

Assemblée générale Urbamet – 15 juin 2001

**Information électronique professionnelle :
tendances et innovation**

Michel VAJOU, Journaliste

Je vais commencer par me présenter. J'ai été journaliste, comme l'indique le programme. De 1982 à 1984, puis de 1996 à 2000, j'ai été responsable d'une revue qui s'appelait Infotecture, lettre d'information sur l'industrie des bases de données. Je n'ai pas été trotskiste, mais j'ai quand même été documentaliste encarté à l'ADBS de 1980 à 1982. Plus sérieusement, l'essentiel de mon parcours professionnel s'est déroulé en Italie où j'ai été pendant sept ans responsable des développements européens de l'entreprise Seat Pagine Gialle, l'éditeur italien des pages jaunes lié à l'opérateur de télécommunications. Comme tous les journalistes vieillissants, je suis devenu consultant individuel sous l'enseigne M. V. Etudes et Conseil ce qui me permet non seulement de travailler à proximité de ma thèière et de mon fauteuil mais aussi de rencontrer des gens intéressants comme aujourd'hui. Sans travailler dans le secteur de l'urbanisme, j'ai essayé de trouver quelques illustrations s'y rapportant. Ma vision sera cependant beaucoup plus générique. J'espère qu'elle vous intéressera.

Avant de développer ma présentation, je souhaite prendre trois précautions oratoires.

Premièrement, ce panorama de tendances aura sans doute un aspect très fragmenté. Or je crois que depuis deux ans et pour les cinq prochaines années, la tendance la plus importante porte sur la césure dans le monde de l'information électronique professionnelle. A mon avis, cette césure est aussi importante que l'apparition du time sharing dans les années 60-70. Elle n'a pas encore donné tous ces effets. C'est le message principal que je voulais vous délivrer.

La deuxième précaution porte sur la méthode. Je ne suis pas chroniqueur de mode, je ne vous parlerai donc pas des palettes de coloris pour la saison. Pour tenter de repérer les tendances sans se fier simplement à son instinct, il faut disposer d'un outil, un traitement de type data mining. Cet outil de type statistique de l'ensemble des communiqués de presse de l'industrie de l'IEP est apparu en novembre dernier : la première version de ce travail a en effet été présentée à Londres au OnLine Information 2000. Le communiqué de presse est l'écume de la communication. Pris isolément, il est peu informatif à moins de savoir le lire. Néanmoins, en tant que corpus, la communication de presse est un mode de communication idéal dans une problématique de veille. Pourquoi ? Parce qu'il traite de sujets suffisamment innovants pour se distinguer de ces camarades et en même temps, il n'est pas trop en avance sur les problématiques marché puisqu'il faut aussi être compris du marché. Ainsi, convenablement traités, ces corpus de communication de presse représentent un outil de travail très intéressant pour faire de la veille. J'en ai 200 ici. Il existe évidemment des biais méthodologiques : ces communiqués étant essentiellement issus du secteur marchand, il est beaucoup plus facile de pointer les tendances des opérateurs de ce secteur que du secteur associatif et coopératif tel qu'Urbamet.

La dernière précaution tient au fait que je ne suis pas spécialiste de l'urbanisme. J'ai essayé de trouver des illustrations concrètes dans les champs qui sont les vôtres. J'espère que vous amplifierez ces illustrations par vos expériences.

Cadrage sur l'IEP et sur les marchés

Evolution du marché français de l'IEP

J'ai réalisé pour le GEFI une étude de quantification et d'analyse des tendances du marché de l'industrie française de l'IEP. Cette étude est disponible in extenso sur le site Internet gouv.fr dans la rubrique rapports 2001. Pour 2001, le marché de l'IEP représenterait plus de 7,3 milliards de francs. C'est dès aujourd'hui le second vecteur d'information professionnelle, après la presse professionnelle. En 2005, sans verser dans l'hyperbole, ce marché atteindra 15,5 milliards de francs. Il sera le premier vecteur d'information des professionnels.

