

Le centre d'informatique juridique

L'informatique juridique est « l'application au droit des techniques modernes permettant de mettre en mémoire et d'utiliser les divers éléments de l'information juridique ». Telle est la définition qu'en donne le lexique de termes juridiques publié chez Dalloz.

Il faut tout d'abord distinguer l'informatique juridique — qui est le traitement des textes et décisions judiciaires — de l'informatique judiciaire qui est quant à elle l'application du traitement informatique à l'administration de la justice et aux fonctions connexes.

En second lieu, il ne faut pas confondre informatique juridique et droit de l'informatique. L'informatique juridique est un aspect de l'informatique, appliqué au Droit ; le droit de l'informatique est un ensemble de règles qui seront applicables en matière d'informatique.

I. — Le traitement de l'information, en particulier dans le domaine juridique n'est pas une nouveauté. Les fichiers auteurs et matières d'une bibliothèque sont déjà un procédé élémentaire de traitement de la documentation imprimée. L'innovation réside dans le recours à des moyens automatiques.

Quelques organismes en France se sont intéressés au traitement par l'informatique de la documentation juridique : l'IRETIJ à Montpellier, qui enregistre sous forme d'abstracts, la jurisprudence de certaines cours d'appel ; le CRIDON de Lyon qui a vocation pour la documentation notariale et le CEDIJ, Centre d'Informatique Juridique.

Le CEDIJ est une association de la Loi de 1901, créé en 1970 par M. Lucien Mehl, Conseiller d'Etat, avec le concours de quelques ministères et organismes publics tels que le Ministère de la Justice, le Ministère de l'Intérieur, le Conseil d'Etat, la Cour de Cassation, etc. Le Centre est issu d'une équipe qui avait été constituée au Conseil d'Etat, par M. Mehl, en 1966.

Il a pour objectif l'automatisation de la documentation dans le domaine juridique. Il enregistre les textes législatifs et réglementaires ainsi que la jurisprudence et certains éléments de doctrine.

Les matières actuellement interrogeables sont les suivantes :

- Jurisprudence
 - (Conseil Constitutionnel
 - (Conseil d'Etat
 - (Tribunal des Conflits
 - (Tribunaux Administratifs
 - (Cour de Cassation

- Législation
 - (Commerciale (sociétés commerciales et faillite)
 - (Fiscale
 - (Circulaires fiscales
 - (Urbanisme
 - (Municipale
 - (Sociale (Code du Travail)

- Travaux parlementaires

Ces bases de données sont interrogeables à l'aide d'un programme conversationnel qui permet une recherche par étapes successives, selon un dialogue avec le système.

De nouvelles bases seront prochainement interrogeables :

- Recueil du droit de la fonction publique ;
- Mise à jour du droit de l'urbanisme (nouveau Code de l'Urbanisme) ;
- le nouveau Code des Communes 1977 (+ les textes sur les régions, les départements et les communes) ;
- Traités fiscaux bilatéraux.

A. — Les documents enregistrés sont regroupés en bases de données — ensemble de documents présentant une certaine unité. Par exemple : Code pour la législation, arrêts d'un ordre de juridiction, d'une juridiction ou même d'une chambre, pour la jurisprudence.

On peut voir une certaine ressemblance entre une base de données et un fichier de bibliothèque : ils regroupent l'un et l'autre un ensemble homogène de documents : classés soit par auteur, soit par matière.

Le CEDIJ a choisi d'enregistrer la législation en texte intégral et la jurisprudence sous forme de résumés en langage clair. Ce système n'exclut cependant pas une certaine indexation, le plus souvent combinée avec le résumé. Cette indexation et la confection des résumés sont effectuées à la source de l'information (en particulier pour la jurisprudence), ce qui diminue sensiblement les risques d'erreur.

L'unité documentaire de recherche et de réponse retenue est en principe l'article pour les codes, lois et règlements, et la décision dans son ensemble pour la jurisprudence. Les recherches se faisant sur des documents enregistrés en langage naturel, tous les termes significatifs du document pourront donner lieu à une interrogation. Il n'y aura donc pas, lors de l'interrogation à recourir à une liste de mots-clés préétablis.

Avant leur enregistrement, les documents recensés font l'objet d'un certain traitement qui diffère selon qu'il s'agit de textes législatifs et réglementaires ou d'analyse de jurisprudence.

Avant leur enregistrement, les documents recensés font l'objet d'une préparation qui diffère selon qu'il s'agit de textes législatifs et réglementaires ou de décisions de jurisprudence.

La jurisprudence est saisie sous forme de résumés précédés, pour le Conseil d'Etat, des titres et d'une indexation libre, et pour la Cour de Cassation de titres principaux et de références. Pour chaque décision est rempli un cadre d'enregistrement dont les diverses rubriques vont contenir des éléments, extraits de la décision elle-même. Il s'agit d'une part de rubriques d'identification (pour la Cour de Cassation, on rencontre notamment l'indication de la chambre, la date, le numéro de pourvoi, la publication, les noms des demandeurs et défendeurs, la composition de la formation de jugement, etc.), et d'autre part de références à d'autres textes, et notamment ceux qui sont cités par la décision. Parmi cet ensemble de rubriques, certaines, qui sont de longueur fixe (zones structurées) peuvent être utilisées à l'aide de fonctions spéciales du programme de recherche, afin d'obtenir, automatiquement certaines sélections ou tris. Lors d'une recherche, on peut alors limiter les réponses aux documents postérieurs à une date déterminée, on les trier sur cette date, etc.

