

# L'acte de juger est-il modélisable ? De la logique à la justice

Danièle BOURCIER

*Directrice de recherche au CNRS (CERSA/Université de Paris II)  
Comité d'éthique des Sciences*

Résumé. – Faut-il dire le droit ou rendre la justice ? Cette question a parcouru toute l'histoire de la philosophie du droit mais l'e-justice est en train de renouveler ce débat. Nous verrons d'abord si les logiciels de décision, à base de règles ou à base de cas, sont capables *techniquement* de simuler la décision juridique. Puis nous examinerons si ces logiciels ont la possibilité *juridique et éthique* de le faire. La conclusion la plus raisonnable est de réaffirmer la spécificité humaine de l'activité de décision pour penser les limites de la machine à juger.

Mots-clés : décision - intelligence artificielle - machine à juger – connexionnisme - système expert juridique – pouvoir discrétionnaire – langage juridique

La philosophie du droit fait appel à des apories, des dilemmes, des oppositions qui révèlent la difficulté fondamentale de l'acte de juger. On peut citer certains de ces couples pris dans divers domaines du droit : jugement analytique/jugement synthétique<sup>1</sup>, règles/principes<sup>2</sup>, pouvoir lié/pouvoir discrétionnaire<sup>3</sup>, égalité devant la loi/individualisation de la peine<sup>4</sup>, *easy cases/hard cases*<sup>5</sup>, logique/argumentation<sup>6</sup>. Ils ont en commun de représenter le conflit qui préside à tout jugement humain : faut-il dire le droit ou rendre la justice ? Il s'agit

---

1 E. Kant, *Logique*, Paris, Vrin, 1997, § 36.

2 R.Dworkin, *Prendre les droits au sérieux*, Paris, PUF, 1995.

3 A. de Laubadère, *Droit administratif*, Tome 1, LGDJ, 2001.

4 R. Saleilles, *L'individualisation de la peine, étude de criminalité sociale*, Paris, Alcan, 1898.

5 H.L.A. Hart, *Le concept de droit*, Bruxelles, Publications des Facultés universitaires Saint Louis, 1980.

6 G. Kalinowski, *Introduction à la logique juridique*, Paris, LGDJ, 1968.

[p. 37-53]

D. BOURCIER

*Arch. phil. droit* 54 (2011)

aussi de confronter ou de réconcilier deux visages du droit et de la justice, calcul objectif ou raisonnement subjectif, droit prévisible ou droit flexible. La querelle autour du syllogisme judiciaire qui s'est développée depuis l'École de l'Exégèse a réhabilité la rhétorique au moment même où se développait une nouvelle branche de la logique, la logique déontique : pour Perelman, le droit, truffé d'antinomies et de lacunes interdit simplement au raisonnement judiciaire de n'être qu'un syllogisme et d'être mu par des opérateurs modaux.

Ces mêmes oppositions continuent à alimenter les nombreux débats qui se posent autour de la décision artificielle<sup>7</sup> : en effet les sciences de l'artificiel pénètrent toutes les sphères de la décision (médecine, économie gestion) Le droit ne pouvait y échapper. Ce rêve de *mathesis universalis* a parcouru toute la pensée occidentale depuis Ramus et Leibniz. En 1748, un médecin, philosophe des Lumières, publiait à Leiden l'*Homme machine*<sup>8</sup>. Sa principale thèse était d'opposer le déterminisme de la machine – l'homme fait partie de la machine en question – à l'exercice d'une quelconque volonté interne. L'opposition entre la machine et la volonté définie comme la faculté pour l'être humain de n'être soumis qu'à son libre arbitre (sans être contraint par une quelconque « force extérieure » comme le dit Descartes) est devenue depuis trois siècles le thème central du débat qui allait opposer les monistes et les dualistes.

L'homme machine c'est le titre qu'a donné La Mettrie au livre pamphlet de cette ode au mécanisme intégral. Malgré son radicalisme, il anticipe les débats qui se sont développés dans la deuxième partie du XX<sup>e</sup> siècle autour de la robotique, de l'intelligence artificielle et des neurosciences. Ce n'est plus le mécanisme qui intéresse les ingénieurs philosophes mais la machine cognitive : peut-elle simuler des comportements humains, comme l'exercice du jugement, du raisonnement, de la mémoire ? En réponse à cette question, une hypothèse théorique va se dessiner avec le développement de l'informatique : si le cerveau fonctionne peu ou prou comme une machine (de Turing), compte tenu de moyens de calculs nouveaux offerts par l'ordinateur actuellement, des « organismes formels » ne peuvent-ils apprendre et effectuer des actions proches (voire plus satisfaisantes car mieux informées) de celles que réalisent les humains ?

Un autre débat critique allait émerger dans les sciences cognitives sur l'incarnation de l'esprit (*embodiment of mind*) qui apporta un contre-argument au cerveau-machine. L'esprit pour être opérationnel doit être « incarné » dans le corps matériel et organique. Autrement dit, l'algorithme représentationnel ne suffit pas à simuler la pensée, il faut le contexte physique et corporel dans lequel elle s'inscrit : comment transposer cette hypothèse dans

7 D. Bourcier, *La décision artificielle, le droit la machine et l'humain*, Paris, PUF, 1995.

8 J. Offray de La Mettrie, *L'homme machine* (1747), Éditions Bossard, 1921.

[p. 37-53]

D. BOURCIER

*Arch. phil. droit* 54 (2011)

un programme d'ordinateur ? L'argument devient difficile à contredire et ni les neurosciences ni les sciences cognitives n'ont résolu cette impasse<sup>9</sup>.