Au-delà de son importance quantitative très nette, ce vecteur a un réel poids qualitatif. L'information numérique se prête parfaitement à de multiples usages de veille, de prise directe sur les applicatifs et les logiciels métiers. L'IEP est non seulement au centre des transactions, notamment financières, mais aussi au centre des systèmes de modélisation de l'univers physique, que ce soit au travers de données météorologiques ou géographiques.

La ventilation du marché de l'IEP par secteurs

Trois segments représentent à eux seuls plus de 70 % du marché. Premièrement, l'information financière est très présente puisque toutes les transactions financières s'appuient en temps réel sur des données financières. Deuxièmement, l'information de solvabilité permet de se renseigner sur les capacités de paiement des clients. Troisièmement, l'information marketing établit toutes les données de relevé sur panels et les fichiers de base adresses.

Les dépêches d'actualité, principal consommable des sites portail Internet, sont également un segment important. Tous les autres segments de marché se situent entre 1 à 3 %.

Un constat s'impose : le marché de l'IEP privilégie actuellement une information de veille directe très opérationnelle et à horizon relativement court. C'est particulièrement vrai pour les marchés financiers et les marchés de l'information de solvabilité. Il y a peu d'outils de modélisation prédictive ; c'est la nouvelle frontière de l'IEP. Le second constat porte sur le degré de numérisation de l'information. L'information la plus numérisée se trouve dans les trois premiers secteurs évoqués, donc dans ceux où elle est directement en prise avec des logiciels qui augmentent la productivité des utilisateurs. A contrario, l'information est moins numérisée lorsqu'elle porte sur la culture générale. On parle alors d'information défilante. La presse électronique ne représente ainsi que 3 % du chiffre d'affaires de la presse professionnelle et spécialisée. Il faut faire une distinction fondamentale entre une "information carburant" qui alimente le moteur des systèmes de production et une information "ère du temps" qui permet de voir défiler l'actualité.

Tendances structurelles

Prédominance stratégique des grands éditeurs producteurs d'information et marginalisation des serveurs

Des salons comme le OnLine Information de Londres ou IDT en France illustrent cette prédominance. Il y a 5 ans, des agrégateurs appelés serveurs dominaient ces salons. Citons Dialogue, Questel-Orbit, PergamonSDC. Aujourd'hui, la structuration de cette industrie est uniquement organisée autour des éditeurs producteurs d'information qui ont complètement

renversé le rapport de forces. Les activités internationales d'IEP se répartissent entre quatre grands acteurs : RidElsevire, Wolters-Kluwer, Thomson Publishing et Reuters. Durant les cinq dernières années, les serveurs se sont marginalisés. Toutefois, face à la prédominance stratégique des grands éditeurs, les réseaux documentaires connaissent une revalorisation concomitante. Urbamet en est un très bel exemple. L'association américaine de bibliothèques et des centres de documentation spécialisés mène actuellement une guérilla farouche contre le rachat de l'éditeur Harcourt par RidElsevire. Ces bibliothécaires ont calculé qu'au terme de ce rachat, 50 % des budgets d'information tomberont dans l'escarcelle de RidElsevire. Cette problématique de prédominance stratégique des grands éditeurs a des implications considérables, y compris sur la structuration de la demande.

Prédominance des grands groupes anglo-néerlandais

Tous ces grands groupes sont européens. Dans le domaine de l'IEP, il n'existe pas de groupe américain puissant au niveau international. Il y a 10 ans, ces groupes étaient mixtes, leurs portefeuilles se composaient aussi bien d'activités B2B que B2C. Au début des années 90, pour faire face à la crise des recettes publicitaires, ils ont transféré toute leur activité sur le B2B, en particulier sur les bases de données, en un mot sur l'information récurrente qui est une activité économiquement peu cyclique. Ils ont systématiquement revendu le cœur historique de leur métier, les activités de presse grand public, comme l'a fait dernièrement le Néerlandais VNU.

Parallèlement, pour garder le même niveau de chiffre d'affaires, il fallait pratiquer des acquisitions. Les seules acquisitions adéquates étaient les acquisitions sur le marché américain, marché linguistiquement unifié de 350 millions d'utilisateurs potentiels. Les sociétés anglo-néerlandaises se sont ainsi recentrées sur le B2B et ont traversé l'Atlantique pour effectuer des rachats. Il n'y a pas eu de levée de boucliers aux Etats-Unis lorsque Wolters-Kluwer ou RidElsevire ont racheté un éditeur local, mais ce fut le cas lorsque le rachat a confiné à l'abus de position dominante, comme le rachat d'Harcourt par RidElsevire.

