

«IDT 85»

Le sixième congrès national sur l'information et la documentation s'est tenu à Versailles, sous l'intitulé «IDT 85 - Information - Documentation - Transfert des connaissances» du 12 au 14 juin 1985. Il était organisé par l'ADBS (Association française des Documentalistes et des Bibliothécaires spécialisés) et l'ANRT (Association nationale de la Recherche technique) avec le concours de la MIDIST (Mission interministérielle de l'Information scientifique et technique). Plus de 700 participants français et étrangers assistaient à ce congrès où 60 orateurs environ ont fait le point sur l'évolution rapide des nouvelles technologies de l'information (informatique, télématique) et sur les perfectionnements nouveaux des matériels.

Ces nouvelles technologies ont fortement modifié les données de la profession de documentaliste dont les fonctions et les qualifications sont souvent sous-évaluées. Une troisième enquête nationale effectuée par l'ADBS a permis de mieux connaître la situation d'ensemble de la profession de documentaliste désigné souvent maintenant comme agent de transfert. Nous y constatons, d'après les premiers résultats, une accentuation de la féminisation, un développement de la documentation en province, et un éventail très large des salaires.

Les réponses à un questionnaire adressé à un échantillon de 500 documentalistes italiens ont permis pour la première fois, d'identifier trois types de profils : celui du documentaliste (âge, sexe, profession), celui de l'organisation dont il fait partie (administration publique ou entreprise privée, secteur économique concerné), celui des fonctions qu'il exerce (types d'informations, types d'utilisateurs).

Les fonctions documentaires : acquisition de docu-

Geneviève HARTMANN*

ments, rédaction de résumés ou de synthèses bibliographiques, indexation, diffusion, stockage des documents, réponses aux questions, abonnements à des profils, saisie des données, achat de microordinateurs ou terminaux, sont en grande partie une activité de production. Il faut donc gérer la documentation comme on gère toute entreprise, c'est-à-dire en y introduisant les notions d'innovation, de productivité, d'analyse de la valeur, et surtout de dimension économique. La profession ne peut en tirer que des avantages.

Chacun de nous doit pouvoir accéder, quelle que soit sa localisation géographique, à l'information dont il a besoin. C'est ainsi que dans les milieux régionaux, au moment où se développe le mouvement de décentralisation, se sont créés des réseaux locaux départementaux. Nous citerons : le projet de Club d'utilisateurs de banques de données pour les collectivités locales mis en route par les grandes villes de Rhône-Alpes, le système ORIADOC pour recenser le potentiel d'informations dans les institutions et associations, l'élaboration des plans régionaux et des contrats de plan Etat-Région, la création de banques de données municipales, de systèmes locaux d'information (en Aquitaine, notamment), l'action des ARIST (Agence régionale d'Information scientifique et technique). Tout ceci a exigé le concours des professionnels de la documentation ou de formateurs conduits par l'ADBS, dont les enseignements ont été adaptés aux changements administratifs et technologiques. En effet, pour voir les régions acquérir les moyens de collecte et de diffusion de l'information économique et sociale (IES), la formation de spécialistes s'imposait.

Outre les régions de France, les pays en voie de développement ont besoin, eux aussi, de l'information scientifique, technique, économique et sociale. L'UNESCO les aide à traiter et à diffuser les documents utiles. Citons deux organismes parmi d'autres : AGRIS, Système international d'information pour les sciences et la technologie agricoles qui dépend de l'Organisation des Nations Unies pour l'alimentation et l'agriculture (FAO) et PADIS, Système pana-

* - CIRAD - 42, rue Scheffer - 75116 PARIS.

fricain de documentation et d'information de la Commission économique pour l'Afrique.

Le Centre pour le Développement industriel (CDI) créé en 1977 dans le cadre de la Convention de Lomé a pour but de favoriser la création d'entreprises industrielles et le transfert de technologie dans 65 pays ACP (Afrique - Caraïbes - Pacifique). Les demandeurs de données technologiques, économiques et statistiques disponibles à l'échelon international peuvent les recevoir et en effectuer le choix.