En ce qui concerne le droit, les mêmes métaphores ont accompagné le développement de ses formalisations. Pour Frédéric de Prusse déjà, le code devait être aussi précis qu'une horloge. Ce n'est donc pas la première fois qu'avec l'intelligence artificielle, se pose la question du droit comme machine. De nombreux travaux de cybernétique puis de systémique depuis Leibniz tentent de faire du droit un mécanisme<sup>10</sup>.

En 1959 paraissait à Londres un ouvrage sur la mécanisation des processus de la pensée (*Mechanisation of the Thought Process*). Lucien Mehl, qui a lancé la juricybernétique en France dans les années 50 continuait les conférences Macy<sup>11</sup> bien connues des cybernéticiens Heinz von Foerster et John Mac Carthy et intitulait son chapitre « Automation in the legal world ». Il distinguait deux types de *law machines* : la machine à informer et la machine à juger. On retrouve cette distinction plus de cinquante ans après. Les technologies après avoir été utilisées pour des opérations relativement simples (des tris et des calculs comme la recherche d'un mot-clé dans une base de textes) simulent des activités cognitives beaucoup plus nobles dans l'échelle des activités humaines comme apprendre, traduire, apprécier, qualifier, évaluer, diagnostiquer, décider, juger. En quelque sorte ces machines ont entrepris de se substituer à ce qui nous caractérise en tant qu'humain, des tâches de haut niveau qui impliquent notre libre arbitre.

Dans une première partie, j'examinerai la question de la capacité *technique* des machines à simuler l'acte de juger. Dans une deuxième partie, la capacité *juridique et éthique* de ces machines à simuler un tel acte.

## I. — LES MACHINES ONT-ELLES LA CAPACITÉ TECHNIQUE DE SIMULER L'ACTE DE JUGER ?

Une machine n'a en général pas la capacité technique de remplir une tâche humaine de haut niveau cognitif mais elle peut en réduire la complexité<sup>12</sup> pour la simuler. Les travaux d'apprentissage cognitif des tâches physiques ou mentales lancés dans les années 70-80 ont conduit à des recherches complètement

9 P. Johnson-Laird, *The Computer and the Mind : An Introduction to Cognitive Science*, Cambridge, MA : Harvard University Press, 1988.

10 L. Mehl, « Automation in the legal world. From the machine processing of legal information to the law machine », *Mechanisation of Thought processes*, vol. 2, National physical Library, Symposium n° 10, 1959.

11 J.-P. Dupuy, *Aux origines des sciences cognitives*, Paris, La Découverte, 1994 – à propos des Conférences Macy.

12 La science des systèmes complexes vise à maintenir et traiter cette complexité plutôt que la réduire. Voir, en droit, D. Bourcier, R. Boulet, P. Mazzega (éds), *Complexité et politiques publiques*, Paris, Hermann, 2012 (à paraître).

[p. 37-53]

D. BOURCIER

*Arch. phil. droit* 54 (2011)

nouvelles qui en réalité ont consisté à approfondir la description opératoire des activités humaines comme première étape de la simulation par une machine. La deuxième étape sera confiée à l'ingénieur cognitif qui concevra un logiciel qui transposera ce *travail* cognitif en programme d'ordinateur. La tâche humaine (traduire, catégoriser, évaluer, décider, justifier,...) devra être correctement décrite pour que le résultat fourni par la machine ne soit pas différent de ce que ferait un expert humain dans les mêmes conditions. On appelle d'ailleurs ces dispositifs des *systèmes experts*.

La modélisation consiste à partir d'une conceptualisation de ce qu'est la décision du juge, de tenter d'en simuler les différentes démarches cognitives dans une machine et d'en évaluer les limites et les impacts en milieu ouvert. Que l'on veuille mécaniser ou simuler des activités ne signifie pas qu'il faille procéder à un réductionnisme mécanique des opérations de jugement ni qu'il faille remplacer définitivement les procédures institutionnelles par des processus automatiques. Dans tous les cas, la question scientifique et surtout cognitive de la représentation décisionnelle mérite une vaste réflexion.

Cette double démarche théorique et applicative permet d'observer les effets de tels outils dans le monde institutionnel auquel ils sont destinés. Se soumettre à l'application complète mais aussi à l'expérimentation en lieu réel permet d'apporter un retour d'expérience efficace. Des laboratoires du CNRS<sup>13</sup> au carrefour du Droit et des Sciences cognitives ont été impliqués dans de tels projets avec des spécialistes de l'Intelligence artificielle mais aussi avec les professionnels du droit opérant dans leur contexte institutionnel durant tout le processus d'élaboration. Ce fut là une exigence épistémologique et technique qui devait assurer une complète évaluation de ces nouveaux outils et de leurs usages.

Deux modèles ont été explorés en intelligence artificielle. Les modèles à base de règles (les systèmes experts) et les modèles à base de cas (les systèmes neuronaux ou connexionnistes). Ils correspondent aux deux modes d'apprentissage dégagés par Piaget<sup>14</sup> dans la psychologie du développement de l'enfant – par les règles et par les exemples. Mais ils renvoient aussi à deux systèmes de droit – le monde romano-germanique légicentré et le monde de la *Common Law* à raisonnement casuistique.