Les producteurs indépendants d'information démontrent leur vitalité

D'une part, le développement de la coopération documentaire en réseau répond à cette concentration croissante de l'offre. D'autre part, les grands groupes ne s'intéressent qu'aux marchés qu'ils peuvent servir sur une base globale aux problématiques similaires des deux côtés de l'Atlantique. Leurs actuelles opérations de cession le prouvent. Ils se désengagent financièrement des marchés de niche pour lesquels la mise en œuvre de problématiques transversales est impossible. Dans leur optique, le marché de l'urbanisme est sûrement un marché de niche. Ce désinvestissement laisse de fortes opportunités aux acteurs de taille moyenne. Par exemple, les Français Consodata (information marketing), GLTrade et Fininfo (information financière) sont des challengers qui obtiennent de très bons résultats au plan européen et qui ont trouvé leur place aux côtés de géants comme Reuters.

Internet ne modifiera pas profondément la structure de l'offre

Il y a deux ans, à Londres, le thème était la " Déstabilisation des business models de l'IEP ". Il est vrai que, désormais, de nombreux services sont gratuits sur Internet. L'interrogation sous-jacente était la suivante : les start-up du Net vont-elles prendre la place des grands éditeurs, un peu comme AOL s'est emparé de Time Warner ? Aujourd'hui, il n'en est rien. Il n'y aura pas d'effet de restructuration du secteur lié à l'initiative des start-up de l'Internet. Ce sont au contraire les grands groupes qui sont les mieux placés, au travers des rachats, pour profiter des développements Internet passés par le canal des start-up.

Les start-up de contenus payants sont les seules à tirer leur épingle du jeu

L'étude dont j'ai parlé comportait pour la première fois cette année un volet sur les start-up. Après analyse des chiffres d'affaires, on s'aperçoit que les seules start-up qui réussissent sont celles qui produisent et vendent du contenu. Toutes celles qui ont adopté la gratuité ont peu de chance d'être bénéficiaires sur des projections à terme. Les start-up qui produisent des contenus payants réaliseront dès cette année plus de 1,2 milliard de francs de recettes. Ainsi, Consodata né en 1995 sur le créneau des bases de données comportementales va réaliser dès cette année plus de 100 millions d'euros de chiffre d'affaires. C'est considérable.

En comparaison, les sites de portail généralistes et spécialisés semblent moins prospères. Les sites de portail généralistes ne réaliseront probablement cette année pas plus de 183 millions de francs de chiffre d'affaires. C'est un paradoxe car derrière les sites de portail généralistes se trouvent souvent des opérateurs puissants comme Cegetel, BNP Paribas. Il est toutefois évident que ce modèle échoue à décoller. Les sites portail spécialisés se portent un peu mieux avec 440 millions de francs de recettes prévus pour cette année. Ce chiffre n'est lié qu'à la multiplicité de l'initiative. La France compte plus de cent portails professionnels Internet spécialisés. Lorsqu'on rapporte ces 440 millions au nombre d'initiatives, les chiffres d'affaires unitaires sont limités.

En termes de chiffre d'affaires, les agrégateurs de contenu s'en sortent bien car ils sont moins nombreux. Le total global s'élève à 164 millions de francs. Ces agrégateurs sont divers : Nfactory (start-up en cessation de paiement), PRLine ou Company News (services d'agrégation de communiqués de presse). Ils gagnent de l'argent en faisant payer les entreprises qui émettent les communiqués de presse et non pas ceux qui consultent l'information.

Le dernier secteur est celui des infologiciels, qui connaît de belles réussites en France. Les infologiciels regroupent tous les outils de visualisation, de traitement, d'indexation, de recherche, de filtrage de l'information.