Parmi ces rubriques, on trouve, par exemple, pour les décisions de la Cour de Cassation, le numéro significatif d'identification, le code de publication, toujours en un caractère : A pour les arrêts publiés, B pour les arrêts non publiés, mais qui figurent au fichier de la Cour de Cassation, C pour les arrêts non publiés et non fichés, etc.

Un traitement différent va être appliqué à la législation. Chaque article de loi constituera une unité documentaire et donnera lieu à la réduction d'un bordereau. En plus de la création de rubriques d'identification et de références, certains traitements vont être appliqués au texte lui-même. En effet une des particularités offertes par le système DOCILIS (DOCUMENTS INTERROGATIONS LIBRES), mis au point par le CEDIJ, est la possibilité de lutter contre le silence et de réduire le bruit. Pour cela, on recourt entre autre à l'explicitation. Ce procédé consiste à réinjecter dans le texte — entre des signes spéciaux, pour les distinguer du texte original — des notions implicites du document (explicitation du pronom, du groupe de mots incomplet, des concepts de définition, délai, durée, explicitation des polysèmes : Action * action en justice *, action * action mobilière *).

Comme complément à l'explicitation on a procédé à l'exploitation par programme des tables des matières. Celles-ci en effet contiennent souvent des concepts plus généraux qui n'apparaissent pas au sein des documents et qui sont utiles lors de l'interrogation. Ex. IRPP.

Aux fins de lutte contre le silence on a établi aussi des listes de substituts (ou équivalents). Ces listes explorées automatiquement ou sur demande par la machine feront porter la recherche sur les mots substituables à ceux de la question. L'emploi de listes de substituts grammaticaux (forme grammaticale, forme des adjectifs, des verbes) de substituts syntaxiques (termes d'une même famille syntaxiques ex. agricole, agriculture, agriculteur) et de substituts conceptuels (liste de synonymes : ex. redevable et contribuable) permet d'alléger la formalisation des questions. Il faut noter que l'automatisation d'un thésaurus multibase (dictionnaire destiné à faciliter la recherche en informatique juridique et contenant pour chaque mot-clé de la nomenclature les expressions apparentées par synonymie, analogie, etc.) en cours de constitution

intégrera toutes ces listes. On a dressé aussi une liste des expressions bruyantes comme avoir lieu, tenir lieu, au lieu de. Les mots repérés dès l'entrée sont accolés fictivement dans l'expression et la recherche porte sur l'expression.

L'intervention de l'ordinateur implique une phase supplémentaire par rapport au traitement manuel de la documentation. Il s'agit de la saisie des données. Après corrections, les documents sont introduits en mémoire d'ordinateur et on peut ensuite passer à l'interrogation.

B. — L'interrogation est faite à partir d'un terminal IBM 3275, console de visualisation composée d'un clavier de machine à écrire et d'un écran sur lequel vient s'inscrire le dialogue. A cette console est connectée une imprimante IBM 3284, qui imprime les documents à la vitesse de 15 caractères/seconde. Le programme DPS d'IBM a été abandonné en 1974 au profit du programme STAIRS/VS d'IBM.

Le début du dialogue commence par une procédure qui demande au système de connecter le programme spécial de recherche, ensuite l'ordinateur entame le dialogue et des descripteurs sont tapés sur le clavier. En réponse le système donnera des informations sur le nombre de document correspondant aux descripteurs demandés.

Le système conversationnel employé propose à l'utilisateur une formulation des questions lui permettant d'affiner sa recherche en fonction des résultats successifs obtenus.

La recherche documentaire sur le texte (SEARCH) est conduite à l'aide de trois éléments :

1. La possibilité d'utiliser des restricteurs de distance à côté des opérateurs classiques (opérateurs Booléens ET, OU, SAUF).

On pourra demander à ce que deux notions soient dans le même paragraphe (opérateur PARG), à ce que deux notions soient dans la même phrase (opérateur PHR) ou soient adjacentes (opérateur ADJ).

Ex. : VALEUR ADJ AJOUTEE.

2. L'utilisation du masque (la troncature à droite) permet de retrouver à partir d'une racine commune les différentes formes dérivées d'un mot et donc d'étendre la recherche.

Le masque peut être illimité, c'est-à-dire porter sur un nombre indéterminé de caractères derrière le radical.

La technique du masque doit être utilisée avec prudence, car si la racine est trop étroite, il y a risque d'introduire un certain bruit au niveau des réponses.

Ex. : Secret\$

donnera les documents contenant le mot secrétaire.

Il y a possibilité de limiter le masque à quelques caractères après la racine.

Secret\$1

donnera secret et secrets.

