DONNÉES SENSIBLES ET RISQUE INFORMATIQUE
DE L'INTIMITÉ MENACÉE À L'IDENTITÉ VIRTUELLE

PAR

Danièle BOURCIER

Directeur de recherche au CNRS (CURAPP)

L'ordinateur traite *virtuellement* des individus, des territoires, des informations, des organisations ou des décisions mais agit *réellement* sur le monde[1]. L'informatique rend le monde *sensible*. Depuis le début des années *welfare* puis de la libéralisation généralisée des échanges[2], le management, qu'il soit public ou privé a incité massivement à la création de traitements informatisés. En retour, le pouvoir de mémorisation, de discrimination et de connexion des ordinateurs a suscité de nouveaux projets d'automatisation des activités humaines. Les sciences de l'artificiel pénètrent désormais toutes les sphères de la décision (médecine, économie, droit, gestion) : les effets induits sont incommensurables.

Un de ses effets les plus reconnus est de rendre plus *sensible* ce qui l'était déjà, et même de créer de nouvelles zones de vulnérabilité, particulièrement dans les domaines où les activités humaines paraissaient les plus complexes et donc les moins justiciables d'un traitement automatique.

1. A ce propos, je signale que cet article a dû être entièrement réécrit car il avait disparu lors de son transfert sur une disquette physiquement défectueuse. Ce cas n'est pas rare. Les modes de précaution ou de protection, les utilitaires de récupération aussi sophistiqués soient-ils ne suppriment pas ce risque redoutable...

2. La directive européenne du 24 octobre 1995 dont nous reparlerons par la suite indique d'ailleurs dans ses premiers considérants que les systèmes de traitement de données doivent contribuer au développement des échanges et que l'article 8 A du Traité de Maastricht sur le fonctionnement du marché intérieur va augmenter les échanges de données à caractère personnel entre administrations nationales et développer des flux transfrontaliers entre les entreprises privées des différents états membres.

CURAPP - *Questions sensibles*, PUF, 1998.

Le mot "sensible" n'a pas en soi de connotation positive ou négative: dire qu'un phénomène est sensible ne dit rien de plus que ses états sont instables et que les déséquilibres sont relativement imprévisibles et potentiellement incontrôlables. Cependant si l'on considère l'informatisation de la société, la notion de *donnée sensible* évoque plutôt des menaces supplémentaires. Certes, tous les acteurs d'un phénomène ne perçoivent pas les mêmes dangers et ne sont pas confrontés aux mêmes risques. Après avoir opposé l'informatique aux libertés, on tente de réconcilier protection des personnes et circulation des données. Est-ce vraiment un progrès? Il n'en reste pas moins que c'est autour de cette notion de sensibilité, des degrés de l'atteinte, des protections, des exceptions, et des limites qu'elle implique, que s'est construit le domaine "informatique et... libertés".

Notons d'emblée que malgré sa centralité dans le champ qui nous intéresse, la notion n'a jamais été définie. Est-ce une lacune, ou ne fait-elle pas partie de ces notions indéterminées mais "ouvertes", qui délimitent simplement les conséquences que son emploi *en contexte* déclenche ? L'analyse et l'interprétation des règles de droit peuvent être alors pertinentes pour cerner la question générale qui nous préoccupe à savoir : que dit-on *réellement* d'un phénomène — naturel, social, expérimental — quand on lui affecte la qualification de sensible et quelle décision éventuelle ce "dire" peut provoquer ? Dans le champ de l'informatique, repérer ce qui a été qualifié de sensible revient donc à analyser les dysfonctionnements du progrès technologique[3], à décrire les pouvoirs, contre-pouvoirs et garanties qui se sont développés autour de la "tentation de la puissance" impliquée par ce nouvel outil et, "en creux", ce qui doit irréductiblement, échapper à l'intelligibilité de la machine.

Nous ferons un bref panorama des tendances sensibles de l'informatique. Puis nous verrons comment le droit protège les données, traitements et utilisations sensibles de l'informatique. Nous donnerons ensuite quelques cas qui ont donné lieu à un approfondissement de la notion de sensible. Nous conclurons cette exploration en terre juridique par une réflexion sur la fonction pragmatique de la notion de sensible et sur les rapports du sensible avec la complexité des systèmes dans lesquels nous vivons.

I - DES DONNÉES SENSIBLES AUX PROFILS DE PERSONNES

On étendra la notion de données sensibles en informatique à trois types de phénomènes : les données *initiales*, liées à la vie privée que l'on collecte de plus en plus, de façon massive, et souvent à notre insu, les données *dérivées*, traitées pour créer de nouvelles applicationssensibles, enfin les *traitements* à

3. On parle aussi en sciences de la gestion du paradoxe de Solow par lequel l'informatique n'est pas synonyme de productivité, bien au contraire ; d'où les nombreux travaux sur l'amélioration des systèmes d'information par la recherches d'indicateurs "sensibles" de pilotage.

risque en tant qu'ils peuvent provoquer des effets inattendus voire nettement indésirables. Les deux premiers effets renvoient au rôle croissant de l'information sur les personnes et au risque particulier que représente l'informatique en tant qu'elle peut stocker et croiser ces données. Le troisième concerne non plus l'information mais les modèles de décision qui, élaborés à partir des données sur les personnes peuvent leur être opposables.

Nous nous focaliserons sur le "sensible" dans le traitement informatique des personnes, en ce qu'il concerne leur vie privée (erreur, intrusion ou diffusion) et leur identité (éclatement, reconstruction).

A) Qu'appelle-t-on "donnée sensible" ? Les frontières floues de l'intimité

Certaines données étaient sensibles avant l'informatique. Le droit s'en préoccupait dans un but de protection de la vie privée et des libertés fondamentales. Le droit civil en a affirmé le principe[4]. Le droit pénal a créé les infractions d'atteinte à la vie privée, au secret professionnel et au secret des correspondances[5]. Mais l'informatique a aggravé ces risques, et le développement de la société de l'information a étendu le type de données dont l'utilisation peut, aux frontières de l'intimité, constituer de véritables atteintes à la personnalité.

La loi Informatique et libertés[6] ne donne pas de définition de la notion de données sensibles : mais elle crée une catégorie particulière de données particulièrement protégées. L'article 31 en donne une liste qui s'est lentement stabilisée : ce sont les *"données nominatives qui, directement ou indirectement font apparaître les origines raciales ou les opinions publiques, philosophiques ou religieuses ou les appartenances syndicales ou les mœurs des personnes"*. N'y figurent ni les données patrimoniales ni les données professionnelles. La liste dont le cadre a été fixé depuis 1981 dans la convention du Conseil de l'Europe a été reprise dans la directive européenne[7] : ce sont les données à caractère personnel qui révèlent *"l'origine raciale ou ethnique, les opinions politiques, les convictions religieuses ou philosophiques, l'appartenance syndicale"*[8]. Les données sur la santé ne figuraient pas dans la liste des données sensibles visées par la loi de 1978. Celles concernant les moeurs ont été introduites en 1992.