La société de l'information tient de plus en plus une place croissante dans les pays en développement, mais peu ont, jusqu'à présent, élaboré une politique nationale d'information. Le programme NATIS de l'UNESCO, vers les années 1970, consistant à développer des infrastructures d'archives, bibliothèques et services de documentation, n'a pas eu beaucoup d'efficacité. Le remède est que, pour être effective, une politique nationale d'information doit être formulée conjointement par les diverses parties prenantes, différents organismes spécifiques regroupant les décideurs du plus haut niveau. Un effort d'organisation de l'IST a été entrepris par les pays du Sahel.

L'Institut du Sahel, créé par le CILSS (Comité permanent inter-Etats de Lutte contre la Sécheresse) a pour but depuis 1973 de collecter, analyser et diffuser les résultats de la recherche. En décembre 1976, il fut décidé la mise en place d'un réseau régional d'information et de documentation scientifiques et techniques, connu sous le nom de RESADOC, dont les activités ont débuté en mars 1979. Ce réseau couvre les informations relatives à l'agriculture, élevage, pêches et forêts, ainsi que les informations socio-économiques et démographiques.

En mars 1985, RESADOC se composait de 30 centres sectoriels dans cinq pays et 20 centres régionaux. Les systèmes informatiques sont PCISIS (Bamako) et MINISIS (Dakar); RESINDEX, un index général, paraît deux fois par an.

RESADOC utilise le Macrothésaurus de l'OCDE et AGROVOC celui de la FAO, enrichis de listes complémentaires comprenant les termes spécifiques à la région et des descripteurs ethniques, géographiques, et institutionnels. Un projet avec le CIRAD (Centre de Coopération internationale en Recherche agronomique pour le Développement) et l'ORSTOM est en cours d'étude grâce à l'appui du Ministère des Relations extérieures. Système décentralisé, RESADOC, dont l'évolution est positive, a besoin du concours des pays du Sahel eux-mêmes pour arriver à la mise en place d'un système régional cohérent.

Pour que l'information diffusée soit efficace et atteigne le but qu'elle s'est assigné, ne doit-elle pas être bien comprise, bien interprétée par chacun ? Elle ne peut l'être que si elle est transmise d'une façon intelligible à l'utili-

sateur, c'est-à-dire dans sa propre langue. Nous abordons maintenant les problèmes linguistiques. La traduction humaine classique est souvent lente et coûteuse. Depuis 1978, la Commission des Communautés européennes a développé le système SYSTRAN d'origine nord-américaine avec post-édition et syntaxe libre. TITUS IV, par contre, utilisé par l'Institut textile de France et le CIRAD, est un système à syntaxe contrôlée. Les résumés sont entrés de façon interactive, chaque phrase devant être validée avant de passer à la suivante. Les résumés doivent contenir des termes figurant obligatoirement dans le lexique préalablement établi (français, anglais, espagnol, allemand). Des grammaires génératives transforment chaque phrase en langage pivot très condensé pour les stocker. Chaque document est traité ensuite par les grammaires transformationnelles qui en assurent la traduction, la sortie et l'édition.

ITS (Interactive Translation System) est un système de traduction automatique de seconde génération. Le problème est de résoudre les ambiguïtés des textes destinés au grand public ; on utilise l'interaction et les interlocuteurs humains doivent être des spécialistes et avoir une certaine connaissance du domaine.

EPISTLE est considéré comme un traitement de texte dont le traitement linguistique comprend une partie lexicale, grammaticale et stylistique. Les problèmes de traduction sont résolus d'une manière interactive avec les auteurs mêmes des textes à traduire et non avec les linguistes.

La langue française a retenu l'attention des spécialistes de la communication. Voilà pourquoi les chercheurs des Universités de Grenoble, Lyon et Fribourg (Suisse), formant le groupe SYDO, essaient de décrire d'une manière exhaustive l'organisation morphosyntaxique d'un texte en langue naturelle. De son côté, l'Institut national de la langue française à Nancy, laboratoire du CNRS, a un programme précis portant sur la langue et la linguistique françaises : traitement de bases de données textuelles, lexicologiques et bibliographiques. C'est ainsi qu'a été constituée, entre autres, une textothèque de plus de 1800 textes littéraires du XVIIe au XXe siècle.