La convergence information/ logiciels " métier " est une clé de structuration forte

L'objectif des grands groupes est d'avoir une continuité maximale entre la source d'information et le logiciel qui la traite dans une logique de productivité. Les grands producteurs d'information doivent prendre appui sur le segment des applicatifs métier pour être certains de s'enraciner durablement dans la station de travail de leur clientèle. On peut changer de fournisseur si l'offre est plurielle, mais il est difficile de changer de fournisseur de logiciel CAO. La stratégie transversale présente dans tous les domaines, aussi bien dans l'information juridique que dans les brevets, est l'intégration de l'information aux outils. La France a tendance à oublier les enjeux stratégiques qui en résultent. Je cite souvent l'exemple d'une fusion réalisée l'an dernier aux Etats-Unis. Le numéro un des bases de données sur les composants électroniques a racheté le numéro un des systèmes de CAO de circuits électroniques. Ainsi, la base de données sur les composants électroniques devient un input direct du système de CAO des composants. Le système de CAO n'utilisera plus que les composants référencés dans la base de données. Vous imaginez bien que cela pose d'importants problèmes de visibilité de marché. C'est un phénomène majeur pour les utilisateurs et pour la structuration de l'industrie dans son ensemble.

Faiblesse de l'offre d'infotlogiciels destinés aux utilisateurs

Les infotlogiciels sont en plein développement depuis 4 ans. Bientôt, il n'y aura plus d'information sans logiciels de traitement de l'information. Aujourd'hui, ce secteur est encore très atomisé, les grands mouvements de concentration ne font que s'amorcer. Les offres du secteur restent tournées vers les éditeurs, peu de produits sont destinés à la station de travail de l'utilisateur final. Les sociétés de taille moyenne n'ont en effet pas les moyens d'irriguer le marché capillaire de l'utilisateur final ; il est plus simple de vendre un logiciel d'édition sur Internet à une cible limitée d'éditeurs que de vendre un outil pour station de travail. L'évolution vers les infotlogiciels équipant l'utilisateur sera un grand chantier des années à venir.

Tendances produits et marché

Répartition des annonces par secteur d'appartenance

J'ai opéré cette répartition entre information juridique, entreprise, marketing, presse financière, scientifique et médicale (STM) à partir du traitement des 200 communiqués de presse.

Information STM, pôle d'innovation le plus actif

Depuis 2 ans, le secteur STM a retrouvé son rôle d'innovation des années 60. Son poids relatif est fort puisqu'il représente 31 % des annonces. Aujourd'hui, les outils auteur, les outils de workflow, les outils "collaboratifs", les bases de preprint, les bases de liens de navigation s'inventent dans le secteur STM. Ces inventions ont vocation à migrer dans l'ensemble des secteurs de marché. Tout ce qui est relatif à la veille économique et financière reste important : l'économie et la finance représentent plus de 26 % des annonces. Le reste est moins significatif.

STM : différenciation des positionnements

J'ai plaisir dans l'enceinte d'Urbamet à évoquer la revalorisation des rôles de référencement. Il y a 3 ans, on prédisait la fin des données bibliographiques au profit du texte intégral. Cette évolution ne s'est pas concrétisée. Aujourd'hui, il est essentiel d'avoir le contrôle des outils de référencement, en particulier des bases de liens dans le contexte de l'Internet. Elles sont le poste d'aiguillage permettant d'accéder à l'information et donc la porte d'entrée vers l'e-commerce de l'information. Des initiatives dynamiques apparaissent sur ce créneau de l'édition secondaire : l'Américain ISI (Information Scientific Institute) s'est constitué en 2 ans un fabuleux catalogue de bases de données bibliographiques par rachat ou par contrat d'externalisation à un moment où tout le monde se désintéressait des bases bibliographiques. Le Cambridge Scientific Abstract suit cette voie. Ces bases bibliographiques sont différentes des bases conçues il y a 15 ans : désormais, un lien se trouve derrière la base. La notice bibliographique est un sas d'entrée vers l'information primaire.