Je ferai l'hypothèse, après ces années d'expérience, que l'activité de juger en droit correspond aux contraintes suivantes (*check-list* que l'on peut aussi rapprocher des cas d'ouverture du recours pour excès de pouvoir) : identifier si le décideur-utilisateur est compétent, sélectionner des données de faits et de droit, qualifier les faits sous la règle, utiliser le pouvoir d'appréciation et d'interprétation à la fois des faits, du contexte et des règles, motiver de façon spécifique, et citer les références légales et valides. C'est un tel *design* qui sera utilisé pour construire des systèmes d'aide à la décision capables de

13 UA 430 *Informatique Droit linguistique*, Conseil d'État, Palais royal (1984-1996) et CERSA (Université de Paris 2).

14 J. Piaget, B. Inhelder, *La psychologie de l'enfant*, Paris, PUF, 2011.

[p. 37-53]

D. BOURCIER

*Arch. phil. droit* 54 (2011)

« simuler » l'activité et les connaissances dont se sert l'expert humain. Le modèle connexionniste tend à résoudre des questions de flexibilité de la décision et a vocation à s'intégrer aux systèmes du premier type fondés sur des règles explicites<sup>15</sup>.

#### A. — *Justice représentationnelle : les systèmes à base de règles*

Ces systèmes experts ou à base de connaissances se sont développés à partir des années 2000. On en a répertorié en France plus d'une centaine relevant des systèmes d'aide à la décision juridique<sup>16</sup>. Ils sont désormais le plus souvent intégrés dans des systèmes informatiques de gestion et de contrôle, pour des profils de comportements, des évaluations de risques, des bases d'apprentissage, ou des activités de formation.

#### **Quel est leur mode cognitif de fonctionnement ?**

Ces systèmes sont fondés sur la séparation entre la base de connaissances et le mécanisme de raisonnement. Leur finalité est d'apporter une réponse ou une *solution prospective, créative* à un problème posé. Ils se distinguent des systèmes documentaires (comme Legifrance) qui apportent une réponse *rétro-spective* comme trouver un ensemble de documents pertinents (des cas ou des règles) qui répondent à des mots-clés représentant un problème. Les bases de connaissances compilent les règles d'un domaine susceptibles d'être utilisées par le juge ou le décideur pour résoudre un problème juridique ou proposer une solution (la décision comme *problem solving*, suivant la terminologie de Popper). Les connaissances sont représentées sous forme de règles de production (ou d'implication) en « si A, alors B » c'est-à-dire telles conditions impliquent telle conclusion :

Ex. du Système BRUITLOG

SI chantiers publics ET SI non-respect des conditions d'utilisation ET  
dépassement des valeurs d'émergence ALORS contravention d'atteinte  
à la Tranquillité Publique SAUF si bruit ambiant < 30 dB.

Les faits et données (y compris des mesures) sont fournis au système (base de faits). Le mécanisme de raisonnement parcourt en fonction des faits donnés par l'utilisateur le réseau des règles qui s'appliquent au cas jusqu'au moment où aucune règle n'est plus candidate : la dernière règle explorée donne la solution à l'arbre de décision. La machine conclut. Le raisonnement se fait en chaînage avant (hypothético-déductif) et en chaînage arrière (inductif ou régressif).

15 D. Bourcier, « Modéliser la décision administrative, réflexions sur quelques paradigmes » in *Droit administratif en mutation*, Paris, PUF CURAPP, 1993.

16 D. Bourcier & J.-P. Costa, *Systèmes experts et systèmes d'aide à la décision*, Paris, STH, 1992.

[p. 37-53]

D. BOURCIER

*Arch. phil. droit* 54 (2011)

Ainsi BRUITLOG<sup>17</sup>, un système simulant la décision du maire en matière de police du bruit pouvait proposer plusieurs solutions et déclenchait plusieurs activités face à une nuisance sonore : poursuivre au pénal, lancer des mesures d'information ou de médiation, rédiger un arrêté municipal.

### En quoi sont-ils aptes à représenter la spécificité de l'acte de juger ?

Ces systèmes sont des systèmes fondés sur des règles ré-écrites à partir de la règle de droit : cet isomorphisme (règle de droit/règle de production) facilite la représentation des règles mais aussi des concepts (un concept est de façon analytique considéré comme une règle). Ils sont donc adaptés *a priori* à la forme de la connaissance juridique.

Le raisonnement utilisé est *hypothético-déductif* mais aussi inductif. On peut partir des faits établis mais, lorsque le juge connaît intuitivement la solution mais ne sait pas comment y arriver (par quels faits, par quelles règles), le système peut partir de la conclusion supposée et vérifier auprès du décideur si toutes les conditions sont remplies pour cette conclusion : il peut aussi lui suggérer de sélectionner telle condition qui lui permette d'arriver à une décision satisfaisante au sens de H. Simon. L'activité de raisonnement du juge est liée aussi à l'activité de *justification*. Ces systèmes permettent de restituer les règles qui ont été activées ainsi que les enchaînements entre les règles. Enfin les systèmes experts peuvent *motiver* parce que tous les faits et règles utiles sont stockés au cours de la session et sont donc restituables. C'est le raisonnement *ad hoc* suivi par la machine qui sert à restituer un « raisonnement juridique ».

Revenons au cœur de la prise de décision juridique. La spécificité de l'acte de juger réside dans le fait que des parties de la décision sont « calculables » (modélisation logique) et d'autres ne le sont pas (langage ouvert). Ainsi le langage naturel dans lequel sont écrites les règles de droit implique toujours une autre interprétation possible. Le métalangage juridique est caractérisé par son rôle pragmatique dit illocutionnaire ce qui signifie que le discours en langage naturel, en étant formulé dans certaines conditions « réalise » certaines opérations cognitives. Pour aider le décideur, des actions d'interprétation ont été mises à sa disposition : par le déploiement des définitions en extension ou en compréhension, et de raisonnements extensifs ou analogiques déjà réalisés. On peut apporter ainsi, sinon un résultat décisionnel, une aide à l'acte d'interprétation. Ainsi parmi les phénomènes non calculables, vont être représentés les concepts flous (comme la notion *d'activité inadaptée* ou de *danger*) et le pouvoir discrétionnaire. Ces deux types d'opérations linguistico-cognitives interrompent la chaîne des raisons et des causalités car elles s'effectuent sans règles *a priori*. Le système ne sait donc pas comment les représenter.