Le travail informatique consiste à faire le relevé et le comptage de tous les « mots pleins » d'une oeuvre, de mettre en évidence les occurrences ou les cooccurrences, c'est-à-dire les apparitions d'une unité linguistique dans le discours, de signaler des termes placés entre guillemets ou mis en italique, d'analyser enfin l'environnement syntaxique d'un mot.

Si cette exploration informatique peut contribuer à mieux connaître l'histoire des faits, institutions et idées au cours des siècles, l'intelligence artificielle étudiée il y a plus de 30 ans aux Etats-Unis, particulièrement au MIT (The Artificial Intelligence Laboratory Massachusetts Institute of Technology) a, elle, un tout autre but. Son

objectif est de comprendre la nature de l'intelligence en construisant des programmes d'ordinateur imitant l'intelligence humaine, capables d'accomplir des tâches réservées exclusivement à l'intelligence humaine.

L'intelligence artificielle s'intéresse aux problèmes cognitifs mis en oeuvre par l'être humain lors de l'accomplissement des tâches intelligentes. Celles-ci sont très variées ; elles peuvent s'appliquer à résoudre un problème de mathématiques, à jouer au bridge ou aux échecs, à établir un diagnostic médical. Les caractéristiques des programmes d'intelligence artificielle sont la manipulation de symboles autres que numériques, le traitement des problèmes dont on ne connaît pas l'algorithme, l'utilisation de méthodes heuristiques, et la représentation des connaissances, c'est-à-dire une correspondance entre le monde extérieur et un système symbolique permettant de raisonner.

Un des domaines de l'intelligence artificielle est en ce moment celui des systèmes experts, très fertile en applications. La construction d'un système expert nécessite de nombreux entretiens avec un expert du domaine. Ce dernier peut expliquer ses connaissances de façon désordonnée, imprécise ; celles-ci doivent donc être soumises aux critiques d'autres experts. On obtient ainsi une base des connaissances, non figée mais demandant une mise au point vigilante et une mise à jour régulière. A côté de la base des connaissances, la base des faits regroupe l'ensemble des informations propres au problème précis que l'on désire résoudre et à partir duquel le système expert pourra appliquer ses connaissances et ses mécanismes de raisonnement.

La partie intelligente du système expert qui permet de tirer à partir de la base des faits, des conclusions et décisions appropriées, est le moteur d'inférences. Ce dernier est un mécanisme qui effectue le calcul des propositions logiques.

Les systèmes experts sont particulièrement utilisés en biologie et en médecine. Nous citerons, parmi beaucoup d'autres : MYCIN, diagnostic des infections bactériennes développé à l'Université de Stanford (1977) ; PIP (Present Illness Program), diagnostic des problèmes rénaux ; CASNET (Causal Association Network), diagnostic de glaucomes ; INTERNIST/CADUCEUS, développé par l'Université de Pittsburgh, comprenant des connaissances sur environ 350 diagnostics de la médecine interne.

Parmi les systèmes experts en sciences physiques et chimiques, citons : DENDRAL (1971), qui propose des structures chimiques possibles pour des molécules dont le spectre de masse est connu ; PROSPECTOR, spécialisé dans la prospection minière et la recherche pétrolière. Enfin dans le domaine agronomique, le système TOM, de l'INRA, renferme les connaissances, le mode de raisonnement et les expériences des spécialistes de la tomate. Sont

étudiées 60 maladies, dont 23 mycoses, 5 bactérioses, 6 viroses, 2 mycoplasmoses, des maladies non parasitaires, carences, anomalies génétiques, et ravageurs. Le moteur d'inférences fonctionne suivant un algorithme de filtrage qui examine si les prémisses des règles qu'il possède sont vérifiées par des faits contenus dans la base des faits ou par les réponses aux questions qu'il pose chaque fois qu'il lui manque des informations pour poursuivre son raisonnement. TOM a l'initiative du dialogue avec l'utilisateur qui lui répond («oui», «non», ou réponse numérique).

