

Production, diffusion et exploitation des banques de données en ligne constituent aujourd'hui une véritable filière industrielle, en pleine évolution pour s'adapter aux nouveautés imposées par Internet.



Les banques de données commerciales (ASCII, Vidéotex)

Quels en sont les principaux acteurs ?

■ **les auteurs** : Ils sont à l'origine de l'information dite « primaire » par la publication d'articles, d'ouvrages... dans le cadre de leurs travaux de recherche ou d'enseignement.

■ **les producteurs** : Les organismes qui, par un travail systématique de collecte et de sélection des informations, fournissent le contenu des banques de données.

- Les centres de documentation des grands pôles de recherche scientifique sont tout naturellement devenus les premiers producteurs. En France, le CNRS (banques de données Pascal, Francis), le CEA ou la NASA et Chemical Abstracts Service aux Etats-Unis qui produisaient des bulletins signalétiques imprimés, se sont tournés vers la télématique comme moyen complémentaire de diffusion.

- Rapidement, entreprises, associations professionnelles, organismes publics... sont devenus des producteurs au même titre que la communauté scientifique comme c'est le

cas actuellement sur Internet.

■ **les diffuseurs** : Ils sont chargés par les producteurs de la structuration des banques de données et du développement des outils qui permettent de les interroger.

Pour les banques de données accessibles en ligne, le rôle de diffuseur est assuré par les centres serveurs. Ces derniers soumettent les données fournies par les producteurs à des traitements informatiques qui les rendent interrogeables à distance : création d'index, harmonisation des formats de l'ensemble des banques diffusées.

- Les serveurs gèrent également la commercialisation et la promotion des banques de données inscrites à leur catalogue.

- Parmi les grands serveurs ASCII diffusant au niveau mondial, Dialog-Datastar, devenu Dialog Corporation à la suite de son rachat par la société anglaise M.A.I.D. (celle-ci offre, à travers son service Profound, l'accès via Internet à un ensemble de sources en



texte intégral dans le domaine des affaires), propose un catalogue de plus de 800 banques de données qui couvrent tous les champs de la connaissance. On peut citer également Lexis-Nexis, serveur spécialisé en information juridique, presse et affaires, qui compte le plus grand nombre d'abonnés (plus de 800 000), OCLC avec son service Firstsearch, STN plus tourné vers l'information scientifique et technique.

- En France, le serveur le plus ancien Questel.Orbit, donne accès à 250 banques de données qui se répartissent en trois grandes lignes thématiques : la propriété intellectuelle et industrielle, les affaires et l'actualité, l'information scientifique et technique. Il compte 35 000 abonnés. L'Européenne de données, filiale du groupe OR Télématique offre quant à elle un catalogue de banques de données dans le domaine de l'information juridique, de l'actualité générale et de l'économie.

■ **les transporteurs** : Il s'agit des opérateurs de réseaux de télécommunications qui assurent l'acheminement des données d'un point à l'autre du globe.

Ces réseaux fonctionnent selon des normes d'échange communes (transmission de données par paquets) suivant une architecture complexe d'interconnexions internationales.

- Dans de nombreux pays européens, l'Etat possédait encore le monopole de la transmission de données par réseau spécialisé : par exemple, France Télécom a créé le réseau Transpac qui permet d'accéder aux banques de données sur le territoire national mais aussi aux serveurs mondiaux, par l'intermédiaire d'un noeud de transit international.

- Dans d'autres pays, les USA en particulier, l'infrastructure des télécommunications est gérée par des sociétés privées et concurrentes (réseaux Tymnet, Telenet, Uninet...).

En 1998, l'ouverture du grand marché européen a mis fin aux situations de monopole existant encore en Europe et va généraliser le système concurrentiel.

■ **les intermédiaires** : On appelle ainsi les organismes ou personnes qui jouent le rôle de courtiers (« brokers »). Ils effectuent contre rémunération des recherches ponctuelles pour le compte de clients réguliers ou occasionnels. Ces structures sont privées ou publiques.

Les bibliothèques universitaires peuvent être rangées dans cette catégorie. Elles disposent de services de recherche documentaire en ligne (RDI) pour interroger les principaux serveurs en complément de services d'interrogation directe de banques de données sur cédéroms ou via Internet qu'elles offrent aux utilisateurs.

■ **les formateurs** : L'utilisation efficace des banques de données requiert une formation pratique et méthodologique particulière.

Au sein des universités, les URFIST (Unités régionales de formation à l'information scientifique et technique) ont pour mission d'assurer la formation des professionnels de l'information, des étudiants, des enseignants et des chercheurs à l'usage des nouvelles technologies d'information (cf infra p. 62 présentation des Urfist). Les bibliothèques universitaires assurent également ce rôle pour leurs utilisateurs.

gement dominé par les acteurs anglo-saxons.

- Les stratégies en particulier au niveau des serveurs s'illustrent par des prises de contrôle ou des fusions, destinées à créer des pôles de diffusion de taille internationale. L'achat d'Orbit par Questel, il y a quelques années, entraine dans cette logique de même que celui de Dialog-Datastar par M.A.I.D plus récemment.

On assiste également à de nombreuses acquisitions dans le domaine de l'édition dans la perspective du développement de l'édition électronique : les éditeurs néerlandais comme Elsevier ou Wolters-Kluwer se montrent particulièrement actifs dans ce domaine.

- Autres évolutions notables ces dernières années, les serveurs déclinent leur offre pour toucher un public plus large, notamment l'utilisateur final, avec des produits de plus en plus ciblés et personnalisés.

Cela s'est traduit dans un premier temps, en France, par une offre parallèle d'accès vidéotex guidés. De nombreux utilisateurs (recherche, entreprises) recourent à ce mode d'accès simplifié pour leurs besoins ponctuels.