Les concepts flous vont alors être traités au cours du dialogue en proposant les interprétations variées qui ont déjà été activées. Au cas où il n'y a pas assez

---

17 Système élaboré au laboratoire Informatique Droit Linguistique avec le soutien de S.A. Lyonnaise des eaux (Laboratoire d'Intelligence Artificielle, Compiègne).

[p. 37-53]

D. BOURCIER

*Arch. phil. droit* 54 (2011)

d'interprétations enregistrées, le décideur va introduire sa propre interprétation. Cette interprétation pourra être incorporée dans la base ultérieurement (accroissement *dynamique* des connaissances).

Quant au pouvoir discrétionnaire, on peut le définir comme un pouvoir qui s'exerce sans règle explicite : il existe une jurisprudence du Conseil d'État qui sanctionne d'ailleurs la décision discrétionnaire qui fixerait les règles dont elle serait issue. Dans le cas d'exercice du pouvoir discrétionnaire (ou non lié), le système interrompt la chaîne des questions et s'adresse directement au décideur : « y a-t-il atteinte à la tranquillité publique ? Oui ou non ? ». En fonction des faits (des mesures en cas de bruit par exemple) ou de réponses précédemment fournies par le décideur, le système va traiter le degré de cohérence de la description et suggérer que sa décision POURRAIT être annulée.

### B. — *Justice artificielle : Les systèmes à base de cas*

Ces systèmes sont appelés connexionnistes, ou réseaux de neurones artificiels (RNA). Ils se déclinent aussi sous le nom de *neurolaw*. Ils sont utilisés par exemple en Australie dans les procédures de médiation ou de décision (calcul de pension alimentaire ou répartition des frais d'entretien de l'enfant) ou en France, dans le calcul d'indemnité de licenciement.

#### **Quel est leur mode cognitif de fonctionnement ?**

Le système à base de cas est par définition un système qui raisonne sans règles. Sa connaissance vient donc d'une base d'exemples : soit les règles ne sont pas données par le législateur, c'est l'indétermination légale (pas de définition de *l'ordre public* dans le Code général des collectivités locales) soit parce qu'il y a une variété de paramètres qu'il est difficile de représenter (qu'est-ce que *gérer en bon père de famille* ?). Ce que les organismes vivants arrivent à résoudre globalement en s'adaptant sans décliner les types d'interactions et causalités possibles relève de ce que Alain Berthoz en biologie appelle la *simplicité*<sup>18</sup> : ce syncrétisme peut être utile pour décrire des évolutions juridiques sans règles.

On crée donc, pour un domaine de décision donné, une base d'apprentissage de cas (de la jurisprudence ou des dossiers) qui va constituer la base de référence. Ce phénomène est parfois comparé à ce qui se passe pour l'apprentissage humain. Ces données fournies en vrac vont former la couche d'entrée dans le système. En sortie, soit le système est binaire, (coupable, non coupable) soit il choisit entre plusieurs solutions (diversité des sanctions). L'action du réseau consiste donc à convertir la valeur de ses entrées en un signal qui se propage le long de ses connexions. Ce signal subit plusieurs

18 A. Berthoz, *La simplicité*, Paris, O. Jacob, 2009.

[p. 37-53]

D. BOURCIER

*Arch. phil. droit* 54 (2011)

traitements mathématiques (notamment le passage par une fonction sigmoïde) et finit par être envoyé par le neurone de sortie sous la forme d'une valeur que l'expérimentateur peut interpréter. L'hypothèse que l'on pose est que, dans le cas d'une sortie binaire (allant de 0 à 1), certains neurones ont, via les connexions, tendance à diminuer le signal alors que d'autres ont tendance à l'augmenter. En faisant la somme de l'influence de chaque connexion relative à chaque entrée et en les comparant les unes aux autres, on peut déterminer le sens de l'influence de chaque entrée (le sens global de la source) et l'importance relative de leur influence sur le résultat final.

La base de cas doit avoir une certaine stabilité, un poids calculé entre les différents cas et les mots d'entrée. Quand cette base est stable, qu'elle a appris les poids de chacun des faits dans l'ensemble des vecteurs connectés à l'intérieur du réseau et qu'elle a créé ses couches intermédiaires, on dit qu'elle a *appris*. On peut alors lui fournir un jeu de données pour tester ce que le système peut « décider » en fonction des cas enregistrés : l'acte était-il légal ? La décision était-elle conforme ?

Nous avons utilisé cette modélisation dans plusieurs expériences. NEUROLEX était un système qui était capable de décider « comme » le Conseil d'État dans le champ de la police municipale. Une autre expérience a été menée avec une Cour d'appel sur les motifs de licenciement. Dans ce dernier cas, l'algorithme (une réduction des situations possibles) a été implémenté dans un réseau de neurones modélisant un contentieux faisant appel à 14 critères. La conclusion du contentieux est binaire : applicabilité ou non applicabilité d'une clause de non-concurrence dans un contrat de travail.

Un test sur 30 situations différentes sélectionnées au hasard et comparées avec les justifications proposées par un magistrat a permis de donner un aperçu des performances de l'algorithme. Le contentieux qui a été « enseigné » au réseau de neurones est celui de l'applicabilité des clauses de non-concurrence, tel qu'il se présentait en juin 2002.

