelles sur les résidents de l'UE
- sont concernées les entreprises hors-UE !

Périmètre d'application

- les entreprises privées ou publiques qui :
  - proposent des biens et services dans le marché de l'UE
  - collectent et traitent des données personnelles sur les résidents de l'UE
- sont concernées les entreprises hors-UE !

« Sphère de sécurité »

- RGPD et transferts trans-frontaliers
- Qualification des prestataires : sous-traitant ou responsable de traitement ?
- *Privacy Shield* : passerelle droit Eu/droit US
  - personnalité vs. donnée marchande

« Sphère de sécurité »

- RGPD et transferts trans-frontaliers
- Qualification des prestataires : sous-traitant ou responsable de traitement ?
- *Privacy Shield* : passerelle droit Eu/droit US
  - personnalité vs. donnée marchande

Max Schrems

- Citoyen autrichien, étudiant en droit
- Fondateur de http://europe-v-facebook.org/
  - 2011 : 1200 pages de données Facebook pour 22 plaintes
  - 2013 : révélations de E. Snowden sur la surveillance de masse de la NSA, nouvelle plainte contre Facebook (à Dublin, siège Eu)
  - 06/10/2015 : invalidation du *Safe Harbor* par la Cour européenne de justice
  - Depuis le 01 août 2016 : *Privacy Shield*…

https://www.privacyshield.gov

Max Schrems

- Citoyen autrichien, étudiant en droit
- Fondateur de http://europe-v-facebook.org/
  - 2011 : 1200 pages de données Facebook pour 22 plaintes
  - 2013 : révélations de E. Snowden sur la surveillance de masse de la NSA, nouvelle plainte contre Facebook (à Dublin, siège Eu)
  - 06/10/2015 : invalidation du *Safe Harbor* par la Cour européenne de justice
  - Depuis le 01 août 2016 : *Privacy Shield*…

https://www.privacyshield.gov

Idées-forces

- Minimisation (*need-to-know*)
  - proportionnalité et finalités légitimes
  - fragmentation
- Souveraineté
  - péremption, notification de transfert/usage
  - anonymisation (neutralisation)

Idées-forces

- Minimisation (*need-to-know*)
  - proportionnalité et finalités légitimes
  - fragmentation
- Souveraineté
  - péremption, notification de transfert/usage
  - anonymisation (neutralisation)

RGPD, 5 ans après

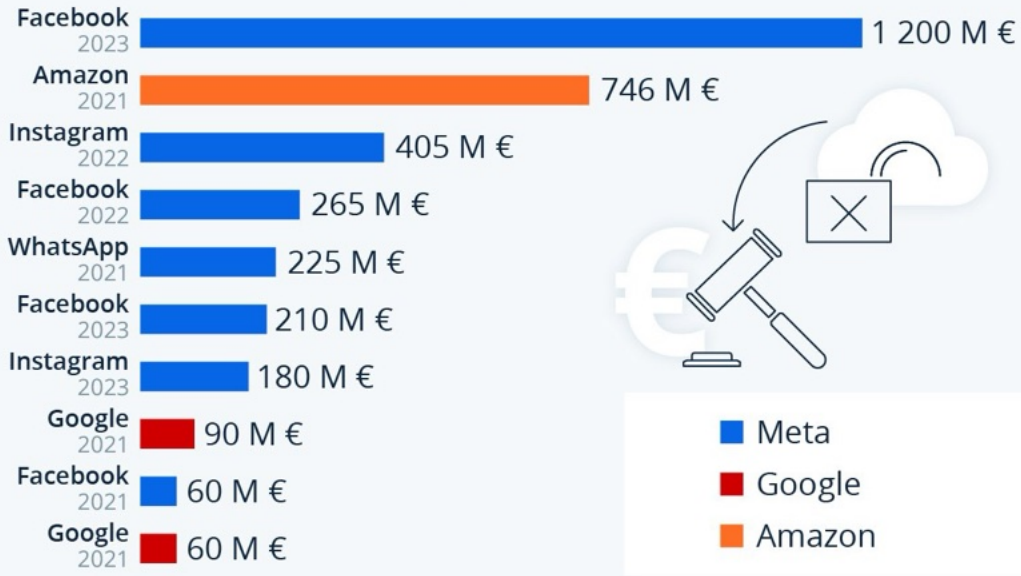
- 70+ pays ont adoptés des textes similaires
- 2+B€ amendes administratives dans l'UE
- Quelques grosses affaires, mais encore peu de petites amendes
- Bilan CNIL 2022 en France :
  - 21 sanctions (1+M€ d'amendes)
  - 147 mises en demeure