La plupart des services professionnels sont diffusés sur les hauts paliers : 08 36 28, 08 36 29 et 36 16, 36 17. Vous en trouverez quelques exemples à la fin de cette brochure p. 65 et 66.

Autre stratégie couramment adoptée : l'offre de logiciels conviviaux qui intègrent des commandes d'interrogation guidées avec des fonctions d'édition... facilitant ainsi la réalisation de véritables dossiers documentaires. Questel, par exemple, propose le logiciel Imagination 2 (cf p 20 et 21).

- Enfin, aujourd'hui, les principaux développements se font autour d'offres Internet-Intranet pour permettre aux utilisateurs d'interroger les ressources via une interface Web en exploitant les potentialités offertes par l'hypertexte. Dialog Datastar offre Dialog Web et Datastar Web tandis que des acteurs plus récents comme Ovid (banques de données sur cédérom et via Internet dans le domaine médical) connaissent un certain succès du fait de leur positionnement rapide sur ce type d'offre.

De nouveaux acteurs sont en cours d'apparition sur ce créneau : ce phénomène est particulièrement sensible dans les domaines des affaires et de la médecine.

- Parallèlement, l'accès direct au document en texte intégral dont l'importance n'a cessé de s'accroître depuis la fin des années 1980 pousse les serveurs commerciaux à améliorer constamment les logiciels de recherche de l'information. Amélioration de la recherche en **texte intégral** et possibilités d'interroger en **langage naturel** sont deux enjeux importants.

Les coûts d'utilisation

Voici les différents éléments qui interviennent dans le calcul du coût d'interrogation :

- **le producteur** : il est en général rémunéré par des redevances versées par les serveurs. L'utilisateur n'a donc pas en principe à le payer directement.

- **le serveur** : L'utilisateur est lié au serveur par un contrat. La facturation des consommations est, le plus souvent, mensuelle. Les

structures tarifaires et le niveau de tarification sont actuellement en pleine évolution. Cependant, on retrouve généralement, les principaux postes de coûts suivants :

- dans certains cas, **des frais d'abonnement** ou un **droit d'entrée** annuels.

- un **coût horaire** de consultation très variable selon les types de données et les domaines concernés. Actuellement, ces coûts horaires baissent régulièrement voire

L'évolution de l'offre

L'industrie de l'information en ligne a connu au cours de ces 10 dernières années de profonds bouleversements, aussi bien sur le plan économique que sur celui de la technologie.

La production comme la diffusion des banques de données se révèlent aujourd'hui - surtout dans le domaine économique et financier - un secteur très concurrentiel, lar-

disparaissent au profit du paiement du document seulement.

- un coût **à la référence**, à la page de texte intégral ou au document, variable selon les domaines et selon différents paramètres comme le format de visualisation choisi ou les opérations effectuées (consultation, impression, téléchargement).

De nouvelles formules de tarification accompagnent l'adaptation des services à une consultation par Internet :

- paiement à l'acte
- forfaits déclinés selon différents modèles (de l'abonnement mensuel correspondant à un quota d'heures et/ou de consultations et/ou de documents, aux formules clubs, aux licences de sites négociées par les bibliothèques, par exemple).
- formules mixant les deux précédentes.

La facture intègre également (pour les serveurs ASCII) les frais de télécommunications.

■ **Les réseaux** : les coûts de télécommunications comportent plusieurs éléments :

- Les frais de téléphone pour accéder au réseau Transpac (facturés directement par France Télécom).
- les frais de télécommunications (réseau Transpac facturé par le serveur) calculés selon plusieurs paramètres dont le débit des liaisons.

L'accès par Internet constitue désormais une autre alternative (cf schéma technique p. 6 et 7).

Ces différents coûts s'additionnent.

■ D'autres prestations optionnelles peuvent également être facturées :

- **la commande du document** : le coût varie en fonction du fournisseur et du mode de transmission (poste, télécopie, courrier électronique). A titre d'exemple, l'INIST (Institut de l'information scientifique et technique du CNRS) propose à ses clients la transmission d'un document de dix pages en service courant (photocopies par voie postale), moyennant le débit de 29 unités sur un compte ouvert préalablement par l'utilisateur (coût d'une unité : 1,21 F TTC). Le même document transmis par fax dans les 2 heures entrainera un débit de 130,5 unités sur le compte.

- **la diffusion sélective de l'information (DSI et SDI en anglais)** : l'utilisateur peut faire établir un profil de recherche personnalisé sur une thématique précise définie par ses soins. Il pourra ainsi recevoir régulièrement le signalement des nouvelles références entrées dans une ou plusieurs banques de données. La tarification de ces prestations dépendra du nombre moyen de références obtenues lors de l'exécution automatique du profil de recherche.

Ce type de service s'est largement popularisé sur Internet : de nombreux serveurs proposent à l'utilisateur de lui envoyer à son adresse électronique les dernières nouveautés parues sur le serveur, les éditeurs proposent d'envoyer les sommaires de leurs journaux...(cf p. 35 et 36).

La tarification des services en mode vidéotex

Elle est toujours fondée sur la durée et varie selon les paliers Télétel et pour certains, selon la base de facturation choisie par le fournisseur de service. A titre d'exemple, le coût du 36 17 varie de 2,23 F TTC à 5,57 F TTC la minute, le coût du 08 36 28 s'élève à 5,57 F TTC la minute, le 08 36 29 à 9,21 F TTC la minute.

L'ensemble des coûts liés à la consultation des banques de données en mode vidéotex est directement imputé sur la facture téléphonique de l'utilisateur. France Télécom reverse ensuite une partie des sommes acquises aux producteurs, en fonction du palier de tarification dans lequel ils ont placé leur service.