Quatre conditions doivent être réunies pour que la clause soit applicable :

1. La clause doit avoir été portée à la connaissance du salarié (3 critères)
  - a. Soit par le contrat de travail
  - b. Soit par la convention collective, auquel cas le salarié doit être informé de son existence au plus tard le jour de l'engagement
2. La clause doit protéger les intérêts légitimes de l'entreprise, sans être excessive (regroupant 5 critères)
3. La clause ne doit pas avoir été levée par le salarié (4 critères)



Celui-ci peut la lever si le contrat ou la convention collective prévoyait le paiement d'une contrepartie financière et que celle-ci n'a pas été payée

4. La clause ne doit pas avoir été levée par l'employeur (2 critères)

Celui-ci peut la lever si une telle procédure est prévue dans le contrat de travail ou dans la convention collective

Dans chaque exercice, l'algorithme a montré qu'il parvenait à reconnaître sans erreur les critères qui avaient eu une influence positive sur le contentieux (clause applicable) de ceux qui avaient une influence négative (clause non applicable).

#### **En qui sont-ils aptes à représenter la spécificité de l'acte de juger ?**

Quand on travaille sur les questions de transposition de règles de droit en règles informatiques, on s'aperçoit rapidement qu'on ne peut traduire toute la loi sous forme de règles<sup>19</sup>. Même si c'était possible, on s'apercevrait – comme on vient de le signaler plus haut – que les éléments de base de la loi (son ontologie) peuvent être flous (*activité dangereuse*) ou indéterminés (*véhicule à moteur*) ou discrétionnaires (*atteinte à l'ordre public*) et qu'en réalité, la loi délègue au juge le soin d'ajouter des règles implicites d'interprétation de ces éléments face à des cas réels : c'est ce qu'on appelle la marge d'appréciation du juge. Les RNA sont particulièrement adaptés à ces connaissances floues, incertaines et bruitées. De nombreux tests ont été faits pour utiliser des modèles connexionnistes dans le traitement du droit<sup>20</sup> et l'on a reconnu leur pertinence pour traiter les parties les plus ouvertes du raisonnement malgré les critiques qui leur sont faites.

Indépendamment du fait que ces systèmes sont bien adaptés au régime de *common law* (plutôt fondé sur des cas de jurisprudence), ils sont utiles quand on n'a pas de règles explicites qui fondent la décision. On leur a reconnu d'autres intérêts : ils peuvent intégrer des cas fondés sur des raisonnements prenant en compte des faits variés et des pondérations différentes, c'est-à-dire des décisions divergentes.

Leur principal défaut réside dans leur difficulté à exprimer les règles sur lesquelles sont construits leurs résultats (la représentation est « mathématique »).

Dans certains domaines (reconnaissance des visages), l'expérimentateur n'a pas besoin de savoir quel pixel ou quel neurone a fait basculer la décision du réseau. En revanche la justification est fondamentale en droit, pour le magistrat et pour les parties.

<sup>19</sup> Il reste notamment les principes, qui nécessitent une autre représentation.

<sup>20</sup> L. Bochereau, D. Bourcier, P. Bourguin, « Extracting legal knowledge by means of a multilayer neural network application to municipal jurisprudence », *Proceedings ICAIL'91*, New York, ACM, 1991.

La lecture d'un arrêt montre que le juge se justifie généralement *en citant les éléments de faits ou de droit qui ont orienté sa décision*. Telle clause n'est pas applicable car le contrat de travail prévoyait le paiement d'une contrepartie financière et que celle-ci n'a pas été payée ; tel licenciement est sans cause réelle et sérieuse car la lettre de licenciement n'énonce pas de motifs réels et précis ; tel privilège de juridiction est écarté car une convention prévoit des règles de compétences particulières. En ce sens, une justification satisfaisante émanant d'un modèle connexionniste consisterait à ce que ce système nous informe des éléments de fait ou de droit (critères) qui ont fait basculer sa décision.

Ce qui est aisé pour un système-expert (comme justifier son raisonnement) ne l'est pas pour un réseau de neurones, et ce, pour une raison de structuration de la connaissance. Le réseau de neurones disperse l'information à travers sa structure et le poids de ses connexions. Ceci rend particulièrement complexe la procédure de récupération des informations pertinentes. Cependant les RNA peuvent mettre en évidence des processus cognitifs plus fins<sup>21</sup>. Dans certains exemples, le RNA a utilisé un « raccourci » lui permettant de généraliser de façon à obtenir de bons résultats sur la totalité des exercices (les 16 384 combinaisons ont été testées dans le cas de l'applicabilité de la clause de non-concurrence). Le réseau de neurones a pu se servir de sa puissance de généralisation de façon à échapper à la justification que l'on voulait lui imposer et à trouver, au moyen des autres critères d'entrée, une nouvelle manière cohérente de parvenir à un résultat satisfaisant. On lui accorde ainsi une propension à prédire de nouveaux raisonnements ou simplement à en adopter de plus efficaces.

Ainsi, la modélisation de la décision juridique pour un traitement par une machine cognitive – et non mécanique – est par définition toujours réductrice mais conduit à expliciter les processus décisionnels que l'argumentation juridique traditionnelle ne permettait pas de faire.

## II. — LES MACHINES PEUVENT-ELLES JURIDIQUEMENT SIMULER L'ACTE DE JUGER ?

Cette question recouvre les aspects juridiques de la décision artificielle et les aspects éthiques regroupés autour d'une nouvelle discipline : l'éthique computationnelle.

### A. — *L'encadrement juridique de la décision*

La plupart des machines dont nous venons de parler traitent des données personnelles, d'informations individuelles et de traitements discriminatoires.