On peut prévoir la prolifération de nombreux systèmes experts dans les domaines suivants : banque, assurances, droit. Ils sont susceptibles de remplacer les programmes traditionnels (algorithmiques) qui sont souvent très difficile à mettre à jour et à améliorer. La recherche va donc continuer, visant à des systèmes experts encore beaucoup plus performants.

Il est possible de réaliser des systèmes experts sur microordinateurs, mais l'utilisation d'un langage approprié comme LISP, PROLOG, s'impose.

LISP, créé dans les années 1960, permet de traiter des listes de symboles. Des versions de LISP sont disponibles sur la majorité des microordinateurs : Apple 2, IBM PC, HP 150.

PROLOG, créé en 1970, permet de résoudre des problèmes d'objets et de relations entre ces objets. C'est un langage conversationnel ; on peut tenir une espèce de conversation avec l'ordinateur.

PROLOG attend qu'on lui donne des règles et des faits en rapport avec le problème que l'on veut résoudre. C'est un langage non procédural, c'est-à-dire que chaque partie du logiciel est indépendante des autres. Disponible sur de nombreux microordinateurs, il est surtout un élément essentiel du programme de cinquième génération au Japon.

La micro-informatique, dont la qualité la plus remarquable est la facilité de mise en oeuvre, voit son domaine s'étendre de plus en plus. Une enquête réalisée par le groupe MICRO-DOC de l'ADBS a donné un aperçu de l'utilisation des logiciels documentaires sur microordinateurs en France. 35 réponses ont été exploitées. Citons notamment les logiciels MICRO-QUESTEL de TELESYSTEMES conçu par Tamil, fonctionnant sur Micromega 32 de Thompson, BIBLIO 8000, GESTDOC, TEXTO (Chemdata) utilisé entre autres par le CIRAD et implanté sur OLIVETTI M 24. Les caractéristiques essentielles de TEXTO sont : le mode conversationnel pour toutes les fonctions, la maîtrise des applications, la possibilité de modifier en permanence le contenu ou la structure des fichiers, l'obtention d'éditions sophistiquées définies par l'utilisateur, d'index et de sous-fichiers, et la possibilité d'avoir des interrogations complexes. VALSE, créé en 1979

au CNRS, fonctionne sur APPLE 2, IBM-PC, avec disquettes ou disques durs, système d'exploitation (CP/M ou MS/DOS) et langage (BASIC). MIKRODOC est utilisé par le centre de recherche SANDOZ et implanté sur MICRAL 8031 C du groupe BULL. L'inconvénient de ces micro-ordinateurs est l'impossibilité, pour la plupart, de gérer des thésaurus. Des études en cours sont faites pour y remédier.

Cette explosion de la micro-informatique qui s'amplifiera encore dans les prochaines années, est accompagnée de l'édition électronique qui, rassemblant de nouvelles technologies, apporte de nouveaux services et bouleverse les techniques documentaires.

La **production électronique** concerne des documents diffusés sous forme traditionnelle, mais avec utilisation de machines de traitement de texte pour informatisation de la saisie, et de petites imprimantes à laser.

La **diffusion électronique** concerne la télécopie, l'accès en ligne à des bases de données en texte intégral, l'obtention de disquettes, vidéodisques, disque optique numérique, disque compact.

La **publication électronique** rend accessible des informations, sans édition intermédiaire, sur support électronique. C'est le domaine des vidéotex (terminal, réseau de télécommunications, serveur) et de la messagerie électronique, c'est-à-dire la distribution instantanée de messages aux utilisateurs en laissant au destinataire le choix du moment et du lieu à partir duquel il prendra connaissance des messages qui lui sont destinés. Prenons comme exemple le Catalogue Collectif National des publications en série (CCN) dont la messagerie électronique permet une transmission instantanée des demandes de documents. Une expérience de télécopie permettra d'intervenir sur les délais de transmission du document lui-même.