Comment situer la sensibilité de ces données par rapport aux autres données nominatives protégées par la loi ? Le texte français, rappelons-le, fait une

4. Code civil article 9 : *"Chacun a droit au respect de sa vie privée"*.

5. Code pénal articles 226-1 et 226-13 à 226-15.

6. Loi 78-15 du 6 janvier 1978.

7. Directive 95/46/CE du 24 octobre 1995 relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données.

8. Directive du 24 octobre 1995 : le mot "sensible" apparaît dans le 34 ème considérant.

distinction entre celles *"qui ne portent manifestement pas atteinte à la vie privée et aux libertés"* et qui ne sont soumises qu'à une déclaration simplifiée de conformité *et* les données nominatives qui (par opposition implicite) peuvent porter atteinte à la vie privée. Ces dernières, suivant qu'il s'agit d'un traitement pour le compte de personnes publiques ou privées, doivent faire l'objet d'une déclaration *ou* d'une demande d'autorisation.

Puis, une troisième catégorie concernent des "données" nominatives (et non plus des informations), celles que nous avons énumérées dans l'article 31 comme données sensibles. Enfin on ajoutera une quatrième catégorie qui concerne les données du NIR (numéro de Sécurité sociale) et qui sont implicitement considérées comme sensibles. En effet contrairement à d'autres pays, ce numéro[9] en France n'est pas déterminé de façon aléatoire : il est porteur d'information sur l'état civil de la personne comme le sexe, sa date et surtout, le lieu de naissance. On peut alors opérer des traitements discriminants sur des catégories de populations et notamment sur l'origine, à partir des données sur la naissance à l'étranger.

On a là une catégorisation implicite des différents types de données concernées par la loi. Ce classement en catégories, induit à partir *des différences de régime applicable*, sera brièvement commenté dans la section suivante.

Dans le texte suédois de 1973, la notion de sensibilité est beaucoup plus explicite : le mécanisme est "gradué" et fixe le régime de sécurité et de responsabilité attaché à trois types de fichiers. En effet, la notion de sensibilité peut être influencée par plusieurs facteurs en dehors de la *nature* des données. Ainsi la quantité (les détails sur une personne) peut devenir un facteur de sensibilité. De même, un fichier peut être sensible globalement, indépendamment de la nature des données individuelles. Les classes ne sont d'ailleurs pas définitivement fixées: des données peuvent *devenir sensibles* suivant les circonstances. Ainsi le niveau 1 implique des personnes qui ont des connexions avec l'administrateur du fichier : client, membre, employé etc. Dans la classe 2, figurent les catégories de la convention de Strasbourg. D'autres données sont ajoutées: celles concernant une information sur un crédit ou un recouvrement de dettes, une aide financière ou l'octroi d'une allocation de protection sociale donc des données patrimoniales et économiques. Données auxquelles il faut ajouter les fichiers qui contiennent une *évaluation* de la personne, et les fichiers personnels qui comprennent une *grande partie de la population* d'une région. Dans la classe 3, ce sont tous les dossiers personnels en matière judiciaire ou policière qui sont concernés.

La loi de 1978 fait donc une première partition en fonction des opérateurs et des traitements possibles et une deuxième en fonction des informations traitées, qui recouvre en partie[10] la liste traditionnelle des données liées à l'intimi-

9. Numéro d'identification au répertoire national mis en place par l'INSEE.

10. Les données relatives à la vie privée n'ont pas toutes la même valeur et la même sensibilité.

té. Il semble cependant que cette liste pourrait être augmentée car la société de l'information combinée au *welfare* a développé de nouvelles applications et de nouvelles menaces. La mise en œuvre des procédures d'attribution de prestations mais aussi la gestion des informations relatives aux usagers peuvent traiter des indicateurs sur la situation financière et économique, la vie professionnelle et les rapports à l'éducation de familles particulièrement démunies et déjà fragilisées : c'est le cas d'ANIS-ASE, système informatique de gestion de l'aide sociale à l'enfance sur lequel nous reviendrons. En revanche, les données sur la santé traitées pour la recherche médicale ont fait objet d'une protection soutenue de la part du législateur depuis le début des années 1980 à l'occasion de la constitution de fichiers de registre sur le cancer.

Donc la sensibilité concerne d'abord *une partie des catégories traditionnelles de la vie privée* vis-à-vis desquelles la nature même de l'outil informatique augmente les risques d'atteinte. La généralisation de l'informatisation accroît les capacités de collecte et de stockage systématique, l'opacité des fichiers numérisés et la difficulté de contrôle par les citoyens (droit à la rectification et à l'oubli).

B) Les données dérivées : la sensibilité des procédures identificatoires

L'usage de l'ordinateur implique une autre capacité que celle de constituer seulement des fichiers de données sur "la vie privée" de chaque citoyen (article 1 de la loi de 1978). L'article 1 énonce en effet, à côté des atteintes à la vie privée, celles concernant "l'identité humaine". C'est l'informatique en tant que telle qui rend plus *sensible* l'identité humaine en créant des données virtuelles, parfois à l'insu des personnes, qui sont entièrement nouvelles par rapport à celles consenties au départ. Il en est ainsi des méta-connaissances (généralisation) ou de connaissances indirectes (par agrégation, connections ou induction). Ces données peuvent alors être particulièrement sensibles : elles ne sont soumises à aucun contrôle et toute manipulation abusive ou erronée peut provoquer des conséquences imprévisibles. Enfin, d'autres données physiques, physiologiques ou génétiques peuvent faire l'objet de techniques identificatoires plus ou moins liées à l'informatique. Il s'agit à mon avis d'un autre aspect sensible de l'informatisation du corps de la personne. A ce titre, la directive en parlant de *"traitement sur des catégories particulières de données"* rend compte de ces nouvelles pratiques identificatoires.