21 F. Borges, R. Borges, D. Bourcier, "Artificial neural networks and legal categorization", in *Legal knowledge and Information systems*, JURIX 2006, Amsterdam, IOS Press, 2006. [p. 37-53]

D. BOURCIER

*Arch. phil. droit* 54 (2011)

Ces machines en traitant de données personnelles opèrent en effet des raisonnements particulièrement sensibles<sup>22</sup>. En outre, ces raisonnements ne sont pas *a priori* disponibles et accessibles. Ce traitement de l'intimité par ce type de machine implique des précautions. Ce n'est donc pas un hasard si la première loi Informatique et Libertés en 1978<sup>23</sup> a visé ce type de traitement et plus généralement l'informatisation de la décision quand celle-ci « traite » des données psychologiques ou cognitives liées à la vie privée.

Il existe depuis la loi de 1978 modifiée un droit subjectif nouveau : le « droit de ne pas être soumis à une décision automatisée ». Une jurisprudence est en train de se constituer peu à peu à travers la Commission nationale Informatique et Libertés et les tribunaux. On attend un recours spécifique.

Que dit la loi de 1978 sur la décision automatique ? Des données personnelles figurent massivement dans les banques de données. Ces données sont prélevées sur les réseaux sociaux ou sur les fichiers commerciaux sans que le consentement (quant à son étendue) soit clairement demandé. En outre avec le mouvement mondial de l'Open Data qui consiste à rendre massivement accessibles des données, des fichiers sont constitués systématiquement et pour des finalités différentes des objectifs premiers (comme les embryons surnuméraires). Ces données (génétiques, de modes de vie ou de consommation) peuvent très facilement être analysées et utilisées pour établir le profil d'une personne et extrapoler sur ses intentions, ses actions ou même ses évolutions possibles. Ce déterminisme qu'implique la décision d'une machine au sujet du comportement prévisible ou extrapolable d'un être humain est attentatoire au principe de libre détermination qui est au cœur du principe de liberté.

L'utilisation de données personnelles n'est pas interdite *a priori* pour les décisions non judiciaires, notamment si le résultat fait l'objet d'une intervention humaine. Ce qui l'est, en revanche, c'est d'utiliser un tel procédé automatisé, qui évalue certains aspects de la personnalité, comme base pour une décision négative concernant la personne, et qui aurait pour elles des conséquences dommageables.

Nous pouvons comparer les deux étapes de ce « droit à ne pas être soumis aux décisions automatisées » entre la loi de 1978 et la même loi modifiée et transposée.

---

22 D. Bourcier, « Données sensibles et risque informatique », in *Données sensibles*, Paris, PUF, CURAPP, 1998.

23 La loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés a été modifiée et transposée par la loi n° 2004-801 du 6 août 2004 relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel a été publiée au *JO* du 7 août 2004.

Prenons la loi n° 78-17 du 6 janvier 1978 :

Article 2 (ancien)

« Aucune décision de justice impliquant une appréciation sur un comportement humain ne peut avoir pour fondement un traitement automatisé d'informations donnant *une définition du profil ou de la personnalité de l'intéressé*.

« Aucune décision administrative ou privée impliquant une appréciation sur un comportement humain ne peut avoir pour seul fondement un traitement automatisé d'informations donnant une définition du profil ou de la personnalité de l'intéressé. »

La décision de justice ne peut en aucun cas être fondée sur un dispositif artificiel même s'il y a intervention d'un juge sur le résultat obtenu. Il devra être clair que ces éléments issus de la machine n'ont pas « influencé » outre mesure sa décision finale.

Article 3 (ancien).

« Toute personne a le droit de connaître et de contester les informations et les raisonnements utilisés dans les traitements automatisés dont les résultats lui sont opposés. »

Cet article consacre le droit à connaître les justifications, connaissances et raisonnements utilisés. On remarquera que nos systèmes experts étaient fondés sur ces contraintes.

La loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés a été modifiée et transposée par la loi n° 2004-801 du 6 août 2004 relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel et modifiant. Elle a introduit quelques modifications en ce qui concerne la décision :

Article 10 (nouveau)

Aucune décision de justice impliquant une appréciation sur le comportement d'une personne ne peut avoir pour fondement un traitement automatisé de données à caractère personnel destiné à évaluer certains aspects de sa personnalité.

Aucune autre décision produisant des effets juridiques à l'égard d'une personne ne peut être prise sur le seul fondement d'un traitement automatisé de données destiné à définir le profil de l'intéressé ou à évaluer certains aspects de sa personnalité.

Dans cette formulation, la décision administrative et la décision privée sont regroupées sous le label « décision produisant des effets juridiques ». Le terme « éléments de sa personnalité » est restrictif : s'agit-il des « données dites

sensibles » ?<sup>24</sup> Les données sur la santé sont-elles considérées comme des données destinées à évaluer certains aspects de la personnalité ?

Ne sont pas regardées comme prises sur le seul fondement d'un traitement automatisé les décisions prises dans le cadre de la conclusion ou de l'exécution d'un contrat et pour lesquelles la personne concernée a été mise à même de présenter ses observations, ni celles satisfaisant les demandes de la personne concernée.

Le rôle du consentement du destinataire – ou son intérêt – est fondamental dans la qualification de « décision automatisée ». C'est pourquoi, cette interdiction ne s'applique pas aux décisions prises dans le cadre d'un contrat (action volontaire) qu'il s'agisse d'un prêt ou de la souscription à une assurance. Les autres décisions visées concernent surtout les décisions administratives (non contractuelles) c'est-à-dire les actes unilatéraux, à moins qu'une intervention humaine n'interrompe la chaîne déductive. Enfin, la décision judiciaire reste exclue de ce genre de machine.