RGPD, 5 ans après

- 70+ pays ont adoptés des textes similaires
- 2+B€ amendes administratives dans l'UE
- Quelques grosses affaires, mais encore peu de petites amendes
- Bilan CNIL 2022 en France :
  - 21 sanctions (1+M€ d'amendes)
  - 147 mises en demeure

RGPD : Meta cumule les amendes monstres

Plus grosses amendes infligées pour violation des données personnelles dans les pays de l'UE (non-respect du RGPD)



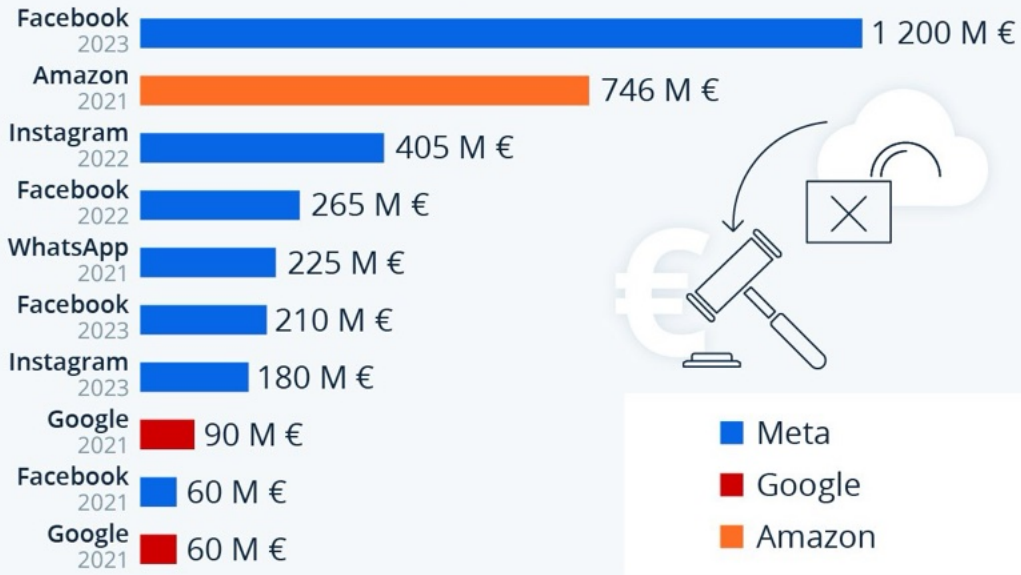
En date du 23 mai 2023.

Sources : CMS GDPR Enforcement Tracker, European Data Protection Board

statista

RGPD : Meta cumule les amendes monstres

Plus grosses amendes infligées pour violation des données personnelles dans les pays de l'UE (non-respect du RGPD)



En date du 23 mai 2023.

Sources : CMS GDPR Enforcement Tracker, European Data Protection Board

statista

Baromètre

Data Legal Drive

RGPD EN 2021

1 entreprise sur 2 a un bon niveau de conformité RGPD

La **situation sanitaire aura finalement été favorable à la gouvernance des données personnelles** au sein entreprises et des organismes publics qui sont près la moitié à estimer avoir atteint un niveau de complétude supérieur à 70%.

47%
Des sondés estiment avoir atteint un niveau de complétude supérieur à 70%.

37%
Des structures auraient un taux de complétude inférieur à 50%.

Au sein des entreprises, le RGPD est vécu comme :

Toutefois, des actions claires, des processus et gammes opératoires transverses et une organisation efficace **nécessitent encore d'être améliorés**, automatisés, accompagnés afin d'atteindre un meilleur niveau de conformité.

48% des interrogés perçoivent le RGPD comme une démarche transverse, permanente et vertueuse.

Baromètre

Data Legal Drive

RGPD EN 2021

1 entreprise sur 2 a un bon niveau de conformité RGPD

La **situation sanitaire aura finalement été favorable à la gouvernance des données personnelles** au sein entreprises et des organismes publics qui sont près la moitié à estimer avoir atteint un niveau de complétude supérieur à 70%.

47%
Des sondés estiment avoir atteint un niveau de complétude supérieur à 70%.

37%
Des structures auraient un taux de complétude inférieur à 50%.

