

LE MOT " INFORMATION " EN INFORMATIQUE ET LIBERTES ETUDE TERMINOLOGIQUE

Anne BRYGOO

Dans le domaine relatif aux fichiers et aux traitements d'informations nominatives le terme d'"information" est très fréquemment utilisé. Il nous a semblé intéressant d'étudier les diverses acceptions de ce mot et des qualificatifs qui lui sont régulièrement associés.

Les motivations qui sont à l'origine de ce travail tiennent à la constatation qu'en informatique (comme dans toute autre discipline) les mots prennent petit à petit une signification bien spécifique. L'oubli des différents sens d'un même mot rend difficile les échanges interdisciplinaires. Ceci est particulièrement vrai en "Informatique et Libertés" où la signification technique de certains mots se heurte aux significations plus "grand public" de ces mêmes mots. Ainsi mettrons-nous en parallèle pour un même terme les divers sens qui ont cours suivant que l'on se situe dans les domaines de la technique informatique, du grand public ou celui du droit. En effet les traitements automatiques d'informations nominatives concernent conjointement ces trois domaines.

A plus long terme donc, l'objectif de ce travail serait une contribution pour faciliter des échanges plus approfondis et donc plus fructueux entre informaticiens, juristes, utilisateurs et usagers de l'informatique et non d'établir le lexique dépourvu d'ambiguïté d'une "novlangue".

Concrètement dans ce texte nous adopterons pour chaque mot la présentation suivante : définitions extraites de dictionnaires généraux (1, 2 et 3) et informatiques (4 et 5), suivies de commentaires sur ces définitions. Les commentaires mentionneront si nécessaire des articles de la loi "Informatique, fichiers et libertés" du 6/1/78, qui régleme les traitements d'informations nominatives, et ceux de la Convention européenne "pour la protection des personnes à l'égard du traitement

automatisé des données à caractère personnel" faite à Strasbourg le 28/1/81 et ratifiée le 15/11/1985. Quand cela sera nécessaire, nous préciserons si le qualificatif associé au mot "information" se rapporte au mot lui-même ou s'il fait référence implicitement au traitement qui l'utilise.

Commençons tout d'abord, par le mot "information".

Information :

D'après le Robert (1), l' "information" est :

- 1- "en droit, l'ensemble des actes qui tendent à établir la preuve d'une infraction, et à en découvrir les auteurs,
- 2- le renseignement que l'on prend sur quelqu'un, sur quelque chose, et par néologisme, l'ensemble des renseignements obtenus par quelqu'un,
- 3- le renseignement ou l'événement qu'on porte à la connaissance d'une personne, d'un public, et par néologisme l'ensemble des informations et l'action d'informer le public".

D'après le dictionnaire d'informatique Morvan (4) l'"information" est un objet à la base de la communication des connaissances et d'après le GUNGUAY-LAURET, autre dictionnaire d'informatique (5), l' "information" est la "signification que l'on attribue à une expression conventionnelle ou "donnée" de telle sorte qu'elle constitue pour l'observateur un élément de connaissance". La "donnée" est, elle, un fait "ou notion représenté sous une forme conventionnelle convenant à une communication, une interprétation ou un traitement soit par l'homme soit par des moyens automatiques" (définition ISO). La "donnée variable" est, "dans les traitements périodiques, une donnée prenant à chaque période des valeurs différentes comme par exemple le nombre d'heures travaillées" (5).

Il est important de constater qu'au mot "information" peuvent donc être associées plusieurs notions (disjointes ou non) :

- a- objet à base de communication (à prendre et/ou à donner) ;
- b- la signification, le sens, que l'on donne à cet objet ;
- c- l'attribution de cet objet à une chose, à un événement, à un individu ;
- d- l'utilisation de cet objet.

En informatique, la distinction donnée-information, correspondant à la distinction forme-sens ou objet-signification (a et b), ne fait pas l'unanimité. Par contre dans les applications de l'informatique, au mot "information" est associé non seulement l'objet mais aussi la signification que l'on donne à cet objet.

L'attribution de l' "information" (c) est, dans le domaine qui nous intéresse ici, relative à un individu.

L'utilisation de l' "information" (d) doit être faite pour servir la finalité déclarée du traitement automatisé. Cette liaison "information"-finalité du traitement, qui est une notion primordiale en "Informatique et Libertés", tend à disparaître avec le développement des banques de données. En effet, ces dernières, contrairement aux fichiers, sont conçues techniquement pour permettre tous les traitements possibles sur des "données". Et un marché commence à se créer où des "données" sont commercialisées en tant que telles, indépendamment du contexte d'où elles sont extraites et indépendamment de la finalité du traitement qui les utilisera (Cf. les qualificatifs "nécessaire", "utile" et "collectée").