La Communauté européenne a participé au financement d'une série d'expériences qui ont exploré plusieurs solutions possibles pour la réalisation de systèmes de fourniture électronique de documents. TRANSDOC, projet piloté par le CNRS, améliorera le service manuel existant de fourniture de documents pour la base de données bibliographiques INPI (Institut national de la Propriété industrielle), le fichier biomédical PASCAL et la base de données EDF pour la littérature grise. Le projet comparera l'utilisation du disque optique numérique (DON) et les microfiches comme moyen de stockage. Les documents seront commandés via TELESYSTEMES-QUESTEL, puis imprimés sur place par laser ou transmission par télécopie.

Commencée en février 1983, l'expérience TRANSDOC où l'utilisateur enchaîne recherche et commande en ligne, doit se terminer fin 1985.

L'édition électronique, malgré son mouvement ascendant,

ne pourra pas remplacer, faire mourir le « papier ». Celui-ci est bien vivant, et chacun de nous peut y trouver l'information dont il a besoin. On parle souvent de sur-information face à la multitude de documents, surtout de type technologique. Au Japon, c'est par millions que se mesure le nombre d'informations à caractère technologique qui circule en une année. Comment identifier le document pertinent, cohérent et valable ? La veille technologique, représentée en France par le CPE (Centre de Prospective et d'Évaluation - Ministère du Redéploiement industriel et du Commerce extérieur) aide les PMI et autres organismes à connaître ce qui se fait dans les domaines de la recherche, du développement, à mieux appréhender les mutations technologiques qui affectent le développement industriel. Elle offre les moyens d'avoir une vision cohérente du futur et combat la désinformation, la rétention de l'information scientifique, technique, et industrielle. Le CPE a créé un journal électronique VIDEOTECH-CPE, en collaboration avec Télésystèmes EURODIAL et Innovation 128. D'autre part, des initiatives prises par l'ACTIM (Agence pour la Coopération technique industrielle et économique) consistent à informer la presse étrangère sur les réalisations techniques françaises. C'est ainsi qu'une cellule de presse, installée à Paris, donne des informations techniques faisant l'objet de communiqués. Deux bureaux de presse à Francfort (RFA) et Toronto (Canada) traduisent les dossiers de l'ACTIM et d'autres organisations et les proposent aux rédactions de publications techniques locales.

La veille technologique consiste aussi à s'assurer de la valeur d'un document au cours d'une recherche pour la décision scientifique et industrielle. Un interrogatoire systématique de la base SCIENCE CITATION INDEX dans le domaine de la physique a permis d'apprécier l'intérêt et les limites de cette liste alphabétique d'auteurs.

Dans le domaine des brevets et de la littérature grise, les indexations par mots-clés font souvent défaut ou sont longues à réaliser. Ainsi le système mis en place par le SERPIA (Service d'Études et de Réalisation de Produits d'Information avancés) et le Centre de Sociologie de l'Innovation de l'École des Mines de Paris, avec les programmes LEXIMAPPE et la collaboration de TELESYSTEMES QUESTEL-PLUS, a permis d'étudier les associations de mots qui désignent les sujets de recherche dans la littérature scientifique et technique.

Ainsi, on peut détecter les centres d'intérêt concernant les domaines scientifiques et industriels ou évaluer les sujets de recherche de pointe.

La collecte et le traitement simultané des données techniques et économiques peuvent donner lieu à un véritable **choix technologique**. La démarche de ce choix paraît claire ; elle comporte le plus souvent deux étapes : la prévision des conséquences de l'innovation étudiée effectuée à l'aide de calculs, et l'évaluation de ces conséquences, qui peut entraîner un débat .

Ces deux données économiques nécessitent le recours à un dispositif statistique, qui donne une image résumée des productions industrielles et transactions commerciales. Ce dispositif statistique est donc complémentaire des réalisations en cours à partir de systèmes d'information scientifiques et techniques. Ces différents travaux apparaissent, selon un représentant de l'INSEE, comme les composantes d'un futur «Système d'information technologique», nécessaire à la maîtrise du développement.