En effet, la vie quotidienne dans les sociétés industrielles est scandée par un maillage d'identification des personnes et un éclatement de leur identité en fichiers de plus en plus nombreux. Parallèlement, des technologies d'identification développent, à partir de cette individualisation extrême du sujet, des mécanismes particulièrement sophistiqués de "reconstruction" de sujets virtuels. Les grands opérateurs publics ou privés qu'il s'agisse d'administrations, des secteurs du management, des banques et des assurances sont de plus en

plus assistés par les industriels de l'identification. Il s'agit par exemple de *systèmes biométriques* qui reconnaissent un individu à partir de ses caractéristiques physiques, morphologiques et même maintenant génétiques. On sait par exemple comment les données contenues dans le génome intéressent ceux qui visent à exclure les personnes jugées porteuses d'un trop grand risque médical. On vient d'apprendre récemment[11] que des travaux de "médecine prédictive appliquée au travail" développés par l'Institut national de la recherche et de la sécurité pour la prévention des accidents du travail et des maladies professionnelles visent en toute bonne conscience à identifier, à partir d'une analyse des prédispositions génétiques des salariés, des personnes à risques accrus, et, partant, à élaborer des mesures préventives et une surveillance médicale à leur égard. Il est à craindre que ces suivis conduisent plus à des politiques de sélection que de prévention. On parle déjà d'identification discriminatoire et de "délit de sale gène".

Il existe un autre secteur délicat de l'informatique identificatoire, celui qui gère la reconnaissance des signatures (mots de passe, cryptogrammes, cartes à mémoire, téléphones mobiles...) et qui suit les traces multiples que nous laissons dans les multiples réseaux chargés de renforcer la sécurité et de traiter nos transactions. Les derniers nés appelés "agents intelligents" sont de petits robots fureteurs qui circulent dans l'ombre et en toute liberté sur le net et devisent entre eux pour consulter l'agenda électronique de leurs mandants, stocker des données ciblées ou connaître les caractéristiques de ceux qui consultent certains sites ou sont intéressés par certains produits. Ils utilisent massivement des *cookies* qui sont en réalité des enregistrements d'informations stockés chez le client, par le serveur, dans un fichier texte. Le problème est que seul le serveur auquel le client est abonné peut ensuite y avoir accès, et sans que le client s'en doute le moins du monde. Rappelons qu'actuellement les principaux navigateurs du marché reconnaissent les cookies[12].

Enfin, un autre secteur recouvre les technologies de construction d'individus statistiques, de consommateurs probables ou d'assurés à risque. Ces technologies fondées sur des modèles de *profiling*, de *scoring* et de *matching* reliés à des bases de données sociales ou géographiques, instaurent de nouvelles méthodes de connaissance de l'individu susceptibles de prédire ses comportements et ses choix. Le *profiling* est une technique de gestion de données fondée sur une phase de généralisation de cas multiples (un profil) et une phase d'application de ce profil à un cas individuel nouveau. L'établissement de *profils de groupe* est susceptible de détecter la fréquence avec laquelle un facteur sera répété à travers chaque élément d'un fichier de population par exemple. Dans un *profil individuel*, on peut à partir de données sur une personne voir jusqu'à quel point ces données convergent vers un modèle à risque[13]. On peut

11. Journal *Le Monde* du 7 janvier 1998.

12. A ce propos, le site Web de la CNIL a développé une action pédagogique en ligne sur la façon dont fonctionnaient les cookies (http://www.cnil.fr).

13. Compte tenu des procès en matière de responsabilité médicale aux Etats-Unis, il existe

identifier par exemple les consommateurs de drogue parmi les salariés d'une entreprise (type de retards, âge, etc..). Profils de vie, segments de clientèle fondés sur des systèmes de valeurs ont pour but de caractériser, de prédire des modes de consommation[14]. Le *scoring* affecte des poids ou de coefficients à chaque modalité des variables et agrège au final l'ensemble de notes[15]. Le *matching* permet de rapprocher et de comparer des données à partir de plusieurs fichiers pour en vérifier par exemple la cohérence et faire ressortir les anomalies éventuelles.

Dans ce processus, l'individu en tant que sujet ne participe plus de plein gré à la construction de son *identité sociale* à travers des mécanismes d'interaction explicites, consentis et acceptés. L'attribution de rôles lui échappe et une identité déformée lui est renvoyé par l'attribution de préférences et de choix qui lui sont corrélés "automatiquement". Sa propre autodétermination peut être affectée par ses processus. Mais un autre aspect est important à noter. Etant donné qu'il s'agit de données qui sont construites à partir de fichiers préexistants, le consentement ou l'opposition des intéressés ne sont plus inscrits dans les processus. Or, parallèlement aux libertés fondamentales inscrites dans l'article 1er de la loi de 1978, ont été instaurées des *"libertés de protection"*[16] : ces garanties que constituent le droit à l'oubli, le droit à la loyauté des données, le droit à la confidentialité, et le droit d'accès, deviennent partiellement inefficaces en face de tels traitements.

C) La décision artificielle comme traitement sensible

Les outils d'intelligence artificielle comme les systèmes interactifs d'aide à la décision, les systèmes experts, les réseaux neuronaux artificiels ont pour objectif de simuler des activités cognitives proprement humaines: diagnostic, analyse de stratégie, jugement. Ils utilisent souvent des données issues de comportements individuels et de techniques identificatoires vues précédemment — notamment les techniques de *profils* traduites en règles de sens commun — pour constituer leur base de connaissances. Ils utilisent des questionnaires très détaillés sur les individus à partir desquels "le raisonnement" de la machine simulera non seulement leur comportement mais l'appréciation du décideur et produira des effets à leur égard. Il sont destinés en général à améliorer la

(suite note 13) une liste noire informatisée des patients qui sont susceptibles d'intenter un procès...

14. Voir *Le Monde*, "Le fichage des consommateurs s'accroît et se sophistique", 7 février 1998 où l'on apprend que 2,5 millions de Français figurent dans la méga-base ConsoData et 3,5 millions de foyers ont été répertoriés par sa concurrente Claritas. Il en résulte que 20 % de la population française est listée en fonction d'un millier de critères (90 % aux Etats-Unis)...

15. La Caisse nationale de prévoyance utilise un système expert ANDROMED fondé sur la technique du *scoring* : le profil de santé du client est calculé à partir de ses propres données et de recoupements avec d'autres fichiers ainsi que d'un examen médical (qui fera de plus appel à des examens génétiques).

16. Terme proposé par G. Vedel.

productivité d'un service administratif ou à réduire les risques d'un choix, qu'il s'agisse d'un recrutement[17], de l'octroi d'une subvention ou d'un crédit, du refus d'une autorisation ; ils ont besoin de *données* particulièrement *sensibles*. Mais surtout, leurs effets *dans le monde réel* peuvent porter atteinte aux droits de la personne car les critères retenus peuvent être discriminatoires et les motifs prédéterminés.