Ayant précisé la notion d'"information" voyons maintenant les qualificatifs associés à ce terme.

Absente

Information "qui n'est pas présente", dans le fichier ou la banque de données.

Cette absence peut concerner l'ensemble des individus du fichier, l'information n'a alors pas été prévue à la création du fichier. Cette absence d'information peut aussi concerner un seul individu. Elle peut alors être considérée parfois comme une information en soi. Par exemple, l'absence de références favorables pour un locataire, dans un fichier des locataires, peut être interprétée comme une référence défavorable.

Anonyme, anonymisée :

Est anonyme "celui dont on ignore le nom" ou celui "qu'on n'a pas pu identifier" (2).

En informatique, c'est habituellement le premier sens qui est considéré et cela n'est pas équivalent au deuxième sens ; en effet dans un fichier non nominatif il peut être quand même possible d'identifier un individu avec un certain nombre de critères judicieusement choisis. C'est

pour cette raison que l'article 4 de la loi du 6/1/78 stipule qu'une information est nominative si elle permet directement ou non l'identification d'une personne physique. (Cf. nominative).

Archivée :

Est archivé "ce qui est enregistré sur des documents d'archives", les archives étant "l'ensemble de documents hors d'usage courant, rassemblés, répertoriés et conservés pour servir à l'histoire d'une collectivité ou d'un individu" (2)(3). Par contre, l'archivage, selon le dictionnaire d'informatique (5), est considéré comme "la conservation de documents en vue d'éventuelles consultations ultérieures".

La notion "hors d'usage courant" n'est pas présente dans la deuxième définition. Par contre, bien que non explicitée, il semble, dans les faits, que la notion "servir à l'histoire d'une collectivité ou d'un individu" reste implicitement liée à l'archivage au sens informatique.

Il semblerait donc important de bien différencier la notion d'"archivage" et celle de "conservation" et de préciser, pour chaque traitement, la notion "d'usage courant" et celle de "hors d'usage courant".

Banalisée :

Est banalisé ce qui "est mis à la disposition de tous les habitants d'une commune" (1), ce qui "est rendu à l'usage d'un grand nombre" (2), et est banal ce qui "est non affecté à un emploi particulier" (5).

En fait, en statistiques et en informatique, les informations sont "banalisées" quand les noms des individus ont été supprimés pour permettre un traitement statistique de ces informations. (Cf anonyme).

Codée

Un codage s'effectue "en échangeant l'écriture courante contre les signes conventionnels d'un code" (3). Est codée "une information transcrite dans un code donné (ensemble de règles de présentation des données)" (5).

Implicitement, une information "codée" est une information illisible ou accessible seulement à un nombre restreint d'initiés. Pour être comprise par un observateur cette information doit être retranscrite, décodée. L'article 35 de la loi indique que le titulaire du droit d'accès doit

obtenir communication des informations en langage clair, donc non "codée", conforme au contenu des enregistrements.

Il faut aussi remarquer que l'information a pu être entrée en machine sous la forme de texte intégral (par exemple, dans un dossier médical, la description détaillée des troubles du sommeil), ou bien a dû, avant introduction, subir une transformation pour faciliter ou rendre possible le traitement automatique ultérieur. Dans ce dernier cas il faut effectuer un choix parmi des catégories prédéfinies (de troubles du sommeil dans notre exemple). On parle aussi dans ce cas de codage de l'information.

Collectée

Collecter, c'est "recueillir, rassembler toutes sortes d'éléments concrets ou abstraits en vue d'une action précise" (3). Collecter consiste aussi à "rechercher et à recueillir des informations de nature spécifique, à les regrouper et à les conserver sur un support physique quelconque en vue de leur traitement par ordinateur" (4).

Dans toutes ces définitions la finalité de la collecte est bien présente. Cependant l'évolution actuelle, comme nous l'avons déjà dit précédemment, est que les données, en tant que telles, peuvent être collectées et commercialisées indépendamment d'une finalité de traitement précise ce qui est contraire à l'esprit de la loi.

