

# Chapitre 1

## Identifier les traitements

### 1. Introduction

Une donnée est la représentation d'une information sous une forme permettant de la stocker, de la transmettre et de l'analyser. Matières premières de la révolution numérique, les datas font émerger de nouveaux rapports entre les citoyens, les états et les entreprises. Selon Serge Abiteboul « une donnée est une description élémentaire d'une réalité. C'est par exemple une observation ou une mesure. La donnée est dépourvue de tout raisonnement, suppositions, constatations, probabilités. Étant indiscutable ou indiscutée, elle sert de base à une recherche ou à un examen quelconque ». (Sciences des données. De la logique du premier ordre à la Toile, leçon inaugurale à la chaire Informatique et sciences numériques du Collège de France, prononcée le jeudi 8 mars 2012). La croissance vertigineuse de leur production justifie qu'on s'intéresse tout particulièrement à elle.

Constituent des données personnelles toutes les informations « se rapportant à une personne physique identifiée ou identifiable ».

## 2. Comment identifier une donnée personnelle ?

Constitue une donnée à caractère personnel « toute information se rapportant à une personne physique identifiée ou identifiable ; 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 (UE) n° 2016/679, 4, 1).

### Que retenir de cette définition ?

– Première observation : ne peuvent être des données à caractère personnel que les données permettant d'identifier des **personnes physiques**. Remarquons que la « personne physique identifiée ou identifiable » est dénommée « personne concernée » dans le RGPD. Dans le cadre de cet ouvrage nous reprendrons cette dénomination.

De ce fait, les personnes morales, y compris les associations, ne sont pas concernées par le RGPD.

En revanche, les nom et prénom du dirigeant d'une entreprise sont des données à caractère personnel concernées par le RGPD.

– Deuxième observation : le caractère personnel d'une donnée est intimement lié à son pouvoir d'identification d'une personne. À ce titre, il est important de constater que ce pouvoir identifiant est largement entendu.

– Il peut être « **intrinsèque ou direct** » : une donnée permettant, seule, d'identifier une personne physique est une donnée à caractère personnel.

– Il peut se déduire de la combinaison avec une autre donnée : pouvoir d'identification « **indirect** ». Par exemple, le numéro d'un bureau ou d'une carte de cantine constitue une donnée indirectement identifiante.

– Troisième observation : l'utilisation de l'adverbe « notamment » dans la définition d'une donnée à caractère personnel permet d'indiquer que les caractéristiques énumérées dans l'article 4 du règlement ne sont certainement pas exhaustives. Ainsi toute information permettant de reconnaître une personne est une donnée à caractère personnel et ce quel que soit son format tel qu'une photographie, une vidéo, une correspondance, etc.

- Quatrième observation : il ressort du considérant 14 que le règlement ne couvre pas « le traitement de données à caractère personnel qui concerne les personnes morales, et en particulier les entreprises dotées de la personnalité juridique, y compris le nom, la forme juridique et les coordonnées de la personne morale ».
- Cinquième observation : le considérant 27 du règlement précise qu'il ne s'applique pas aux données à caractère personnel des personnes décédées.

## LES NOTIONS CLÉS



### Donnée personnelle

Toute information relative à une personne physique identifiée ou qui peut être identifiée, **directement** ou **indirectement**



### Fichier de données

Tout ensemble structuré et stable de données à caractère personnel



### Traitement de données

Toute opération ou tout ensemble d'opérations, automatisés ou non, portant sur de telles données, quel que soit le procédé utilisé

**FOCUS : l'adresse IP**

L'adresse IP (pour *Internet Protocol*) est un numéro d'identification qui est attribué de façon permanente ou provisoire à tout appareil connecté au protocole Internet. À ce titre, l'adresse IP est un élément clé du système de routage des messages sur Internet.

Le 19 octobre 2016, la Cour de Justice de l'Union Européenne (CJUE) a dû se prononcer sur le fait de devoir considérer ou non l'adresse IP comme une donnée à caractère personnel (CJUE, 19 oct. 2016, C-582/14). Ainsi, selon la CJUE, « *une adresse IP dynamique enregistrée par un fournisseur de services de médias en ligne à l'occasion de la consultation par une personne d'un site Internet que ce fournisseur rend accessible au public constitue, à l'égard dudit fournisseur, une donnée à caractère personnel au sens de l'article 2, sous a), de la directive 95/46/CE du Parlement européen et du Conseil, du 24 octobre 1995, lorsqu'il dispose de moyens légaux lui permettant de faire identifier la personne concernée grâce aux informations supplémentaires dont dispose le fournisseur d'accès à Internet de cette personne* » (CJUE, 19 oct. 2016, C-582/14, cons. 49). La Cour de cassation a également rendu, très peu de temps après, un arrêt dans le même sens (Cass, Civ. 1ère, 3 nov. 2016, n° 15-22595). Sans détour, elle affirme que « *les adresses IP, qui permettent d'identifier indirectement une personne physique, sont des données à caractère personnel* ».