3. La possibilité d'interroger sur des zones déterminées. Cela permet de faire porter la recherche sur des parties déterminées de l'information globale en utilisant le procédé de la qualification, lorsqu'on est en mode Search. Par exemple, si on recherche des documents relatifs à des problèmes de sécurité sociale, il peut être intéressant de n'interroger que sur le texte, de façon à éliminer tous les bruits causés par sécurité sociale, en tant que partie.

Il est indispensable d'avoir une bonne connaissance du vocabulaire d'une base pour pouvoir interroger. Le CEDIJ élabore actuellement un thésaurus qui permettra de résoudre les problèmes de vocabulaire.

Les différentes fonctions accessibles avec STAIRS sont les suivantes.

SEARCH : recherche sur tout le document à l'exclusion des zones structurées, à l'aide de mots-clés reliés par des opérateurs booléens (ET, OU, SAUF), des restricteurs de distance (PARG, PHR, ADJ) et des extenseurs (troncature, appel aux listes). Comme nous venons de le voir la recherche peut être limitée à certaines classes ou sections par le procédé de la qualification.

SELECT : recherche sur les zones structurées. Recherches permettant par exemple de sélectionner les décisions entre telles et telles dates, les décisions publiées ou non publiées, etc. Cela sert pour des statistiques ou pour affiner une recherche commencée en SEARCH.

RANK : recherche sur la base de données à l'aide de mots-clés, mais avec un classement en fonction d'un algorithme choisi par l'utilisateur. Cette possibilité ne semble pas présenter d'intérêt pour la documentation juridique.

Il est possible de conserver les formalisations des questions (SAVE), puis de les reposer sans avoir à les refrapper sur n'importe quelle base de données (EXEC). Cette fonction peut servir en particulier pour constituer des profils et assurer ensuite une diffusion sélective de l'information.

Une fonction de tri (SORT) permet un classement des documents pour la visualisation ou l'impression suivant un ordre croissant ou décroissant de numéros d'articles, de date de jugement, etc. La fonction BROWSE permet la visualisation d'une partie ou de l'ensemble du document.

Il faut noter que si, au niveau de la recherche, il est possible d'interroger sur une section par le procédé de la qualification, la visualisation et l'impression se font au niveau de la classe (ex. : identification, références). Seules les zones structurées peuvent être visualisées ou imprimées au niveau de chaque zone.

II. — Il nous faut voir maintenant quelles sont les applications possibles de l'informatique juridique dans les bibliothèques.

A. — Un exemple de fonctionnement nous est donné par la bibliothèque du Conseil d'Etat où un terminal a été implanté par le CEDIJ et est effectivement ouvert à l'interrogation quatre après-midis par semaine depuis le mois de novembre 1977. Le CEDIJ y assure une permanence trois jours sur quatre et a formé une deuxième personne, du Conseil d'Etat, à la manipulation du terminal.

La moyenne du nombre des questions peut être estimée aux alentours de 3 ou 4 par séance d'interrogation ; les bases consultées sont essentiellement celles du Conseil d'Etat (la jurisprudence générale étant plus interrogée que la jurisprudence fiscale).

Les autres bases sont plus rarement demandées. Par ordre d'importance, citons : la jurisprudence du Conseil Constitutionnel, de la Cour de Cassation, et la législation fiscale.

Malgré le contenu encore limité de certaines bases de jurisprudence et la nécessité de faire appel à une tierce personne, à des heures déterminées, pour poser les questions, le système suscite un intérêt certain de la part des utilisateurs.

La plupart des questions sont posées immédiatement en présence de l'intéressé. Il y a très peu de recherches en différé. Cette pratique, qui peut restreindre le nombre des questions, facilite néanmoins la sélection des documents pertinents et un changement éventuel de stratégie.

Les membres du Conseil qui ont utilisé le terminal se sont montrés très intéressés par les possibilités de la recherche juridique automatique et paraissent généralement satisfaits des résultats obtenus puisqu'ils reviennent interroger. Un groupe d'habitues est ainsi en voie de formation notamment parmi les auditeurs.

Actuellement, la plupart des utilisateurs se servent du terminal pour compléter ou vérifier une recherche manuelle, mais une bonne connaissance de tous les avantages du système permet de s'en servir de manière plus autonome.

B. — Cette expérience montre les intérêts que peut présenter pour l'utilisateur d'une bibliothèque juridique, l'utilisation de l'information juridique : la possibilité de dialoguer avec l'ordinateur au cours d'une recherche permettant de préciser une question ambiguë en ajoutant d'autres mots-clés et d'autres relations, et la possibilité d'avoir un accès direct au document si on emploie un terme suffisamment spécifique.

Prenons pour exemple la notion de taxe sur les ordures ménagères. Aucune rubrique n'existe dans le plan de classement du Conseil d'Etat. On obtient par contre un accès direct aux documents, par l'ordinateur, en introduisant la notion d'ordures ménagères.

Le bibliothécaire, quant à lui, tirera grand profit de l'utilisation de l'informatique juridique puisqu'il pourra procéder facilement à la diffusion sélective de l'information et qu'il pourra obtenir des sous-produits tels que la constitution automatique des tables d'index (KWIC-KWOC).

Mlle Combescur.