Attardons-nous sur des outils nouveaux liés au développement massif de réseaux interconnectés et à la disponibilité de vastes fonds de données (mégabases) en ligne qui ont pour finalité ce que l'on a appelé le "*data mining*". Le *data mining* est "*le processus de découverte de corrélations, formes et tendances nouvelles et significatives en passant au crible de grandes quantités de données stockées dans des bases et utilisant des technologies de reconnaissance des formes conjointement aux techniques statistiques et mathématiques*" (Groupe Gartner)[18]. Cette technologie qui regroupe l'ensemble des méthodes d'exploitation des données particulièrement est présente dans le marketing mais on peut aussi la repérer dans d'autres domaines de la gestion. Ce mode d'exploitation est né avec l'idée qu'il faut passer des recherches sur les "segments de marché" à une relation avec des individus : pour tenir compte du client individuellement, il faut avoir le maximum de connaissances le concernant : observer ses besoins, se souvenir de ses préférences et apprendre de ses contacts passés[19]. Dans un secteur ultra-concurrentiel, où se trouvent la plupart des industries qui utilisent les plus grands volumes d'informations, il n'y a plus de limite pour exploiter le maximum de données individuelles obtenues à partir de ce qu'on appelle les enregistrements *transactionnels*. C'est pour cette raison par exemple, que les supermarchés sont devenus des courtiers en information. Des clients divulguent eux-mêmes leurs données personnelles pour élaborer des modèles prédictifs qui leur seront opposables... Cette recherche de généralisation prédictive s'appuie sur ce qu'on appelle désormais "la connaissance de communautés". Cependant, ces techniques peuvent s'appliquer aussi bien à la médecine qu'au droit ou à la commande de processus industriels. Les tâches de *data mining* sont la classification, l'estimation, la prédiction, l'analyse de similitudes et l'analyse taxinomique (*clusters*). Le cercle vertueux du *data mining* est de transformer les données en informations, les informations en décisions et les décisions... en bénéfices.

---

17. Ce fut le cas pour des logiciels de recrutement comme SIGMUND qui comprenait 450 questions personnelles (habitudes de vie, relations amicales...) et auto-administrés par le candidat au recrutement... Les temps de réponses étaient eux mêmes saisis et... interprétés.

18. Berry (M.-J.-A.) et Linoff (G.), *Data mining, techniques appliquées au marketing, à la vente et aux services clients*, Paris, Interéditions, 1997, p. 81.

19. La Bank of America s'est lancée dans de tels outils pour classer ses clients suivant leur probabilité de réponse à une offre de prêt hypothécaire. Constituée sur des enregistrements dont certains dataient de 1914, cette base de données comprenaient à peu près 250 attributs par clients. L'outil a appris les différences entre ceux qui avaient eu un prêt et les autres. On a pu ainsi trouver les règles de bon prospect. A cela s'est ajouté le modèle : "personnes avec enfants en âge de fréquenter l'université" et "personnes à revenus élevés mais variables"».

Enfin, l'ordinateur peut devenir aussi un outil sensible en tant qu'il peut provoquer des effets non attendus quand il se substitue au décideur dans des fonctions de responsabilité institutionnelle. L'exemple fameux en est la riposte nucléaire automatique. Comment laisser à un système de déclenchement automatique la responsabilité d'une telle action avec de si tragiques effets ? Mais il est des situations moins stratégiques qui peuvent être aussi sensibles : certains systèmes experts en matière sociale ont pour objectif par exemple de pré-instruire le dossier d'un demandeur et de calculer le montant d'un droit. On a pu observer que leur règles de production, leur mode de raisonnement et leur utilisation en général étaient en contradiction avec un certain nombre de règles et de principes juridiques comme l'appréciation discrétionnaire de certaines données en matière de RMI[20] ou d'allocation parentale[21].

La décision automatique est un sujet sensible : en effet les logiciels eux mêmes peuvent produire des décisions illégales ou injustes sans que la cause en soit immédiatement repérable. La complexité de leur architecture rend leur contrôle difficile. L'ensemble de la combinatoire des cas possibles ne peut être entièrement explorée. Il peut aussi y avoir un erreur dans les données, dans les règles ou dans la logique mettant en jeu la fiabilité et la sécurité des données et provoquant des inférences erronées sur les personnes. Des intrusions volontaires peuvent rendre les données sensibles rapidement accessibles et en détourner les finalités premières (fraude informatique).

II - LE DROIT PROTECTEUR DES UTILISATIONS SENSIBLES DE L'INFORMATIQUE

Le droit n'est peut être pas le meilleur moyen pour se protéger des effets indésirables de l'informatique et particulièrement des accès frauduleux ou indélicats. Les modes de sécurisation des systèmes d'information — et notamment les méthodes de cryptographie — sont plus efficaces que l'encadrement par le droit. Cependant le droit reste le dernier recours des citoyens contre les traitements illégaux de données personnelles. Les questions sensibles visées par le droit concernent d'abord les données puis les traitements : quelles sont les règles, les définitions, les statuts, les régimes qui ont été mis en place pour pallier les atteintes graves aux droits fondamentaux c'est à dire l'ingénierie juridique produite pour tenir compte de ces effets.

---

20. Bourcier (B.), "Les lois sont-elles des logiciels ? L'aide à la décision en matière de législation sociale", *in Revue des affaires sociales*, Paris, Masson, mars 1995.

21. Voir notamment Magnusson (C.), "Introducing knowledge-based systems in the Swedish social insurance organization" *in Information Age*, London, Butterworth & Co, 1988 qui examine le système expert suédois ALTO : *"The present use of computer technology does not on the whole comply with administrative principles and rules. The implementation of knowledge-based systems may aggravate the situation... Who was to decide which legal references (precedents, legislation etc.) were to be included ? The question appeared when a vague rule in the National Insurance Act has to be supplemented... There is a need to investigate problems of legal, technical and political nature"*.

Comment le droit intervient-il vis-à-vis de ce que l'on vient de caractériser de sensible, qu'il s'agisse de données, de méta-données et de décisions ? Va-t-il "protéger" les individus, ou gérer seulement "l'équilibre des intérêts en cause" ? Nous allons voir dans la loi française — que le doyen Vedel définissait comme *"un véritable code de l'informatique et des libertés"* — puis dans la directive européenne de quelle façon l'ingénierie juridique tente de *désensibiliser* le phénomène des traitements informatiques.

A) Les données sensibles : entre l'intérêt du fiché et l'intérêt public

Les données sensibles dont le champ est plus étroit que la vie privée mettent en jeu la liberté de conscience, l'interdiction de toute discrimination et la liberté d'opinion. Les règles applicables (article 31 de la loi) peuvent s'énoncer en plusieurs étapes, des principes aux exceptions :

- Un principe général d'interdiction : il est interdit de *"mettre ou conserver en mémoire informatique certaines données nominatives qui directement ou indirectement ferait apparaître la race les opinions ou les appartenances syndicales, et les mœurs"*... L'interdiction concerne aussi les infractions condamnations et mesures de sûreté car seules les juridictions et autorités peuvent le faire dans un cadre légal.