**Ce qu'il faut retenir : la collecte d'adresses IP constitue un traitement de données à caractère personnel qui devra, dès l'entrée en vigueur du RGPD, être mentionné dans le registre des traitements.**

Parmi les informations collectées, un grand nombre pourront être qualifiées de données personnelles. Il en est ainsi des données de localisation, des identifiants tels que les adresses IP, les adresses de messagerie, les étiquettes d'identification par radiofréquences. Une donnée personnelle n'est pas une information à proprement parler identifiable, il peut s'agir d'éléments semblant neutres mais qui, mis bout à bout, permettent d'identifier une personne physique. C'est par exemple le cas des informations stockées par les témoins de connexion comme les cookies. Par conséquent, les données à caractère personnel ne sont librement disponibles ni dans leur accès ni dans leur traitement. L'accord de l'internaute est nécessaire pour stocker les informations le concernant sur un cookie et les exploiter.

Les données trouvées, indexées, stockées par les moteurs de recherche, et mises à la disposition de leurs utilisateurs, sont également des données à caractère personnel (CJUE, affaire C-131/12 du 13/05/2014).

### 3. Les cas où la donnée personnelle perd son pouvoir identifiant

Comme nous l'avons vu, ce qui caractérise une donnée « personnelle », c'est son pouvoir identifiant. Autrement dit, c'est la capacité de la donnée à isoler un individu parmi d'autres de façon suffisamment précise pour que, directement ou par recoupement, on puisse remonter jusqu'à l'individu personne physique.

Cependant il est des cas où une donnée, qui originellement caractérisait un individu, perd son pouvoir identifiant, ou du moins que ce dernier devienne moins fort et donc moins sujet à des recoupements. Il s'agit de l'anonymisation et de la pseudonymisation.

L'anonymisation est un « traitement » de données à caractère personnel dont le but consiste à empêcher irréversiblement l'identification de la personne concernée. Il n'existe dès lors plus aucune table de correspondance permettant de remonter jusqu'à la personne concernée. Lorsqu'une donnée est anonymisée, elle perd complètement son caractère identifiant, ce qui signifie qu'elle n'est plus soumise aux règles protectrices du RGPD. Aucune réversibilité n'est possible. Cette donnée, devenue une information « neutre », pourra alors être traitée comme n'importe qu'elle autre information.

Il est cependant difficile en pratique de créer un ensemble de données véritablement « anonymes » en conservant suffisamment d'informations pour que ces données soient toujours exploitables dans une tâche déterminée. C'est la raison pour laquelle de nombreux acteurs optent pour une « pseudonymisation » des données personnelles.

La pseudonymisation est un « traitement » de données qui consiste à remplacer un identifiant par un pseudonyme. Il existe donc une table de correspondance et une opération de réversibilité est possible. Attention car la donnée « pseudonymisée » est toujours considérée comme une donnée à caractère personnel. Cependant les risques pour la personne concernée en cas de fuite seront en principe moindres, ce qui fait de la pseudonymisation une mesure de sécurité efficace (voir infra. « La pseudonymisation »).

### **À retenir**

La pseudonymisation des données se différencie de l'anonymisation en ce qu'elle n'a pas pour résultat de supprimer toutes les informations identifiantes et de rendre pratiquement impossible la réidentification de la personne concernée. Concrètement, la pseudonymisation consiste à remplacer l'identifiant initial par un autre identifiant, arbitraire celui-là. Ce procédé, repose notamment sur l'utilisation d'une table de correspondance secrète, ou sur un algorithme de chiffrement, ou encore sur le recours à la fonction dite de « hachage » des données.

Ainsi, sa particularité est d'empêcher que des données personnelles puissent être attribuées à une personne déterminée sans qu'il faille recourir à des informations supplémentaires. Toutefois, la mise en œuvre d'un tel procédé suppose la réunion de deux conditions : les données personnelles doivent être conservées séparément des clés de réidentification et elles doivent être soumises à des mesures garantissant leur confidentialité.

## **4. Interdiction de traitement de certaines données personnelles**

Le RGPD interdit le traitement de certaines données personnelles : ce sont celles qui révèlent l'origine raciale ou ethnique des personnes, leur opinion politique, leurs orientations religieuses ou philosophiques, leur appartenance syndicale, leur santé, leur vie ou leurs orientations sexuelles. Cette liste doit être complétée par les « données génétiques » ainsi que par les « données biométriques aux fins d'identifier une personne physique de manière unique ».