L'article 27 de la loi précise les conditions selon lesquelles les personnes auprès desquelles sont recueillies des informations nominatives doivent être informées. Mais il est à noter qu'avec le développement des cartes nominatives à mémoire de nombreuses informations sont souvent prises à la source puis mémorisées sans que l'utilisateur ait pleinement conscience des divers traitements possibles qui peuvent être effectués sur les informations qui le concernent. Par exemple, dans le cas du retrait d'argent, si la finalité - tenue de compte - est évidente, la possibilité de reconstituer les déplacements de l'utilisateur de la carte grâce aux lieux et heures des retraits est ignorée de beaucoup.

Confidentielle

"Qui se dit ou se fait en confidence, qui contient des informations qui doivent rester secrètes" (2) ; information dont "l'utilisation est réservée à un nombre restreint d'utilisateurs" (2) (3). Dans le

dictionnaire technique (4) "la confidentialité est la sécurité visant à interdire aux personnes, n'ayant pas qualité pour le faire, l'accès à un système informatique ou aux fichiers de ce système".

La qualité de "confidentielle" concerne l'information elle-même et non le traitement de l'information. Le respect de la confidentialité nécessite au préalable la constitution de la liste des personnes autorisées à accéder à l'information et est liée très étroitement aux mesures de sécurité qui peuvent aller jusqu'au refus de mémoriser informatiquement les informations. Ce problème de sécurité en informatique est particulièrement délicat à résoudre dans le cas des banques de données où le nombre d'utilisateurs est élevé. (Cf protégée).

Conservée

Est conservé ce qui "est maintenu en bon état, préservé de l'altération, de la destruction, qui est fait pour durer (1), mis de côté en vue d'emplois ultérieurs" (3).

Cette notion doit être différenciée d'"archivée" (cf. archivée). A cette notion de conservation est implicitement attachée celle de durée de conservation et la loi, dans son article 28, indique que "sauf dispositions législatives contraires, les informations ne doivent pas être conservées sous une forme nominative au delà de la durée prévue à la demande d'avis ou à la déclaration, à moins que leur conservation ne soit autorisée par la commission".

La durée de conservation d'une information est en effet très étroitement liée à la finalité du traitement. Ainsi le nombre d'heures d'absence d'un salarié doit être conservé, le temps du traitement, pour permettre le calcul de la paye, un temps supplémentaire pour permettre contrôle et contestation ultérieurs (archivages pour certains). Au delà de ce temps réglementaire, la conservation de l'information sous forme nominative peut-être liée à d'autres possibilités de traitement (comme la sélection de personnel à licencier en priorité).

Copiée

Information reproduite (1) ; en informatique se dit d'information "provenant d'un fichier copié c'est à dire dont on a fait un double sur un support de même type ou de type différent" (5).

Les copies de sécurité offrent la possibilité de reprendre un travail en cas d'incident. Des copies peuvent aussi être faites pour vendre un fichier ou pour le conserver en vue d'une utilisation ultérieure différente.

Cryptée

L'information est souvent "codée" pour des raisons techniques alors qu'une information est "cryptée" pour renforcer la protection des informations mémorisées ou en transit sur des réseaux (cf. protégée).

Déduite

Est déduit ce qui "est obtenu comme résultat d'un raisonnement logique" (1). Ainsi une information déduite d'une ou de plusieurs autres informations est la conséquence logique qui en résulte nécessairement.

Ainsi est souvent implicitement associé à la déduction le caractère incontestable du résultat mais déduire peut aussi signifier : "conclure en raisonnant avec plus ou moins de rigueur" (3) comme par exemple déduire pour un abonné au journal d'un parti politique son adhésion au parti.

Détruite

"Altérée profondément ; l'aspect, la forme et les caractéristiques essentiels étant perdus (1) ; supprimée (3) ; effacement volontaire ou involontaire des informations contenues dans une mémoire ou sur un support" (4).

Dans le cas de destruction volontaire il peut s'agir d'un acte malveillant considéré comme un délit par la loi du 5/1/88 relative à la fraude informatique. La destruction volontaire d'une information dans un fichier ne se fait pas, bien sûr, automatiquement sur toutes les copies qui ont été faites du fichier et peut s'accompagner aussi de l'archivage de l'information avant destruction.

Fiable

Se dit d'une information qui est digne de confiance (2)(3). Cette notion exclut implicitement l'idée d'erreur. Il est intéressant cependant de noter qu'en informatique la fiabilité est définie comme "la sécurité de fonctionnement apportée par un appareil et mesurée par le MTBF (Mean

Time Between Failures) exprimant le temps moyen qui s'écoule entre 2 pannes (5) ou encore comme la probabilité de fonctionnement sans défaillance d'un dispositif logiciel ou matériel dans des conditions d'exploitation définies et dans une période de temps limitée" (4). Etre fiable correspond alors à un certain degré de fiabilité qui n'exclut pas l'erreur.