- Deux moyens pour "désensibiliser" les données : le principe du consentement de l'intéressé et le principe de l'intérêt supérieur (motif d'intérêt public). Dans ce dernier cas, il faut cependant un décret. A titre d'exemple, pour informatiser le recensement des présentations de candidature à l'Elysée, le Conseil constitutionnel dut demander un tel décret.

- Une exception : les églises, partis et organismes d'opinion peuvent avoir un registre informatisé. Du fait de la liberté d'association, ces organismes *ne peuvent pas être contrôlés* par la CNIL. Le fondement de ces dernières dispositions est lié à la liberté d'opinion et d'association, au risque volontaire pris par l'affilié et à la pesée consciente et présumée de ses différents intérêts. Les données peuvent donc devenir "insensibles". Mais elles peuvent perdre aussi le privilège de la protection renforcée au nom de plus grandes valeurs en balance: intérêt général, santé, intérêt scientifique. Regardons le statut des données concernant la santé. Quand on observe les avis de la CNIL, il s'agit essentiellement de sauvegarder l'intimité du malade par rapport aux investigations considérées comme intempestives de la recherche médicale. La solution a été donnée en ces termes : les données sensibles deviennent... "insensibles" si le consentement exprès de l'intéressé est donné. Cependant ce consentement, qui levait l'interdit de la communication, obligeait aussi à lever l'interdit de l'information du patient sur la maladie dont il pouvait être atteint. Une dérogation peut alors être décidée par la CNIL...

Le traitement informatique de données génétiques à des fins épidémiologiques met en lumière un autre paradoxe. La science peut devenir prédictive si elle traite statistiquement de grands fichiers de données. Mais le résultat de ce savoir peut-il être diffusé sans risque ? Le comité d'Ethique a limité la communication de ces informations prédictives qui ne peuvent désormais être transmises que par le malade. La CNIL a donc essentiellement visé la recherche médicale en proposant en 1994 des articles supplémentaires (Article 40-1 à 40-10) mettant en œuvre deux niveaux de contrôle et le consentement du malade)[22].

La directive (article 8) énonce des règles un peu différentes suivant la logique suivante :

- le principe d'interdiction est réaffirmé, étendant en outre les données sensibles à celle concernant la santé et la vie sexuelle ;

- une première catégorie d'exceptions correspond à l'intérêt ou à un acte de volonté de la personne :

* si l'intéressé donne son consentement explicite (et non plus seulement exprès) ;

* si les intérêts vitaux ("la clause humanitaire") de la personne concernée sont en jeu (si notamment la personne ne peut plus donner son consentement) ;

* si le traitement est effectué par une association à finalité politique, philosophique, religieuse, ou syndicale. Cette disposition paraît plus protectrice que la loi française sur ce point car elle envisage un certain contrôle de son application ;

* si les fichiers nécessaires pour respecter le droit du travail (dans certaines législations, l'employeur doit gérer des "données sensibles", mais non en France) ;

* lorsque des données sont manifestement publiques (déclarations d'hommes politiques par exemple) ;

* traitements nécessaires à l'exercice d'un droit en justice (dossiers d'avocats sur leurs clients par exemple).

- une deuxième catégorie d'exceptions vise des intérêts supérieurs à ceux de la personne :

* les traitements à des fins médicales qu'il s'agisse de diagnostics ou de gestion, *à condition que* le traitement soit effectué par un praticien soumis au secret professionnel ;

* les motifs d'intérêt public important comme la gestion de la santé et de la protection sociale, la recherche et les statistiques publiques ;

* le registre des infractions et condamnations pénales.

22. Le malade a le droit de s'opposer au traitement mais un consentement exprès et éclairé n'est exigé préalablement que dans le cas de prélèvements biologiques identifiants.

En ce qui concerne les autres données particulièrement sensibles, la loi de 1978 a protégé l'utilisation du répertoire d'identification des personnes physiques (NIR). Elle exige dans son article 18 l'avis de la CNIL préalablement à un décret en Conseil d'Etat. Les réactions à sa demande de diffusion en dehors de la sécurité sociale du NIR sont toujours très vives. La CNIL a rendu une trentaine d'avis de refus depuis vingt ans. Saisie pour avis en 1996 du projet de loi de financement de la sécurité sociale prévoyant des échanges d'informations à partir du NIR entre administration fiscale et organismes sociaux, elle a donné un avis négatif considérant que ces traitements excédaient le domaine du financement direct de la sécurité sociale. Les dispositions ont été ensuite retirées du projet (Délibération 96-075 du 1 er octobre 1996). Cette crainte vise, au-delà de la collecte de données propres à chaque individu la capacité de traitements transversaux et globaux que de tels fichiers autorisent et le détournement de la finalité initiale. Le spectre de l'histoire — et du gouvernement de Vichy — n'est pas absent de cette peur[23].

Dans la directive, les conditions de traitement de l'identifiant sont laissées à la discrétion des Etats membres : ce qui est un avantage quand le niveau de protection est élevé comme en France. Le texte européen étend la notion d'identifiant numérique à tout autre identifiant : le profil *génétique* est implicitement visé. Notons que pour l'instant les données génétiques ne sont pas encore considérées comme sensibles.

Une différence qui paraît devoir être notée entre la directive (article 8-3°) et la loi de 1978 concerne l'importance de la dérogation accordée aux traitements des données sociales. En effet, pour comprendre la portée les traitements considérés comme dérogatoires pour un "motif d'intérêt public" (ils ne sont plus interdits), il est utile d'aller se référer aux considérants européens. Or le 34 ème considérant précise qu'il *"s'agit d'assurer la qualité et la rentabilité* (je souligne) *en ce qui concerne les procédures utilisées pour régler les commandes de prestations et de services dans le régime d'assurance maladie"*. Cette précision — il est vrai, mal rédigée — revient à légitimer au nom de l'efficacité de la gestion des prestations une atteinte légitime à des données sensibles. Les garanties de confidentialité restent à définir en exigeant notamment qu'elles soient traitées impérativement par des personnes soumises au secret professionnel. De même, le statut de données sensibles risque d'échapper à des données de santé ou des données médico-sociales dans la mesure où la gestion de l'assurance maladie ou de la protection sociale en général devient un motif d'intérêt public.