Les grands éditeurs primaires que sont RidElsevire, Springer Verlag et Kluwer sont en retrait par rapport à ces problématiques. Ils sont au centre des actions de contre-lobbying des grandes associations américaines d'utilisateurs parce qu'ils ont un peu trop augmenté les tarifs. Les utilisateurs sont de plus en plus inquiets de ces positions dominantes qui se créent à cause de la concentration du secteur. Les éditeurs primaires ont malgré tout de forts beaux atouts dans leur jeu. Si vous avez quelques économies à placer, achetez des actions RidElsevire ou Thomson, vous ne serez pas déçus.

De l'information de registre au " renseignement enquêté "

Le thème de l'information gratuite sur Internet est en train de provoquer la décantation des problématiques. L'IEP se divise en deux types si l'on exclut l'information d'analyse, de commentaire, journalistique.

Le premier type est l'information de registre : cette information résulte du dépôt d'information dans un *repository* pour raisons légales. C'est le cas de l'information brevet, de l'information financière.

Le second type est celui de l'information de renseignement enquêté : le producteur d'information prend la peine de définir un profil d'enquête, d'envoyer des questionnaires et de remonter par leur biais à une information propriétaire. J'insiste sur le mot " propriétaire " car il n'existe que ces deux types d'information et ils suivent actuellement des évolutions radicalement différentes. L'information de registre a vocation à être gratuitement accessible sur Internet. Il y a seulement 3 ans, l'information juridique, l'information sur les sociétés cotées en Bourse, l'information brevet, l'information registre du commerce étaient très chères dans les kiosques, elles sont désormais accessibles gratuitement sur Internet. Le seul refuge pour les éditeurs est de développer une information propriétaire c'est-à-dire une information créée sur la base de cadres propres aux éditeurs. Ces derniers ont bien sûr un peu de mal à gérer la transition, notamment les éditeurs juridiques qui sont fort marrés de l'apparition depuis 2 ans d'un service public d'accès au droit. Ils tiraient leurs bénéfices de la diffusion chère d'une information qui est fondamentalement une information de registre public.

Tendances par types de produits

Répartition des annonces par types de produits

Ce sont les portails spécialisés et non pas les portails généralistes qui arrivent au premier rang de ces annonces, simplement parce que le foisonnement d'initiatives se traduit dans les chiffres.

Les périodiques électroniques représentent la deuxième catégorie d'annonces, par ordre d'importance numérique. Ils portent sur ce qui a trait aux bibliothèques électroniques de périodiques, aux outils acteurs et aux bases de preprint.

La troisième catégorie concerne les bases de liens. La possibilité de naviguer est une valeur ajoutée essentielle de l'Internet. Pour naviguer, il faut créer des liens et les maintenir. La maintenance des liens est une opération coûteuse requérant beaucoup de minutie. Aujourd'hui, la constitution des bases de liens est un travail éditorial à part entière, elle s'effectue dans une logique consortiale. Je pense au consortium CrossRef qui associe les plus grands éditeurs internationaux de l'information HTML. Les bases de liens sont un service d'aiguillage ; chacun a intérêt à ce qu'il soit commun à l'ensemble des opérateurs. J'invite les utilisateurs que vous êtes à être très vigilants sur cette question : qui contrôle la navigation sur Internet contrôle tout l'accès à l'information. La plupart des bases de liens actuelles sont gratuites mais rien n'empêche un consortium comme CrossRef de se transformer en organisation " for profit ".

La quatrième catégorie d'annonces regroupe les annonces sur la revalorisation des bases bibliographiques. Celles-ci sont un sas d'entrée vers les produits d'e-commerce de l'information et un sas d'orientation vers les portails spécialisés, ce qui est une valeur ajoutée importante. Au dernier OnLine Information, RidElsevire ne communiquait curieusement pas sur sa base de 4 500 périodiques en ligne et en texte intégral, mais sur la base bibliographique

frontale de 10 millions de références. Cette inversion de la communication est significative. De nombreuses bases bibliographiques ont une forte notoriété : c'est le cas d'Urbamet ou de Biotechbase, de Mbase dans le champ de la biologie. Ces bases servent de briques de base à la constitution de portails spécialisés plus larges.

