Institut Mines-Télécom

RGPD : ANONYMISATION ET PSEUDONYMISATION

CLAIRE LEVALLOIS-BARTH,
TELECOM PARIS
COORDINATRICE DE LA CHAIRE VALEURS ET POLITIQUES DES INFORMATIONS PERSONNELLES

CERNA - 3 JUILLET 2019

IMT Atlantique Bretagne-Pays de la Loire École Mines-Télécom
IMT Lille Douai École Mines-Télécom IMT-Université de Lille
IMT Mines Albi-Carmaux École Mines-Télécom
IMT Mines Alès École Mines-Télécom
Institut Mines-Télécom Business School
MINES Saint-Étienne Une école de l'IMT
TELECOM ParisTech Une école de l'IMT
TELECOM SudParis Une école de l'IMT
EURECOM
InSIC
MINES ParisTech
MINESnancy
ARMINES
ensiie PARIS-ÉVRY

4 années d'intenses négociations

- Proposition de la Commission européenne du 25 janvier 2012
- Adopté le 27 avril 2016
- Entrée en vigueur le 25 mai 2018

Un texte complexe

- 99 articles et 173 considérants
- Une cinquantaine de déclinaisons possibles et des lois nationales pour les activer

Une évolution, pas une révolution

- Loi Informatique et Libertés (1978)
- Directive (UE) 1995/46/CE de 1995
- Ordonnance n° 2018-1125 du 12 décembre 2018 prise en application de l'article 32 de la loi n° 2018-493 du 20 juin 2018 relative à la protection des données personnelles

Art. 4-1. RGPD : donnée à caractère personnel

- Toute information se rapportant à une personne physique identifiée ou identifiable
- **Personne identifiable** : personne 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

Nom prénom

06 56 89 90

2 89 751 57 093 57

7 décembre 2018 Montparnasse 8h05
7 décembre 2018 Corvisart 8h45
7 décembre 2018 Corvisart 17h30
7 décembre 2018 Montparnasse 7h57
8 décembre 2018 Corvisart 8h40
8 décembre 2018 Corvisart 17h05
8 décembre 2018 Chatelet 17h45
8 décembre 2018 Chatelet 18h45
8 décembre 2018 Montparnasse 19h37

Considérant 26 RGPD

- Information ne concernant pas une personne physique identifiée ou identifiable
- Données à caractère personnel rendues anonymes de telle manière que la personne concernée ne soit pas ou plus identifiable

Pour la personne concernée

- Réduction des risques d'abus et respect des droits et libertés fondamentaux

Pour l'organisme : non application du RGPD

- Sécurisation de l'exploitation des données personnelles

Application du RGPD

- Anonymisation ultérieure
- Anonymisation à bref délai
  - Traitement de données même si l'anonymisation suit immédiatement la collecte
  - Réduction des risques pour les droits et libertés fondamentaux de la personne concernée

Irréversibilité très compliquée

- Détruire **toute possibilité** de pouvoir identifier à quelle personne appartiennent les données personnelles
- Suppression ou modification des données personnelles afin de créer un ensemble de données vraiment anonymes
- Tout en conservant suffisamment d'informations sous-jacentes pour les besoins de la tâche concernée
- Construction au cas par cas

Avis du groupe de travail de l'article 29 (G29) sur les techniques d'anonymisation du 10 avril 2014

- 3 critères pour évaluer une bonne solution d'anonymisation
  1. **L'individualisation** : est-il toujours possible d'isoler un individu ?
  2. **La corrélation** : est-il possible de relier entre eux des ensembles de données distincts concernant un même individu ?
  3. **L'inférence** : peut-on déduire de l'information sur un individu ?

- Si les 3 critères sont atteints : l'ensemble de données est *a priori* anonyme

- Si au moins un des 3 critères n'est pas respecté : l'ensemble de donnes ne pourra être considéré comme anonyme qu'à la suite d'une analyse détaillée des risques de ré-identification

Article 4-5 RGPD : pseudonymisation

- « Le traitement de données à caractère personnel de telle façon que celles-ci ne puissent plus être attribuées à une personne concernée précise sans avoir recours à des informations supplémentaires
- Pour autant que ces informations supplémentaires soient conservées séparément et soumises à des mesures techniques et organisationnelles afin de garantir que les données ne sont pas attribuées à une personne physique identifiée ou identifiable »