Nous avons à partir de l'analyse de ces règles de droit dégagé plusieurs postulats concernant le statut de la décision automatisée :

- le résultat du raisonnement opéré par une machine doit être distingué de la décision humaine ;
- le décideur doit être identifiable dans l'organisation puisque c'est lui qui assumera la responsabilité du traitement ;
- le destinataire doit être mis au courant du traitement automatisé le concernant et avoir accès aux connaissances et aux raisonnements ;
- le maître d'ouvrage (et non l'informaticien maître d'œuvre) est responsable puisque c'est lui qui en contrôle le fonctionnement et les règles, et à qui on a livré le système. Concrètement cette règle est difficile à appliquer car, compte tenu de la combinatoire exponentielle, le bêta-test ne peut pas être entièrement réalisé. Par défaut, en respectant les schémas hiérarchiques de décision, on peut aussi présumer que le responsable de la décision est le responsable du service où ces technologies sont utilisées.

### B. — Une nouvelle discipline : l'Éthique computationnelle

Cette récente discipline est née en 1985 avec le livre *Computer Ethics* de D. Johnson<sup>25</sup>.

La définition de l'éthique computationnelle est la suivante : « analyse de l'impact social des Technologies de l'Information ainsi que la formulation et la

<sup>24</sup> D. Bourcier, in *Données sensibles*, CURAPP, PUF, *op. cit.*

<sup>25</sup> D. Johnson, *Computers Ethics*, Prentice Hall, 2008 (4th edition).

[p. 37-53]

D. BOURCIER

*Arch. phil. droit* 54 (2011)

justification de politiques pour l'utilisation éthique de telles technologies ». De nombreuses questions de responsabilité se posent du côté des impacts de la décision. Mais c'est la perte d'autonomie qui est signalée par les utilisateurs ainsi que la dilution des compétences<sup>26</sup>. Ces dangers ont incité le comité d'Éthique des Sciences du CNRS (COMETS) à rédiger en 2010 un rapport sur ce thème<sup>27</sup> : il a énoncé la Recommandation suivante qui concerne la décision :

« Recommandation 7. Encourager en France les recherches en STIC dans des domaines clés relatifs aux questions d'éthique. *C'est en particulier le cas des recherches sur l'éthique des machines (Moral Machines) et en informatique sociale (Social Informatics)... D'autres domaines (cognition numérique, agents intelligents,...) sont également concernés.* »

#### **A-t-on le droit d'utiliser des outils de décision automatisés ?**

La question de la compétence professionnelle et la responsabilité éthique et déontologique de l'utilisation de machines *intelligentes* sont devenues des sujets de discussion fondamentaux. Un cas réel a été longuement étudié dans les cercles éthiques : une infirmière se sert d'un système expert dans un service d'urgence obstétricale puis fait de mauvaises interprétations des résultats fournis. L'enfant meurt. Ce système était dédié à des médecins, donc interrogé, et interprété par des experts. Le système expert doit être conforme à un niveau d'expertise et ce niveau doit correspondre à celui de l'utilisateur habilité. Dans le cas de l'infirmière, l'habilitation manquait dans le service.

#### **A-t-on le droit de les fabriquer ?**

Pouvons-nous – en tant que chercheur – assumer que la simulation serve seulement à satisfaire notre soif faustienne de connaissance ou est-on limité par notre responsabilité face à la réalisation d'un objet-machine qui pourrait agir de façon plus ou moins autonome dans un monde réel (Frankenstein). Il s'est élevé de nombreuses voix pour considérer par exemple que le droit ne pouvait être informatisé pour des raisons éthiques<sup>28</sup>.

Mais l'éthique peut se poser de façon inverse. Aujourd'hui, peut-on éthiquement se priver de créer des objets de simulation qui nous renseignent sur nos comportements cachés ou simplement sur nos choix implicites ? Cette interrogation ne prend-elle pas un sens particulier quand il s'agit de remettre

26 D. Bourcier, « Pour une éthique de l'intelligence artificielle. Responsabilité et informatisation » in *Sciences de l'homme et de la Société : la responsabilité des scientifiques*, Paris, L'Harmattan, 2001.

27 Rapport *Pour une Éthique de la Recherche en Sciences et Technologies de l'Information*, 2009 : [www.cnrs.fr/fr/organisme/ethique/comets/docs/rapportComets091112.pdf](http://www.cnrs.fr/fr/organisme/ethique/comets/docs/rapportComets091112.pdf).

28 J. Weizenbaum, *Puissance de l'ordinateur et raison de l'homme*, Paris, Éditions d'Informatique, 1981.

en cause l'exercice des principes fondateurs de nos sociétés : la justice, l'équité, l'égalité devant la loi.

Dans cette recherche sur le connexionnisme en droit, on a pu confirmer que les réseaux de neurones artificiels pouvaient être un modèle adéquat pour modéliser la connaissance juridique floue et qu'ils pouvaient compléter les systèmes à base de règles lors de l'exercice du pouvoir discrétionnaire du juge. Nous avons en outre pu lever un des obstacles majeurs à leur mise en œuvre en droit à savoir leur opacité en développant un outil de justification. Enfin, une autre vertu pourrait être attribuée au « juge connexionniste » : sa possibilité de mettre en lumière des tendances dans l'évolution de la jurisprudence ou de proposer des chemins de raisonnement plus courts. Pour ces raisons, le modèle neuronal peut être utilisé comme outil de simulation des comportements cognitifs du juge et comme heuristique pour tester les justifications fondées sur des règles, ou mieux, sur le sentiment de justice.