Au sein des entreprises, le RGPD est vécu comme :

Devoir de transparence & marque de respect 22%
Contrainte technique et/ou juridique 27%
Obligation règlementaire juridique 32%

Toutefois, des actions claires, des processus et gammes opératoires transverses et une organisation efficace **nécessitent encore d'être améliorés**, automatisés, accompagnés afin d'atteindre un meilleur niveau de conformité.

48% des interrogés perçoivent le RGPD comme une démarche transverse, permanente et vertueuse.

IA et RGPD

- EU AI Act : 08 déc. 2023, accord de principe
  - 4 niveaux de risque
- En France, la CNIL s'empare du sujet :
  - création d'un service spécifique pour l'IA (jan. 2023)
  - l'IA est « RGPD-compatible » (11 oct. 2023) :
    - finalité, minimisation, conformité de la collecte, conservation limitée, sécurité
  - plan d'action : LLM, reco. faciale, règlement EU

IA et RGPD

- EU AI Act : 08 déc. 2023, accord de principe
  - 4 niveaux de risque
- En France, la CNIL s'empare du sujet :
  - création d'un service spécifique pour l'IA (jan. 2023)
  - l'IA est « RGPD-compatible » (11 oct. 2023) :
    - finalité, minimisation, conformité de la collecte, conservation limitée, sécurité
  - plan d'action : LLM, reco. faciale, règlement EU

Où en sommes-nous ?

A. À propos des données personnelles
B. Cadre législatif et réglementaire
    1. Données à caractère personnel
    2. RGPD
    3. Anonymisation

Où en sommes-nous ?

A. À propos des données personnelles
B. Cadre législatif et réglementaire
    1. Données à caractère personnel
    2. RGPD
    3. Anonymisation

Pourquoi anonymiser ?

- Donnée « dé-personnalisée » = donnée neutralisée vis-à-vis du RGPD
  - durée de conservation étendue
  - usage détourné : stats, test, prédiction, vente
  - réglementation : publication des décisions de justice
  - Open data
  - dé-responsabilisation : sous-traitance et transfert trans-frontalier

Open data et protection de la vie privée, Rapport d'information de MM. Gaëtan GORCE et François PILLET, fait au nom de la commission des lois du Sénat, n° 469 (2013-2014) - 16 avril 2014

Pourquoi anonymiser ?

- Donnée « dé-personnalisée » = donnée neutralisée vis-à-vis du RGPD
  - durée de conservation étendue
  - usage détourné : stats, test, prédiction, vente
  - réglementation : publication des décisions de justice
  - Open data
  - dé-responsabilisation : sous-traitance et transfert trans-frontalier

Open data et protection de la vie privée, Rapport d'information de MM. Gaëtan GORCE et François PILLET, fait au nom de la commission des lois du Sénat, n° 469 (2013-2014) - 16 avril 2014

De quoi parle-t-on ?

- Anonymat : "—possibilité de suivre une personne unique dans la durée avec— impossibilité de connaître sa véritable identité" (Lexique AFCDP)
- Anonymisation : rupture définitive du lien entre une donnée et une personne

De quoi parle-t-on ?

- Anonymat : "—possibilité de suivre une personne unique dans la durée avec— impossibilité de connaître sa véritable identité" (Lexique AFCDP)
- Anonymisation : rupture définitive du lien entre une donnée et une personne

Mais aussi...

- Pseudonymat : anonymat *réversible* avec une responsabilité juridique

  e-Commerce : pseudos pour particuliers uniquement (LCEN, 2004)

- Hétéronymat : plusieurs identités autonomes (76 écrivains = F. Pessoa)
- Homonymat (?) : banalisation de l'identité

Mais aussi...

- Pseudonymat : anonymat *réversible* avec une responsabilité juridique

  e-Commerce : pseudos pour particuliers uniquement (LCEN, 2004)

- Hétéronymat : plusieurs identités autonomes (76 écrivains = F. Pessoa)
- Homonymat (?) : banalisation de l'identité

Mais encore

- Norme ISO 15408
- Non-chaînabilité : impossibilité—pour un tiers—d'établir un lien entre différentes opérations faites par un même individu
- Non-observabilité : impossibilité—pour un tiers—de déterminer si une opération est en cours
- *pseudo* < *ano* < *non-chaîn* < *non-obs*

Mais encore

- Norme ISO 15408
- Non-chaînabilité : impossibilité—pour un tiers—d'établir un lien entre différentes opérations faites par un même individu
- Non-observabilité : impossibilité—pour un tiers—de déterminer si une opération est en cours
- *pseudo* < *ano* < *non-chaîn* < *non-obs*

Anonymiser
= Supprimer les Ids ?