Application du RGPD

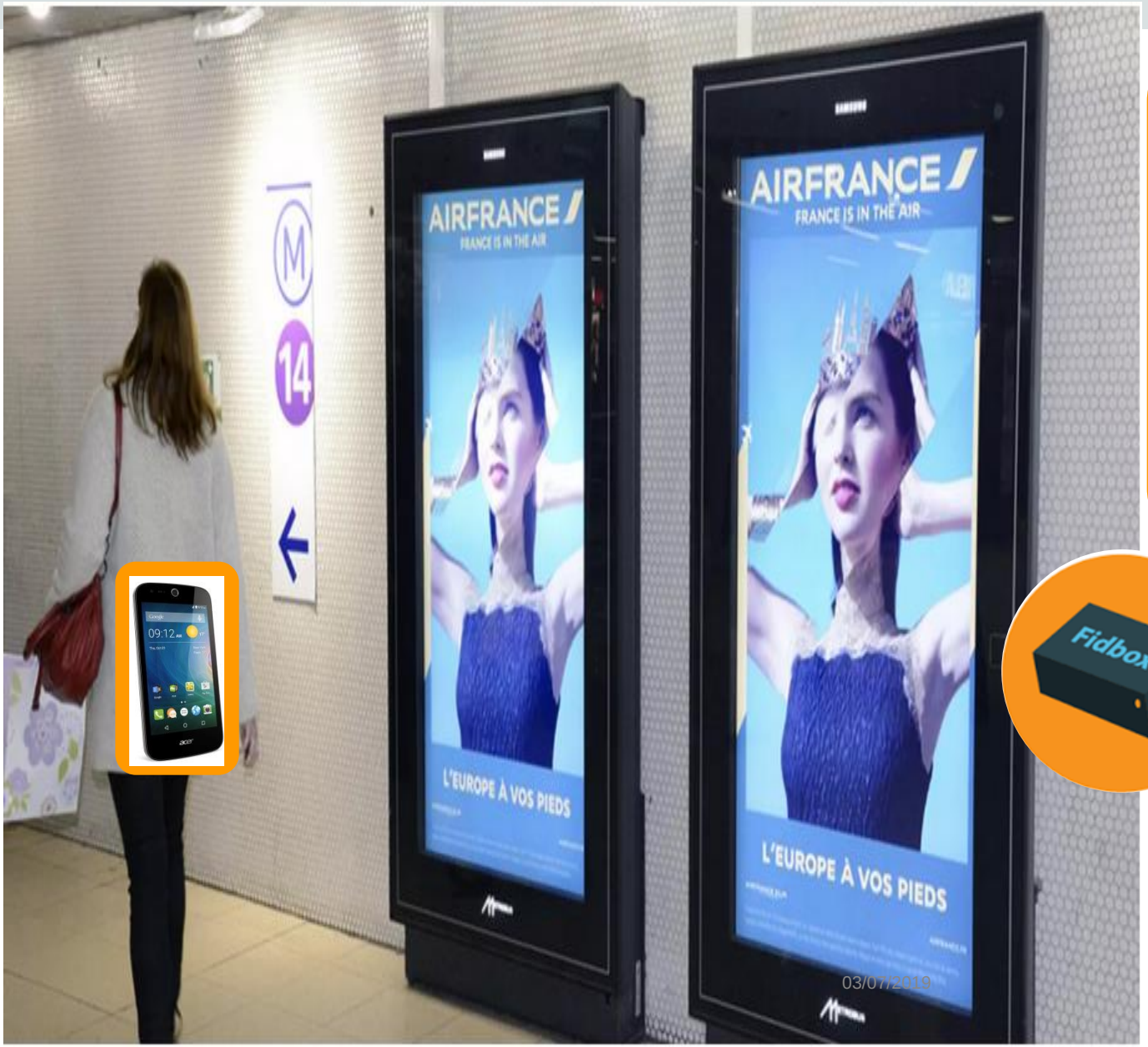
- « Les données personnelles qui ont fait l'objet d'une pseudonymisation et qui pourraient être attribuées à une personne physique par le recours à des informations supplémentaires devraient être considérées comme des informations concernant une personne physique identifiable » (cons. 26 RGPD)

Limiter efficacement le risque de ré-identification directe

- Remplacer un identifiant (plus généralement des données personnelles) par un pseudonyme
  - Tout en permettant l'étude de corrélations en cas de besoin particulier
  - Etre vigilant car une ré-identification peut intervenir à partir d'informations partielles

Exemple de pseudonymisation : CNIL fiche 10 Sécurité

- Générer une clé secrète longue et difficile à mémoriser (une combinaison de caractères aléatoires) qui permet d'établir un lien entre les données personnelles originales des personnes
  - Appliquer une fonction à sens unique sur les données (par exemple, un algorithme de hachage à clé secrète tel HMAC)
- Assurer la confidentialité de cette clé (notamment tracer les accès à cette clé)
- En l'absence de besoin de ré-identification efficace, supprimer la clé secrète pour diminuer le risque de ré-identification