Mémorisée

Est mémorisé ce qui est fixé "dans sa mémoire" (2) ou au sens informatique ce qui "est stocké dans une mémoire quelconque" (4) (cf conservée).

La notion de "mémorisation" contient implicitement la notion de conservation mais contrairement à cette dernière, semble moins attachée à la notion de durée de conservation.

Mise-à-jour

Est mis-à-jour ce qui "a été modifié en fonction d'informations nouvelles" (4) (exemple du changement d'adresse).

Une mise-à-jour implique techniquement la destruction de la valeur précédente de l'information dans le fichier concerné. Contrairement à une croyance naïve elle n'est pas automatiquement effectuée sur les copies qui ont été faites du fichier ni sur tous les fichiers où se trouve cette information. Une mise à jour d'une information peut s'accompagner de l'archivage de la valeur précédente de cette information avant destruction.

L'article 36 de la loi stipule que " le titulaire du droit d'accès peut exiger que soient rectifiées, complétées, clarifiées, mises à jour ou effacées les informations le concernant qui sont inexactes, incomplètes, équivoques, périmées ou dont la collecte ou l'utilisation, la communication ou la conservation est interdite ".

Nécessaire :

Information dont "l'existence, la présence est indispensable pour obtenir le résultat désiré" (1).

La notion de "nécessité" est liée à la finalité du traitement qui utilise l'information (par exemple le nom, l'adresse et la fin de l'abonnement sont les seules informations nécessaires pour l'envoi d'une

publication périodique). Cette notion est à différencier de celle d'utile (cf utile).

Nominative :

"Qui contient énoncé expressément le nom" (1).

La loi du 6/1/78, quant-à-elle, définit les informations nominatives comme des "informations qui permettent, sous quelque forme que ce soit, directement ou non, l'identification des personnes physiques auxquelles elles s'appliquent, que le traitement soit effectué par une personne physique ou par une personne morale".

Objective :

"Qui fait référence à la réalité extérieure indépendante des consciences" (3) ; qui consiste "en une réalité ou un ensemble de faits constatables, dont les résultats sont indépendants de toute préférence subjective, et d'où sont exclus tout préjugé individuel, toute déformation plus ou moins volontaire de la réalité" (1).

Une information peut être objective, mais le choix de celle-ci pour un traitement donné peut être tout à fait contestable.

Pertinente, adéquate, non excessive :

Information qui "est exactement adaptée à l'objet dont il s'agit, qui se rapporte à la question, qui appartient au fond de la cause" (2).

La pertinence d'une information se fait par référence au traitement qui utilise cette information et non pour l'information elle-même (cf. objective).

Protégée :

"Mise à l'abri d'un danger, d'une agression, d'un risque quelconque" (2)(3) ; Information qui, "par des mécanismes logiciels ou matériels, ne peut être consultée ou modifiée que par des personnes autorisées" (4).

L'article 29 de la loi mentionne l'obligation d'assurer la sécurité des informations. La liste des protections matérielles et logicielles possibles est longue mais il n'existe pas de protection absolue. Le choix des méthodes de protection résulte de la conscience des décideurs, de l'importance de la protection des données nominatives et est le résultat

de compromis entre efficacité des méthodes, coût, facilité d'usage et prise en compte de l'importance de l'enjeu pour les utilisateurs.

Sensible :

"Qui peut être senti ou se fait particulièrement sentir (1) ; se dit de ce qui réagit à certaines actions et plus précisément d'un appareil, d'un instrument de mesure qui obéit à de très légères sollicitations" (3).

Par données sensibles on entend habituellement les opinions, les activités religieuses, philosophiques, politiques ou syndicales, l'état psychique, mental ou physique d'un individu, sa sphère intime, les origines raciales, les mesures d'aide sociale ou d'assistance, les poursuites et les jugements pénaux ; mais en fait ce n'est pas tant la nature de la donnée que l'utilisation qui en est faite et le moment de son utilisation qui en font une donnée sensible.

Ainsi l'adresse et le nom d'un réfugié politique recherché par les organes de répression de son pays d'origine sont des données sensibles alors que la religion de Jean-Paul II ou l'appartenance de Jacques Chirac au RPR et de Laurent Fabius au PS font partie intégrante de ces personnages publics appartenant à la sphère du pouvoir. L'approfondissement de la notion de données sensibles nécessite aussi la recherche, l'approfondissement de certaines définitions comme celle de la race dont l'existence est niée scientifiquement mais qui a une reconnaissance sociale.