La dernière catégorie est celle des bases de preprint. Elle est essentielle car la modification des business models de l'édition scientifique et technique est en jeu. Dans le modèle classique de cette édition, non seulement les documentalistes paient l'abonnement mais, en plus, les chercheurs qui doivent publier les articles paient. Les chercheurs ne veulent plus passer deux fois au guichet. Pour y remédier, un puissant mouvement d'apparition de bases de preprint apparaît et il échappe largement aux grands éditeurs ; dans des secteurs que les grandes enseignes de l'édition irriguent mal, comme le secteur de l'urbanisme, ces phénomènes sont voués à se développer. La base de preprint se définit ainsi : c'est une base accessible en ligne où un auteur place son article dès qu'il l'a écrit, sans mise en forme. A j+2, l'article est indexé et directement accessible en ligne, gratuit pour la communauté mondiale. Attention, il ne faut pas entendre preprint au sens graphique du terme, mais bien au sens où, au moment où l'auteur envoie un fichier électronique Word, PDF ou HTML à son éditeur pour une publication dans une revue, il le place en même temps dans une base immédiatement accessible. Le délai de mise à disposition de l'information se trouve raccourci. Dans le modèle classique, il fallait jusqu'à 18 mois entre le moment où un article est soumis et celui où il est publié.

Un bon niveau de créativité sur les produits de l'information

Sur les 120 annonces de Londres, 50 concernaient réellement des produits nouveaux et dépassaient le stade de l'agrément sur des produits existants. Internet et la revalorisation des bases bibliographiques y sont pour beaucoup.

Produits : quête de la valeur ajoutée

Les portails construits autour des produits existants : le concept d'intégration

Un portail se résume souvent à un bouquet de liens derrière un écran d'accueil. Or un vrai portail doit permettre d'intégrer différentes sources d'information à partir d'une même requête utilisateur. Le meilleur exemple d'intégration est le portail grand public Yahoo. En tapant le nom d'une entreprise, on récupère à la fois son cours de bourse, une capsule d'informations décrivant ses activités, ses communiqués de presse et éventuellement une fiche sur ses dirigeants. Chacune de ses informations provient de sources différentes. La valeur ajoutée d'un portail, c'est cette valeur d'intégration. Or la plupart des portails sur le marché traduisent très imparfaitement cette logique. L'intégration nécessite du lien et de l'indexation, ce qui coûte cher. Les portails actuels ne sont cependant pas des outils inutiles. Au-delà de la fonction d'intégration, la valeur ajoutée résulte de la navigation.

Navigation Internet/ base de liens

Cette valeur est directement perceptible par l'utilisateur et permet de se différencier des offres d'informations gratuites.

Les niches d'informations propriétaires

Cette valeur ajoutée émerge en particulier dans le secteur de l'information financière : on trouve les cours de bourse gratuitement sur n'importe quel site boursier grand public. Pour vendre de l'information financière, il faut vendre de l'information propriétaire, par exemple, une information sur les spécificités des titres. Il faut aller vers un niveau plus spécifique de l'information pour créer des barrières à l'entrée.

Les produits résultant de l'extraction de données factuelles sur le Web

Au-delà de son caractère chaotique, le Web est une formidable source d'informations. Pour mobiliser ce potentiel, il faut que des paramètres de productivité extraient cette information. Certains acteurs développent des logiciels agents intelligents qui explorent automatiquement le Web, repèrent les informations pertinentes et les organisent. Par exemple, "Methods Finder" scanne toute la littérature pour chercher les protocoles d'expérimentation dans le domaine des biosciences. La base de données américaine "Compuvirus" extrait les données sur les souches virales informatiques.

L'extension de la couverture

Les grands agrégateurs inscrivent à leur catalogue de plus en plus d'informations, notamment des fonds bibliographiques comme Urbamet. Citons l'agrégateur OVID dans le champ des biosciences et Worldscope dans le champ de l'information financière.

La valeur ajoutée linguistique

L'information professionnelle se déployant dans un champ multilingue, il est essentiel de pouvoir accéder à partir d'une requête en français à une information dans un document de base en anglais. C'est la valeur ajoutée linguistique. En France, l'interface de l'INPI travaille dans une documentation multilingue. En Italie, la CERVED (organisme diffusant les fonds de registre du commerce) dispose d'une petite interface permettant d'accéder aux informations sur les entreprises italiennes avec des masques anglais, français.