Cependant de nombreuses questions éthiques se posent dans la conception de projets de recherche en matière technologique : le pôle de compétitivité Cap Digital a créé un comité d'éthique pour décider s'il fallait supprimer certains projets considérés comme trop sensibles<sup>29</sup>. Aujourd'hui, les risques éthiques des recherches en matière de technologies avancées font l'objet d'une recherche autonome donnant lieu à un rapport d'évaluation particulier sur le process et sur les impacts possibles dans la plupart des projets européens et même dans certains projets nationaux.

### Où la règle de droit devient une règle de gestion

Mais l'e-gouvernement et l'e-justice ne sont plus dans les mains des philosophes. On ne demande plus à la machine de nous imiter dans nos décisions : comme l'humain, son statut s'est dégradé et elle doit appliquer des tâches de gestion définies par le management public avec les contraintes budgétaires mécaniques y afférant. Le droit n'a plus besoin d'être représenté, il suit les impératifs gestionnaires. En 2011, l'exemple emblématique de la vie quotidienne est le *radar automatique*. Attardons-nous sur ce système du point de vue de la décision automatisée. Les radars automatisés ont réuni deux types d'automatisation. L'automatisation de la phase de jugement et l'automatisation complète de la chaîne pénale. En ce qui concerne la phase de jugement, le prélèvement de points sur un permis ne passe plus par le jugement des juges. On citera la décision du TGI de Tarbes (1992) : « Le juge se trouve empêché de remplir déontologiquement ce qui constitue la spécificité de sa mission à savoir la modulation de la peine et sa libre appréciation de la faute commise ».

Mais le radar simule aussi une autre fonction de la e-justice : il gère automatiquement la chaîne judiciaire, là où interviennent plusieurs professionnels du droit : d'abord la phase de constatation des faits (remplie par le policier ou

---

29 Un des projets acceptés était fondé sur un recueil de données sur les « sentiments ».

le gendarme), celle de l'expertise (prise de mesures factuelles et interprétation minimale), celle de la qualification et de jugement (assurée par le juge) ainsi que celle de la sanction (pénale et administrative), celle de la communication du jugement et de l'exécution de la peine (assurée par l'huissier).

Ce cas type est intéressant car il réduit l'acte de jugement à une activité purement mécanique sans que n'intervienne aucune interprétation humaine *a priori*, qu'elle soit factuelle ou juridique. Tous les débats politiques et juridiques dont a été entourée cette politique publique de sécurité routière – quels que soient ses effets sur la sécurité elle-même – dissimule mal l'élément fondamental et implicite qui en définit l'extrême modernité : c'est le premier système de justice automatique entré en application depuis les premiers engrenages de « machine à juger ». Cette application totale (mais limitée à un domaine précis) n'a donné lieu paradoxalement à aucun débat dans le domaine de la logique et de la philosophie du droit. Revenons au droit et à l'e-justice. La machine à juger actuelle, avec le développement des logiques non monotones, des systèmes complexes, et de la théorie du chaos s'est enrichie. Mais comment représenter l'intime conviction dans les engrenages logiques de l'ordinateur ou justifier le raisonnement sur l'absolue nécessité du bon juge Magnaud dans un système expert ? La loi Informatique et Libertés, les Chartes d'Éthique suffiront-elles à en limiter les applications mécaniques et injustes ?

## EN CONCLUSION

Dans le monde de la décision politique, un système expert B6 est désormais utilisé par les gouvernants. Conçu dès les années 80 par une équipe de chercheurs américains et japonais de l'université Wasatch, à Tokyo, B-6 est un outil capable d'analyser une situation et d'apporter une réponse à un problème. Martin Melva, l'un des concepteurs de B-6 en explique le mécanisme : « C'est assez simple. Ce logiciel, comme tous les logiciels d'aide à la décision, repose sur deux éléments : une grosse base de données et un moteur d'inférence. Ça veut dire qu'à partir de certaines règles, des pronostics vont être établis si on change tel ou tel paramètre. »

Ces logiciels sont utilisés dans la plupart des secteurs, surtout en finance, en médecine... Mais B-6 est aussi utilisé par des gouvernements. C'est actuellement le seul capable d'assimiler des données aussi complexes et complètes que celles d'un pays entier ! Le logiciel B-6 travaille sur un modèle à la fois économétrique et « sociométrique » : il simule les conséquences de différentes actions sur ce modèle.

D'un point de vue juridique et éthique, ces machines sont opaques dans leur conception et leurs applications et pourtant elles peuvent avoir un grand pouvoir d'influence dans la décision. Il convient de réfléchir de façon plus générale à la façon dont ces systèmes comme juges mais aussi comme législateurs et gouvernants peuvent être contrôlés. Citons ce que Ph. Souleau annonçait en 1969 au 5<sup>e</sup> colloque des IEJ sur la *Logique judiciaire* dans un



chapitre intitulé La logique du juge : « D'emblée l'idée d'une machine à juger nous choque ; le jugement n'est il pas un des plus nobles privilèges de l'espèce ? Dans cette rumeur d'engrenages qui va se rapprochant, nous percevons instinctivement une offense à notre fierté, une menace à ce qui nous parait l'essentielle intégrité de l'homme, nous sentons monter l'assaut qui, craignons-nous, va emporter les derniers retranchements de l'homme ».<sup>30</sup> À l'heure d'Internet où l'information et les données sont devenues massivement disponibles sans que le raisonnement (inductif, probabiliste, stochastique,...) qui les manipule soit accessible ou justifiable, cette réflexion nous aide à méditer sur les rapports entre jugement et information et sur les modèles qui les relient.

daniele.bourcier@cersa.cnrs.fr



---

30 Ph. Souleau, « La logique du juge » in *La Logique judiciaire*, Paris, PUF, 1969.

[p. 37-53]

D. BOURCIER

*Arch. phil. droit* 54 (2011)