Concept d'information, information ... comment se le représenter, l'interpréter, ce mot qui est apparu tant de fois dans ce texte ? N'apparaît-il pas ambigu, flou ? Est-il possible de connaître le monde de l'information comme le monde physique ? La science de l'information existe-t-elle ? Oui, si on considère qu'elle a su développer des processus de communication entre les hommes, inventer des méthodes d'analyse, qu'elle a découvert des lois (loi de Bradford, qui établit des relations universelles et nécessaires entre l'apparition d'un phénomène et les conditions qui le font apparaître, et permet ainsi de faire des prévisions).

Cependant, ses fondements restent mal assurés. Et le développement national de la recherche, le renforcement des équipes pour cette recherche, la fourniture d'évaluation et de prospective, sont les conditions indispensables de son existence.

En outre, on peut se poser beaucoup de questions sur cette société de l'information, successeur de la société industrielle. Son aspect en est profondément changé ; la notion de productivité a disparu. C'est la société de l'intelligence - le neurone a remplacé le muscle - de convivialité, de mondialisation. L'information est légère, circule; attirée par le mouvement, elle passe les frontières. Il y a uniformisation des échanges. Les marchés de l'information sont désormais mondiaux, et l'information ne peut plus être considérée comme un bien collectif («individualisable»), mais comme un bien individuel, divisible. Ne voit-on pas déjà l'isolement de l'individu chez lui devant la télévision, à l'entreprise ou chez lui encore devant le clavier de son terminal, captant son propre domaine spécifique ? L'information peut être source de distraction, génératrice de passivité. Le message, parfois sans signification ou mal compris, ne passe plus, il n'est pas retenu. Ces propos inquiets, mais lucides sont ceux d'un universitaire.

Les conséquences sociologiques de la société de l'information, elles aussi, ont une certaine importance. Il est difficile de prévoir les demandes et d'anticiper sur la consommation des nouveaux services informatiques. Le facteur de stabilité est le savoir-faire, le «métier» du documentaliste ou de l'éditeur. Dans la mise en réseaux, on

peut constater la priorité du traitement sur le stockage. Ce traitement peut être le fait de certains «dirigeants» il y a donc une menace «d'uniformisation», de standardisation ; une certaine hiérarchisation de l'information peut aussi apparaître.

Sur le corps social, l'informatisation de la société peut avoir plusieurs conséquences :

- les contraintes du temps et de l'espace surtout, disparaissent ; la distance devient sans conséquence, mais cet effet de proximité extraordinaire crée un phénomène de perplexité. Subsistent seulement les contraintes ergonomiques dues aux horaires ;

- l'abondance d'informations utiles étouffe le silence, si essentiel pour la réflexion, et qui ne préoccupe plus ;

- un risque apparaît : le processus de la recherche «à tâtons», aléatoire, disparaît, d'où perte de capacité de conceptualisation, de créativité ;

- parallèlement au développement des grands réseaux, se crée une forme de communication «sans information», à la fois fascinante et dégradante; une espèce de solitude se crée ;

- parallèlement à l'internationalisation, il y a interpénétration de l'espace public et privé. L'informatique n'est pas la conversation.

On se demande alors quel est «l'homme de l'information» ?, celui qui détermine la société d'aujourd'hui. Comment se présente-t-il ? On a de la peine à l'imaginer. Cet homme va-t-il se laisser dominer par l'afflux de machines et robots ? L'informatique (le programme ESPRIT, programme européen sur les technologies de l'information a été lancé en 1984) peut être un risque, un danger.

À côté de cela, en incitant à de nouvelles recherches, en permettant des progrès techniques impensables auparavant, cette informatique demandera encore plus d'intelligence et de compétence pour l'utiliser, et là nous pensons particulièrement aux banques de données et à la quantité de réflexion qu'il faut pour les interroger correctement, à la nécessité de bien connaître son «sujet»!

L'être humain ne doit pas devenir assistant de la machine, il faut qu'il soit au niveau de l'ordinateur. La machine appelle l'homme à grandir, «l'Homme qui, selon Pierre Teilhard de Chardin, dans la Nature, est véritablement un fait, relevant (au moins partiellement) des exigences et méthodes de la Science ... parmi les faits présentés à notre connaissance, nul n'est plus extraordinaire, ni plus illuminant ...». (Prologue du Phénomène Humain, Pékin 1938-1940).