Finalité
tester une méthodologie d'estimation quantitative des flux piétons (optimisation du prix de l'espace publicitaire)

Boitier wifi

- Adresse MAC émise par la carte Wi-Fi
- Horaire de détection de l'adresse MAC
- Puissance d'émission du signal (distance carte Wi-Fi – boitier)

Transmission des données toutes les 2 minutes

- Adresse MAC de la carte Wi-Fi du smartphone tronquée du dernier demi-octet, en utilisant un sel propre à la JCDecaux, puis hachée
- Enregistrement

Serveur

- A la fin de l'expérimentation, les données brutes seraient agrégées à des fins d'analyse pour une estimation
  - Du nombre de détections heure par heure (moyennées),
  - Du nombre de détections uniques par jour/par semaine/par mois
  - Des schémas de mobilité

Demande d'autorisation à la Cnil

- A l'époque, article L. 581-9, l'alinéa 4 du code de l'environnement
  - « Tout système de mesure automatique de l'audience d'un dispositif publicitaire ou d'analyse de la typologie ou du comportement des personnes passant à proximité d'un dispositif publicitaire est soumis à autorisation de la Cnil »

« Ce procédé d'anonymisation constitue une garantie suffisante propre à assurer le respect des droits et libertés fondamentaux »

Notamment à permettre la délivrance de l'information limitée

- Conformément à l'époque à l'article 32-IV de la loi Informatique et Libertés qui limitait les informations à délivrer si les données personnelles étaient appelées à faire l'objet à bref délai d'un procédé d'anonymisation
- Affiche A4 sur ses mobiliers publicitaires mentionnant
  - L'identité du responsable du traitement
  - La finalité poursuivie par le traitement

Le procédé ne constitue pas une technique d'anonymisation

- « Notamment « JCDecaux est en mesure de rejouer le procédé de chiffrement, cette société utilisant un sel qui lui est propre et connu, et en raison du faible taux de collision proposé »
- La finalité du traitement viserait à compter le nombre de terminaux mobiles détectés, mais aussi … la récurrence de passage à proximité d'un mobilier publicitaire JCDecaux
  - Estimer le nombre de passants, leur parcours et le nombre de fois où un même passant repasse sur l'esplanade de la Défense sur une période donnée

Le procédé est une technique de pseudonymisation

- Permettant toujours la corrélation et l'inférence
- Refuse d'accorder son autorisation car l'information des personnes est insuffisante en l'absence de mention de la possibilité
  - D'accéder à ses données personnelles
  - De s'opposer au traitement.

Délibération de la CNIL n°2015-255 du 16 juillet 2015

- « Pour qu'une solution d'anonymisation soit efficace, elle doit empêcher
  1. Toutes les parties d'isoler un individu dans un ensemble de données,
  2. De relier entre eux 2 enregistrements dans un ensemble de données (ou dans 2 ensembles de données séparés)
  3. Et de déduire des informations de cet ensemble de données »
- 3 critères : individualisation, corrélation, inférence

Arrêt du Conseil d'Etat le 8 février 2017 : refuse d'annuler la délibération de la Cnil

- Une donnée personnelle « ne peut être regardée comme rendue anonyme que lorsque l'identification de la personne concernée, directement ou indirectement, devient impossible que ce soit par le responsable du traitement ou par un tiers.
- Tel n'est pas le cas lorsqu'il demeure possible d'individualiser une personne ou de relier entre elles des données résultant de deux enregistrements qui la concernent ».

Réduire les risques

- « La pseudonymisation … peut réduire les risques pour les personnes concernées et aider les responsables du traitement et les sous-traitants à remplir leurs obligations en matière de protection des données » (Cons. 28 RGPD)

Selon le contexte et les risques spécifiques, participe à

- La **minimisation des données** : données adéquates, pertinentes et limitées à ce qui est nécessaire au regard des finalités pour lesquelles elles sont traitées
- La **protection des données personnelles dès la conception**
- La **sécurité** du traitement
  - Mise en place d'un niveau de sécurité adapté au risque en réduisant la probabilité que des personnes soient identifiées en cas de violation de données
- Principe de **responsabilité** : permet au responsable de traitement de démontrer qu'il respect le RGPD

Article 89-1 RGPD

- Pour ce type de finalité, chaque fois que ces finalités peuvent être atteintes par un traitement ultérieur ne permettant pas ou plus l'identification des personnes concernées, il convient de procéder de cette manière

La nécessité d'une approche dynamique

- Plus le volume de données croît, plus les risques de ré-identification par recoupement sont importants
- Revoir régulièrement le procédé d'anonymisation et/ou de pseudonmisation au regard des avancées technologiques
- Mettre en place une organisation interne appropriée