La valeur ajoutée image

Cette valeur émerge alors que l'IEP était essentiellement de l'information textuelle. La mise en ligne de la banque de données sur les modèles et dessins déposés de l'INPI sera une première illustration en France. Ce qui est original, c'est que les modules de recherche incluent des moteurs de recherche image. Les prolongements seront très intéressants dans le domaine de l'urbanisme où l'image est primordiale.

Une plus grande finesse dans la granularité de l'information

Il faut accéder au niveau le plus fin possible. Dans l'information sur le développement urbain, c'est le niveau de l'entité territoriale minimale à partir de laquelle on peut récupérer des données sur l'aménagement. Les éditeurs évoluent vers des cribles géographiques plus fins pour augmenter la valeur de l'information. A IDT, cette année, "The economist intelligence unit" a présenté des séries macroéconomiques relevant des zones urbaines et non plus des marchés nationaux.

Evolution vers les data mining et les outils d'analyse prédictive

Les *data mining* sont des techniques pour produire de l'information à partir de traitements algorythmiques ou statistiques. L'ISI (Institute of Scientific Information) a lancé un produit, Essentiel scientific indicators qui est un traitement statistique de second niveau de fonds d'information déjà acquis. Cette problématique est très importante pour l'information brevet. Nous assistons au passage d'une IEP centrée sur la référence aux outils d'analyse de modélisation et aux outils prédictifs.

L'évolution du marché des infologiciels

Les différents outils

Les outils prédominants sont les outils d'édition Web, c'est-à-dire les applicatifs offerts aux éditeurs. Ceci a permis de baisser les coûts de l'édition électronique.

Les outils dominants sont les outils d'équipement de portails Intranet avec des offres comme celles de Verity, Mediaps. Des offres industrialisées apparaissent : les DRMS (Digital Rights Management System) qui sont des plates-formes fonctionnant en mode serveur de gestion des droits de propriété intellectuelle. C'est très important car les droits de propriété sont aujourd'hui l'un des principaux goulets d'étranglement freinant le développement de nombreux services en ligne.

Les outils linguistiques sont également présents. Le développement du B2C a un effet fantastique sur le développement de l'outil linguistique. Altavista et Yahoo ont acheté ces systèmes parce que l'utilisateur lambda de l'internet ne sait pas poser une équation en langage booléen. Or, il faut du traitement linguistique pour rendre les systèmes B2C performants.

Les outils de bibliothéconomie existent depuis trente ans. Ils ne gèrent plus seulement l'univers bien périmé de la bibliothèque mais tout l'univers de ses liens avec son environnement extérieur.

Enfin, les outils auteurs émergent. Ces outils sont mis à la disposition des chercheurs pour améliorer la productivité de la communication scientifique. L'ISI en fait une clé de son développement car il satellise ainsi les auteurs sur ces bases de preprint. Les outils auteurs ont donc une fonction de rétention des auteurs.

Large dispersion des acteurs et des types d'application

Ce secteur est très fragmenté avec une multiplicité d'offres redondantes dont la notoriété ne dépasse pas le cadre national. Dans les quatre prochaines années, nous assisterons à une concentration certaine avec des acteurs français comme Mediaps.

Faible implication des leaders du logiciel comme IBM ou Microsoft

Les leaders laissent les start-up des infologiciels vérifier la validité des potentiels de marché puis les rachètent. Un changement s'amorce : les leaders publient des annonces sur les outils d'équipement de portails Intranet.

Prédominance des “ technologies pauvres ”

A l'exception des outils linguistiques, la plupart de ces outils logiciels incorporent des technologies pauvres, c'est-à-dire des technologies qui résultent de recherches menées il y a plus de dix ans. Les start-up misent plutôt sur l'industrialisation de technologies développées dans les années 90 que sur des développements très innovants qui restent chers.

Les autres tendances

On constate une prédominance des problématiques d'e-commerce de l'information, avec les problèmes de tarification des périodiques. Les producteurs d'information suivent une double tendance. D'une part, des acteurs comme OCLC ou RoweCom se positionnent vers le e-commerce de l'information. Ils se réapproprient la commercialisation en ligne. D'autre part, de nombreux éditeurs passent un niveau de commercialisation qui est celui de l'article et non plus de la revue.