La loi de 1978 avait mis au cœur de ses mécanismes de protection la finalité des traitements. Chaque renseignement doit en effet être situé dans un processus d'information vu dans un ensemble. Avec les possibilités infinies des

23. Le Parlement fédéral allemand a pris une résolution à l'unanimité déclarant inconstitutionnelle toute tentative d'introduire un tel identifiant.

méga-bases de données, le développement des profils et plus généralement des décisions automatiques, les menaces contre la vie privée proviennent moins des données sensibles par elles-mêmes que du rapprochement de données qui ne le sont pas *a priori* (localisation géographique, par exemple). En revanche, certaines données par les inférences qu'elles autorisent — comme la langue maternelle, dont on demande la qualité dans certains questionnaires visant à l'obtention d'une carte scolaire — peuvent devenir sensibles. Autant que pourra l'être la détermination de certains signes particuliers comme la "couleur de la peau" traités dans le fichier GEVI (Gestion de la violence) des renseignements généraux[24].

Cette menace est renforcée par l'utilisation de certaines catégories de logiciels qui opèrent des raisonnements sur des données en ligne et en continu à la place du décideur. En effet, ils ont pour fonction de rechercher sur l'ensemble du réseau des noms ou des groupes de noms ou des données et de les agréger en fonction d'un certain objectif. Ils simulent en direct des corrélations et peuvent faire certaines déductions qui s'apparentent à des choix multicritères. Considérés comme "technologies nouvelles " au sens de la directive par la CNIL, ces logiciels — derniers avatars du data mining — nécessiteront une veille technologique attentive.

B) Les décisions automatisées : l'exercice d'un droit d'accès et d'opposition est-il possible ?

La loi française comme la directive a été particulièrement soucieuse de parer les risques nouveaux provoqués par les technologies futures[25] : l'aide à la décision (dont la banalisation était encore difficilement prévisible en 1978) en constitue le meilleur exemple.

En effet, la loi de 1978 a conçu les deux articles phares qui ont permis par la suite de parer les effets les plus nocifs des systèmes experts, réseaux de neurones et algorithmes génétiques fondés sur des données personnelles. Rappelons que l'intelligence artificielle connaissait encore peu d'applications à cette époque. Que dit la loi française ? Elle interdit qu'une décision privée ou administrative qui implique un jugement sur des comportements humains puisse être fondée uniquement sur un profil de l'intéressé. Ce qui revient à dire que les systèmes informatiques ne peuvent produire des décisions automatiques opposables. Un agent humain doit rompre la chaîne décisionnelle pour l'évaluer voire la corriger, si besoin est, en fonction d'éléments subjectifs et circonstanciels.

---

24. Délibération n°96-098 du 19 novembre 1996.

25. La notion de "technologie nouvelle" (Directive) ou de nouvelles technologies (délibération de la CNIL du 14 mai 1996) est un concept qu'il paraît nécessaire de conserver car il permet de les soumettre à un examen particulier et en particulier de les considérer comme "à risques", nécessitant par là même une autorisation préalable.

Dans le domaine juridictionnel, où les données individuelles enregistrées sont particulièrement sensibles, parce que le juge doit juger un cas "individué" et, faudrait-il ajouter, parce que de toute façon l'on ne dispose pas d'un "modèle complet du monde", l'aide à la décision est définitivement interdite. Il est intéressant de noter que le risque paraît si important, et le domaine si sensible, que même une simple aide informatique est *a priori* suspecte.

En ce qui concerne la directive, les même dispositions sont reprises dans l'article 15 intitulé "décisions individuelles automatisées" : toute personne a ainsi *"le droit de ne pas être soumise à une décision produisant des effets juridiques à son égard ou l'affectant de manière significative, prise sur le seul fondement d'un traitement automatisé de données destinée à évaluer certains aspects de sa personnalité tels que son rendement professionnel, son crédit, sa fiabilité, son comportement etc."*.

Quelles sont les différences que les deux textes manifestent par rapport à ces décisions très sensibles ?

D'abord, la disposition est présentée au cœur du texte de la directive, dans la section "droit d'opposition de la personne" et non dans les articles de principe dont la place peut être emblématique. Ensuite, la directive définit la décision par ses effets, juridiques ou non, qui peuvent être plus ou moins "sensibles" : mais que signifie "être affecté de manière significative" ? La notion semble faire une certaine part aux circonstances, au type imprévisible de technologie utilisée, voire même à la subjectivité de la personne. Les exemples donnés renvoient plus directement à des décisions privées (demande de crédit, évaluation professionnelle). Mais la principale différence réside dans la présence d'exceptions : la personne concernée ne peut arguer de l'illégalité de cette décision si elle est la conséquence d'un acte volontaire (contrat) ou d'une loi et si des garanties ont été prises. Par exemple, les principes procéduraux de non-discrimination, de motivation et du contradictoire doivent avoir été respectés comme pour n'importe quel acte juridique.

Le principal danger de ces logiciels réside dans le fait qu'ils sont relativement opaques parce qu'organisés sous forme de bases de connaissances — plutôt que de données — et que les mécanismes d'inférence utilisés sont distincts des données. La loi de 1978 avait donc accompagné cette disposition du droit de connaître et de contester les résultats. Mais que signifie connaître ? S'agit-il d'avoir accès à une partie du raisonnement et des données du cas traité ou s'agit-il plus largement de pouvoir porter un jugement sur l'ensemble du raisonnement ? La directive dans ses considérants a prévu la même disposition puisqu'elle donne le droit à toute personne de connaître la *logique* de la décision qui lui est opposée. Mais à la différence de la loi de 1978, elle s'est prémunie contre une interprétation trop large qui donnerait l'accès à l'expertise globale du système: le secret des affaires et le droit de l'auteur du logiciel seront protégés en priorité. L'application de ce droit d'opposition paraît techniquement bien problématique à faire valoir.

Enfin, signalons que la CNIL, dans une délibération du 14 mai 1996, a élaboré une liste de "traitements à risques particuliers" nécessitant le maintien d'un contrôle *a priori* dans l'optique de la transposition de la directive. Cette liste comprenait, entre autres, les cas ci-dessus analysés : les données sensibles et les décisions automatisées. Y figuraient aussi les traitements conduisant à "une exclusion des personnes d'un droit, d'une prestation ou d'un contrat" qui ne sont pas loin de renvoyer directement aux traitements des décisions... automatisées.

III - LE CAS PARTICULIÈREMENT SENSIBLE DES DONNÉES DE SANTÉ ET DES DONNÉES SOCIALES

Il paraît utile en l'état actuel de la discussion sur l'informatique sensible d'insister brièvement sur des domaines qui ont nécessité dernièrement l'intervention de la CNIL. Il s'agit de la protection sociale et de l'aide sociale, domaines qui doivent être distingués de celui de la santé. En effet, ce dernier paraît avoir été particulièrement protégé (trop disent les chercheurs) après avoir suscité de grands débats de société au début des années 90. L'informatisation massive du secteur social, plus feutrée, plus complexe et sans doute plus technique suscite à l'évidence actuellement c'est-à-dire plus de vingt ans après l'affaire qui a été un des premiers dossiers de la CNIL, beaucoup moins de vigilance. Pourtant les questions soulevées sont particulièrement sensibles.