La plupart des fichiers du secteur de la santé sont aussi particulièrement sensibles ; en effet la divulgation d'handicaps de certaines maladies peut perturber la vie familiale, scolaire, professionnelle d'une personne ; assureurs, employeurs... peuvent être intéressés par la présence de clients, de candidats à l'emploi dans certains fichiers (les registres du cancer...) par exemple.

L'article 31 de la loi de 78 indique qu'il "est interdit de mettre ou conserver en mémoire informatisée sauf accord exprès de l'intéressé, des données nominatives qui, directement ou indirectement, font apparaître les origines raciales ou les opinions politiques, philosophiques ou religieuses ou les appartenances syndicales des personnes".

L'article 6 de la convention européenne stipule, quant à lui, que "les données à caractère personnel révélant l'origine raciale, les opinions politiques, les convictions religieuses ou autres convictions, ainsi que les données à caractère personnel relatives à la santé ou à la vie sexuelle, ne

peuvent être traitées automatiquement à moins que le droit interne ne prévoie des garanties appropriées. Il en est de même des données à caractère personnel concernant des condamnations pénales".

Utile :

Se dit de ce qui "est propre à satisfaire un besoin estimé légitime, à procurer à une personne un avantage, un surcroît de bien être ou de bonheur mais aussi se dit, en économie politique, d'un bien ou d'un service qui répond à un besoin ou à un désir, quel qu'il soit" (3).

Cette notion est à différencier de celle de "nécessaire". En effet, par exemple, Hydro-Quebec, monopole d'Etat de distribution d'électricité au Canada veut constituer un super fichier. Il désire mémoriser, en plus des nom, adresse et numéro de téléphone (informations nécessaires) les noms des employeurs, les sources de revenus, les numéros d'assurance sociale et de permis de conduire, les noms, adresses et numéro de téléphone de 2 personnes connaissant l'abonné, les modes de chauffage et de climatisation, le type de chauffe-eau, É, informations non nécessaires mais "utiles" pour Hydro-Quebec¹.

Validée

"Déclarée valable" (1) ; information "rendue valable, autorisée par une opération spécifique comme la validation de la donnée tapée au clavier d'un distributeur automatique de billets de banque" (5). La validité d'un programme est, elle, "la qualité d'un programme dépourvu d'erreur et apte à résoudre un problème donné" (4).

La validation des données peut se faire au moment du recueil ou de la saisie des données. Elle s'effectue souvent indépendamment de la personne concernée et n'a quasiment plus lieu durant la durée de vie de l'information sauf lorsque l'intéressé exerce son droit de rectification (article 36), toute personne ayant "le droit de connaître et de contester les informations utilisées dans les traitements automatisés dont les résultats lui sont opposés" (article 3 et exercice du droit d'accès : chapitre V).

La validation du traitement, quant à elle, doit avoir lieu à trois niveaux : au niveau informatique, au niveau de la Commission Nationale Informatique et Libertés (CNIL) et au niveau des usagers. Au niveau informatique elle doit être le résultat d'une étroite collaboration entre commanditaires du traitement et informaticiens. Au niveau de la CNIL

la validation doit se faire lors de l'examen de la déclaration obligatoire dont doit faire l'objet chaque traitement (articles 15, 16 et 17 de la loi). Au niveau des usagers si, en théorie, toute personne a le droit aussi de connaître et de contester les raisonnements utilisés (article 3), en pratique aucune disposition de la loi ne permet concrètement d'exercer ce droit. Cette dernière validation devrait pouvoir se faire en liaison avec commanditaires du traitement et informaticiens, ces derniers fournissant les raisonnements utilisés dans les algorithmes ou le raisonnement suivi, pour chaque cas, dans un système expert.

Anne BRYGOO

BIBLIOGRAPHIE

- 1 - *Le Grand Robert*, dictionnaire alphabétique et analogique de la langue française. 1983.
- 2 - *Dictionnaire de la langue du 19^e et du 20^e siècle*, CNRS, Gallimard.
- 3 - *Grand dictionnaire de la langue française*, Larousse.
- 4 - *Dictionnaire de l'informatique*, sous la direction de Pierre Morvan, Larousse, 1981.
- 5 - *Dictionnaire d'informatique*, M. GINGUAY, A. LAURET, Masson, 4^e édition, 1990.

¹ Pierrot Peladeau, "Chroniques des sociétés surveillées, lettre d'une des Amériques" Bulletin de liaison du CREIS n°5 Juin 1988.