Signaux faibles des marchés

Les frontières entre B2B et B2C s'estompent

Aujourd'hui les acteurs veulent servir avec le même fond d'informations les marchés les plus larges possible, à la fois professionnels et grand public. Ainsi, les sites boursiers grand public sont équipés de fils d'information qui n'existaient que sur les postes des salles de marché il y a cinq ans. Cette modification des frontières est également perceptible dans le champ de l'urbanisme et de l'environnement. Aux Etats-Unis, de nombreux sites citoyens s'approprient des fonds de données professionnelles. Je pense en particulier aux sites qui répertorient les dépôts de déchets chimiques. C'est une problématique qu'Urbamet doit prendre en compte.

Substitution de la labélisation à l'information

Autrefois pour prendre une décision de solvabilité, il fallait accéder à des bilans et à des rapports. Ce que propose un acteur comme la COFACE, c'est de donner une note de solvabilité en échange de l'information. C'est le système de la notation financière avec les grandes agences comme Moody's, Standard and Poor's. Le modèle économique de l'information est bouleversé puisque c'est l'émetteur d'information qui paie et non plus le consommateur.

Les silences

Les tendances très profondes n'apparaissent pas dans ces annonces liées au secteur marchand.

Les nouvelles formes d'organisation de la demande

De très grands consortiums d'utilisateurs qui vont bien au-delà des logiques des centrales d'achat apparaissent. Par exemple, un consortium américain a créé une dizaine de revues électroniques gratuitement en ligne pour se donner un peu de marge de manœuvre face aux

grands éditeurs. Il crée ainsi une offre alternative. Cette nouvelle forme d'organisation de la demande sous forme de consortium est très présente dans le monde anglo-saxon et aussi en Europe continentale mais avec une problématique plus étroite. Les consortiums anglo-saxons se situent vraiment dans une logique d'acteurs de marché et ne sont pas seulement acteurs d'un Rotary Club de l'information.

Les stratégies de nouveaux intervenants potentiels

Aujourd'hui l'information peut être un produit d'appel gratuit pour des opérateurs vendant autre chose. En 1996, IBM a fait basculer sur un site gratuit l'intégralité des brevets américains depuis 1976. IBM voulait illustrer la puissance de son système de SGBD. C'est aussi vrai pour les Big Five du consulting tel Ernst & Young. L'information est un produit d'appel pour leurs prestations de consulting.

Dessin animé

Il s'agit simplement d'une série de copies d'écrans.

Batirama, archétype du portail

Ce site est décliné par Bertelsman dans sa dimension européenne. C'est l'archétype du portail. On y trouve des informations défilantes, de l'actualité, de nombreux liens mais le niveau d'intégration est faible.

BuildOnline, valeur ajoutée des outils métier

Ce portail est réservé au secteur du bâtiment. Il a misé sur la valeur ajoutée des outils métiers, particulièrement avec la fonction project on line qui est un outil de gestion de chantiers. C'est ce qu'on appelait une armoire à plans. Cet outil permet de gérer toute la conduite de chantiers depuis la mobilisation des fournisseurs jusqu'à la réception de l'ouvrage. La convergence avec les logiciels applicatifs est une des grandes voies de la valeur ajoutée de l'information.

Waste : la valeur ajoutée image

Ce site est signalé sur la rubrique veille du site Urbamet. Ce site recense par des images les dépôts de déchets sur le territoire américain. Il y a aussi le célèbre site de suivis photographiques de la Potsdamer Platz qui présente des panorama successifs depuis 1995.

La dimension cartographique

C'est une valeur ajoutée très importante pour vous. Il est possible d'avoir des outils manipulant des données sur une base géoréférencée, sur une carte. Un site permet même de construire en mode temps partagé une carte indirecte. L'utilisateur rentre ses données et une visualisation apparaît immédiatement.

Vous avez compris que cette présentation n'était qu'un long spot publicitaire pour les prestations de M.V.Services et Conseil.