Exposons deux cas, à titre d'exemples.

Le premier est connu puisqu'il est à l'origine de la jurisprudence de la Commission nationale informatique et libertés. En 1973, un système sophistiqué de protection maternelle et infantile devait permettre de sélectionner *automatiquement* les enfants à surveiller médicalement. Une assistante sociale était alors dépêchée dès le huitième jour puis revenait régulièrement pour prendre en charge un enfant signalé comme prioritaire c'est-à-dire "à risques". Cette sélection était faite à partir de 170 critères qui constituait la base du programme : ce marquage devait constituer le fondement du dossier de l'enfant à l'aube de sa vie et le suivre durant jusqu'à sa vie adulte. Pourquoi ce système a-t-il été considéré comme particulièrement sensible ?

Dans la délibération de la CNIL[26], plusieurs motifs ont été retenus. D'abord, il ne fait aucun doute que cette application entre dans l'application de l'article 2 : il s'agissait d'un profil automatisé. Cependant, étant donné que ce profil n'est pas le seul élément de décision, il aurait pu être autorisé. D'autres arguments — pertinents pour cerner notre questionnement de départ sur le sensible — sont relevés de façon rigoureuse par la CNIL, dont la combinaison rend le traitement définitivement suspect :

26. Délibération du 16 juin 1981.

- l'incertitude du modèle choisi : les présomptions mêmes concordantes contiennent des "facteurs d'incertitude" ;
- l'impossibilité de corriger le modèle ultérieurement, par le contrôle individuel "cas par cas" ;
- le déterminisme de la machine qui exclurait définitivement certains enfants sans possibilité de les sélectionner par la suite ;
- la valeur contraignante de la décision, qui ouvre droit à des allocations ;
- la nécessité de prendre en compte des variantes notamment départementales ;
- l'hétérogénéité du système : certains faits ont objectifs d'autres appréciatifs, certains faits sont prédéfinis, d'autres sont combinés à partir de plusieurs données ;
- enfin, la nature diversifiée des données: administratives, sociales, socio-professionnelles, et surtout médicales.

La CNIL ajoute que cette présélection automatique "contestable" a déjà sensibilisé le milieu puisque diverses associations ont exprimé leur inquiétude. Elle donnera un avis défavorable aux profils personnels mais favorable aux applications statistiques anonymes.

Ce système mettait en place un contrôle généralisé de populations avec un risque d'exclusion sociale irrévocable. On retrouvera certaines de ces craintes dans l'affaire suivante.

En 1997, la déclaration d'un système de gestion informatisée de l'aide sociale à l'enfance (ANIS-ASE) mis en place dans le département de l'Ain est transmise à la CNIL. En réalité, cette affaire s'est déroulée sur plusieurs années car la mise en place a connu des retards considérables (qui ont troublé les orientations de départ ?) et a nécessité plusieurs avis de la CNIL. Saisie en 1994 d'une demande de conseil, celle-ci indiquait déjà que *"ce projet comporte un risque d'atteinte à la confidentialité des informations les plus sensibles et permet d'éventuels détournements de l'utilisation des données par les élus locaux"*. Elle recommandait plusieurs niveaux d'accès sélectif aux données et le troisième niveau devait *"couvrir les données sensibles, médicales ou sociales, qui ne sont accessibles qu'aux professionnels habilités lesquels se chargent également de la saisie"*. Elle exprimait aussi sa crainte de voir se développer *"un fichier global des populations défavorisées"* aboutissant à une *"cartographie de l'exclusion"* reposant sur la définition de profils individuels ou familiaux de précarité. En mai 1995, une délibération de la CNIL autorisait à titre expérimental et pour un an la mise en œuvre d'ANIS. Elle notait positivement entre autres dans son rapport d'activité de 1995 sur les données sensibles du niveau 3 : *"Ces informations qui permettent un suivi personnalisé (par exemple observation du travailleur social, conclusions médicales du médecin de PMI) sont soumises au secret médical ou social"*.

Revenons en 1997. Une délibération prolonge l'expérimentation et une demande d'avis définitif est soumise à la CNIL. La principale question soulevée par ce dossier concerne un point qui n'avait pas fait l'objet de contestations *a priori* (on était dans une phase de test) : la présence de typologies sensibles et leurs conséquences au regard de la jurisprudence de la CNIL en matière de profils fondés sur des données personnelles. Deux cents items — même s'ils sont facultatifs — portent sur l'éducation, la santé, l'état financier, le logement, la capacité personnelle, l'autonomie, le relationnel, les besoins de l'enfant, l'engagement de la personne etc. En analysant les items des différentes listes qui devaient servir de support à ce questionnaire, on s'aperçoit que les sous-catégories utilisées sont particulièrement subjectives ("difficulté rôle éducatif et parental"...). Elles sont en outre discriminatoires et risquent par induction d'affecter des données sensibles comme l'appartenance à une religion ("refus de fréquenter la cantine") ou des données de santé ("difficulté psychologique"). Une grille de codes portant sur le comportement et les mœurs des parents (leurs difficultés, leurs potentialités) impliquent des catégorisations subjectives comme "capacité à établir des liens affectifs", "difficulté psychologique" qui, traitées automatiquement dans la mise en place de politiques sociales, peuvent être particulièrement discriminatoires. L'affaire est en cours de réexamen, notamment en ce qui concerne ses possibilités d'extension à d'autres départements.

D'une façon générale, on assiste dans le secteur du travail social à l'apparition de logiciels d'un type nouveau. Ceux-ci deviennent un allié technique incontournable qui permet de libérer les gestionnaires de contraintes de management (parc de logements, etc.) Mais la propension à collecter des données et donc à les traiter (pour qu'en retour, les décideurs politiques départementaux puissent définir des politiques sociales plus adaptées par exemple), conduit à concevoir des interfaces de plus en plus détaillés sur les individus. Ces interfaces forcent à la *représentation informatique* de l'usager dans des nomenclatures particulièrement rigides[27]. Cette "sur-représentation" peut induire la façon dont la personne va être traitée. En clair, comme le signale un des représentants des travailleurs sociaux qui tentent de réagir avant que le système ne soit définitivement implémenté : "*l'informatique réclame en cette période de déréglementation économique et sociale une vigilance de tous les instants. Elle ne peut se réduire à un outil liberticide qui viserait à contenir les budgets sociaux...*".