- ZIP+Genre+DdN. = 87% population U.S.

Sources : L. Sweeney, *Uniqueness of Simple Demographics in the U.S. Population*, 2000
P. Golle. *Revisiting the uniqueness of simple demographics in the US population, WPES, pp. 77-80,* 2006

- DdN 2 enfants = identité de la mère (Fr)

Source : G. Trouessin, JSSI 2008

- 4 relevés GSM+temps = 95% ré-identification

Source : Y.-A. de Montjoye et al. , *Unique in the Crowd: The privacy bounds of human mobility, Nature*, 2013

Anonymiser
= Supprimer les Ids ?

- ZIP+Genre+DdN. = 87% population U.S.

Sources : L. Sweeney, *Uniqueness of Simple Demographics in the U.S. Population*, 2000
P. Golle. *Revisiting the uniqueness of simple demographics in the US population, WPES, pp. 77-80,* 2006

- DdN 2 enfants = identité de la mère (Fr)

Source : G. Trouessin, JSSI 2008

- 4 relevés GSM+temps = 95% ré-identification

Source : Y.-A. de Montjoye et al. , *Unique in the Crowd: The privacy bounds of human mobility, Nature*, 2013

Révélation d'identité (bis)

- Log de requêtes AOL (2006)
- Concours Netflix (2006-2009) : usage détourné des données
  - https://netflixprize.com/
  - 100M recommandations entre 1999-2005 de 500.000 clients sur 10M titres
  - Croisement avec IMDB : films + notes et #id
  - avec 8 reco. + date (+/- 3 j.) : 96% abonnés identifiés
  - lauréat : *BellKor's Pragmatic Chaos* avec +10% de reco. pertinentes

Source : A. Narayanan, V. Shmatikov.
How To Break Anonymity of the Netflix Prize Dataset, 2006

Révélation d'identité (bis)

- Log de requêtes AOL (2006)
- Concours Netflix (2006-2009) : usage détourné des données
  - https://netflixprize.com/
  - 100M recommandations entre 1999-2005 de 500.000 clients sur 10M titres
  - Croisement avec IMDB : films + notes et #id
  - avec 8 reco. + date (+/- 3 j.) : 96% abonnés identifiés
  - lauréat : *BellKor's Pragmatic Chaos* avec +10% de reco. pertinentes

Source : A. Narayanan, V. Shmatikov.
How To Break Anonymity of the Netflix Prize Dataset, 2006

Moyens techniques

- Chiffrement : masque la donnée de façon réversible. Brique de sécurité fondamentale pour le stockage et les échanges
- Pseudonymisation : attribution impossible sans une « clé », conservée séparément
- Anonymisation : aucune identification possible

Moyens techniques

- Chiffrement : masque la donnée de façon réversible. Brique de sécurité fondamentale pour le stockage et les échanges
- Pseudonymisation : attribution impossible sans une « clé », conservée séparément
- Anonymisation : aucune identification possible

Opinion 05/2014 du G29

- Techniques d'anonymisation

	Is Singling out still a risk?	Is Linkability still a risk?	Is Inference still a risk?
Pseudonymisation	Yes	Yes	Yes
Noise addition	Yes	May not	May not
Substitution	Yes	Yes	May not
Aggregation or K-anonymity	No	Yes	Yes
L-diversity	No	Yes	May not
Differential privacy	May not	May not	May not
Hashing/Tokenization	Yes	Yes	May not

Table 6. Strengths and Weaknesses of the Techniques Considered

Opinion 05/2014 du G29

- Techniques d'anonymisation

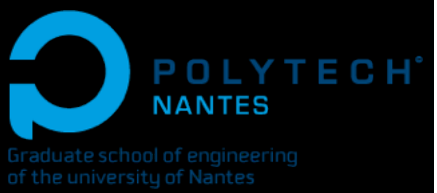
	Is Singling out still a risk?	Is Linkability still a risk?	Is Inference still a risk?
Pseudonymisation	Yes	Yes	Yes
Noise addition	Yes	May not	May not
Substitution	Yes	Yes	May not
Aggregation or K-anonymity	No	Yes	Yes
L-diversity	No	Yes	May not
Differential privacy	May not	May not	May not
Hashing/Tokenization	Yes	Yes	May not

Table 6. Strengths and Weaknesses of the Techniques Considered

91

Merci

contact : guillaume.raschia@univ-nantes.fr



Merci

contact : guillaume.raschia@univ-nantes.fr