L'informatisation du travail social est ainsi en train de développer des phénomènes de "taylorisation virtuelle" qui affectent non seulement les ayants droit ou les citoyens mais aussi les métiers du social. De plus en plus, au fil des expériences départementales (Ain, Haute-Garonne, etc.), des logiciels apparaissent, qui tendent à une modélisation des personnes au détriment d'une

27. Le travail social à l'épreuve des empires virtuels, *in Le journal de l'Action sociale*, mai 1997.

approche globale et intégrée et de la prise en considération de la dimension unique d'une situation. Prises souvent dans un but de meilleure gestion, ces logiciels finissent par traiter de données sensibles et produire des dérives dans les usages et les finalités. Danger renforcé par l'absence de transparence et d'information. En outre l'informatique, définitivement binaire, est malhabile à traiter de catégories floues et indécises : *"Ce qui fait l'originalité irremplaçable de la pensée humaine, c'est le pouvoir de non-pertinence, l'imprévisibilité vraie de l'énonciation, et non plus l'imprévisibilité domptée à l'intérieur d'un système probabiliste. A la proposition binaire «de deux choses l'une» c'est le pouvoir de répondre la troisième qui importe"*[28].

Revenons à la loi française et à la directive. Les données de santé sont considérées comme des données sensibles mais non les données psychosociales. Pourtant celles-ci concernent des informations délicates et susceptibles "d'affecter les personnes de manière significative"... comme le niveau de revenu mais aussi l'octroi d'une allocation sociale ou le refus d'un prêt. Or de plus en plus souvent ces données deviennent sensibles surtout si elles sont agrégées à d'autres données financières ou professionnelles. D'un autre côté, les données sociales intègrent des données qui deviennent potentiellement des données de santé ("incapacité à accomplir acte de vie", "tentative de suicide"). Compte tenu du fait que ces données sont saisies lors d'entretiens où peuvent aussi être intégrées des données administratives dans la même base commune, on assiste à une banalisation de l'échange et du suivi personnalisé qui ne devrait intervenir qu'avec des professionnels soumis à des obligations de secret et de confidentialité.

On remarquera que l'application des ordonnances du 24 avril 1996 — et notamment celle sur la maîtrise médicalisée des dépenses de soins — nécessite de plus en plus de traitements automatisés d'informations nominatives[29]. Les systèmes de gestion de la santé et de la protection sociale auront besoin de dispositifs de plus en plus fiables d'identification des personnes, une généralisation de *"l'identité informatisée"* dont parlait Philippe Lemoine en 1980[30]. Il est donc de plus en plus nécessaire de donner des garanties (droit à l'oubli, droit à la confidentialité...) aux données susceptibles d'entrer dans les dossiers des assurés sociaux. Ces cas décrits suggèrent en outre qu'il est urgent face à ces nouveaux outils de modélisation d'étendre la liste de données sensibles au delà des listes traditionnelles concernant la vie privée et d'y inclure notamment les nombreux traitements qui apparaissent dans la gestion de la protection sociale non pour les interdire définitivement mais pour les entourer du maximum de garanties.

---

28. Escarpit (R.), *Théorie générale de l'information*, Paris, Hachette, 1976.

29. Procédures de télétransmission de feuilles de soins électroniques, carte de professionnel de santé, informatisation des cabinets médicaux, carte électronique individuelle (carte SESAM VITALE), répertoire national interrégimes des bénéficiaires de l'assurance maladie.

30. Lemoine (Ph.), "L'identité informatisée" *in Les enjeux culturels de l'informatisation*, Paris, La Documentation française, 1980.

CONCLUSION : DOMAINES SENSIBLES, SYSTEMES COMPLEXES ET RESPONSABILITÉ

La notion de sensible fait partie d'une catégorie particulière de marqueurs linguistiques qu'on peut appeler "concepts fonctionnels" : non seulement ils ne sont jamais définis mais leur fonction est d'ouvrir des espaces de sens non finis. Ces espaces de sens se recomposent de façon dynamique au fur et à mesure des emplois dans le contexte. Leur fonction est de "déclencher" des raisonnements implicites vers des conclusions de type décisionnel. Autrement dit, utiliser en contexte le marqueur "sensible" ne qualifie aucun élément du monde mais renvoie à la fonction illocutoire du discours. Le locuteur va alors apprécier les effets de l'emploi en contexte. Ces termes s'apparentent aux mots indéterminés (mais non flous). Maintenant il convient de dégager ce que "déclenche" le mot "donnée sensible".

Les effets sont de plusieurs types. En droit, l'emploi de ces mots déclenche des régimes, des statuts, des sanctions. Une donnée sensible est une donnée qui va être particulièrement protégée. Mais elle n'est pas sensible en soi : elle peut être désensibilisée. Celui qui traite des données sensibles voit sa responsabilité renforcée. Et là on rejoint la notion de décision complexe et de systèmes dynamiques que l'informatique tente d'observer et de simuler.

Car le sensible c'est aussi l'instabilité d'une situation du monde ou des effets de l'utilisation d'un objet, d'une idée, d'un outil. Un logiciel peut être sensible en tant qu'il peut "déclencher" aussi des effets non recherchés : en ce sens les missiles à frappe automatique sont des objets sensibles. Tout automatisme peut créer des situations imprévisibles qui accroissent les effets indésirables. On sait qu'un système dont la dynamique est non-linéaire peut donner naissance dans certaines conditions à une évolution apparemment désordonnée. Dans ces dynamiques linéaires, à un point de vue unique s'oppose une multiplicité de points de vue légitimes sur le monde. De plus en plus les gestionnaires veulent relever le défi de décider en situation complexe : ils créent donc des nouveaux outils qui créent des zones de certitude dans lesquelles ils peuvent avoir l'illusion de la gestion. Les systèmes non linéaires fondés sur la notion de "sensibilités aux conditions initiales", montre que les systèmes sont rarement stables. Cela signifie que même dans un espace relativement déterministe des variations infimes peuvent induire des effets très divergents.

La notion de *sensible* renvoie donc aux modèles de la complexité. En quoi décider en environnement complexe doit tenir compte de la notion de sensibilité. Nous croyons penser et agir en termes linéaires : la dynamique des systèmes humains relève en réalité du hasard et du chaos.

La complexité enfin renvoie aux travaux sur le principe de précaution et l'éthique du savoir. Suscitées par les technologies de l'intelligence, ces recherches ont suscité des développements ultérieurs dans le domaine de la responsabilité des décideurs. On ne peut en effet actuellement définir des systèmes d'information sans prendre en compte la notion de responsabilité car de plus en plus la capacité de s'informer et l'accessibilité des données renversent la charge de la preuve. Celui qui exerce un pouvoir ne peut plus affirmer ne pas savoir les effets possibles de ses actes. Ce qui signifie, à l'heure du management du risque et de l'informatisation en réseau, que les systèmes d'information doivent prévoir le rôle de chacun et de tous dans la transmission des connaissances et le traitement global de la décision.