

Catherine Lupovici*Jouve Systèmes d'information*

RÉVOLUTION ÉLECTRONIQUE ET NORMALISATION

LE MARCHÉ de l'information est en train de vivre une révolution technologique qui touche globalement tous les acteurs des différents domaines actuels : les auteurs, les éditeurs, les libraires, les imprimeurs, le monde de la télévision, les télécommunications, le secteur de la musique et du film, les constructeurs de matériels informatiques, les créateurs de logiciels, les fournisseurs de jeux vidéo et tous les nouveaux venus qui peuvent jouer un rôle dans la mise en place des nouvelles relations entre les acteurs traditionnels pour l'offre et la distribution électronique d'information.

Les estimations économiques du marché potentiel de l'édition électronique en Europe sont d'environ 80 milliards de francs en l'an 2000, avec une croissance très importante au-delà. Il est en effet évident que, si la plupart des publications vont continuer de paraître sous forme papier, les produits traditionnels papier perdront leur rôle primordial dans plusieurs domaines où ils ne seront plus l'unité de mesure de l'activité intellectuelle.

Beaucoup d'éditeurs ne sont pas encore conscients des changements radicaux de l'industrie de l'information, d'autant plus qu'un certain nombre d'incertitudes pèsent encore sur les technologies, les marchés, et que le futur de l'industrie de l'information n'est pas encore parfaitement visible.

De même, la plupart des acteurs de l'édition, de l'électronique et des communications n'ont pas encore réalisé que les industries en question sont en train de s'interpénétrer pour construire les bases de l'industrie du futur, avec de nouvelles règles de fonctionnement, une redistribution des rôles et de nouvelles bases pour la concurrence.

Le succès de l'information électronique dépend aujourd'hui uniquement de son appropriation par les producteurs et les utilisateurs. En effet, les technologies existent et la meilleure preuve en est que la normalisation est en cours et qu'elle se déroule transversalement aux secteurs traditionnels et marque parfaitement la convergence et l'interpénétration de domaines auparavant autonomes.

Esquisse du futur marché de l'édition électronique

Le marché prévisible aujourd'hui possède diverses caractéristiques.

– Le produit de l'information électronique est multimédia au sens où le contenu de l'information est un mixage des différents types d'information (textuelle, sonore, image animée, etc.). Les produits électroniques resteront vraisemblablement *off-line* jusqu'en l'an 2000, ce type de produits tels que CD-ROM ou CDI ayant la faveur des éditeurs traditionnels car il a toutes les caractéristiques de conception, de diffusion et de consultation des produits papiers.

Cependant, à plus long terme, on pense que les produits *on-line*, *off-line* et papier coexisteront de manière complémentaire, voire imbriquée. L'édition d'un produit sera nécessairement sur supports multiples, ce qui aura des répercussions sur les modes de distribution.

– Les Etats-Unis et le Japon sont en avance et ont davantage d'atouts que

l'Europe. En effet, ils ont plus de savoir-faire dans l'industrie de la musique et du film, et l'avènement du multimédia est dans leur intérêt économique.

La normalisation se construit généralement sur un ou plusieurs usages et quelquefois sur des standards de fait

– L'Europe possède de son côté une grande maîtrise de l'information textuelle et les éditeurs traditionnels ont tendance à camper sur leur position de force du tout papier associé au droit de propriété de l'information textuelle et aux relations bien établies avec les auteurs et les lecteurs.

Ils ont généralement des difficultés à imaginer et créer des produits électroniques. Pour justifier leur hésitation à investir dans ce domaine, ils invoquent le manque de normalisation, l'inadéquation de la protection des droits de propriété, l'hétérogénéité des plates-formes matérielles, l'immaturité du marché. Cette attitude de retrait se prolonge dans leur non-engagement effectif dans les processus de normalisation.

– L'émergence du marché de l'information électronique dépend non seulement du secteur de la production de l'information, mais aussi de l'industrie de l'électronique (ordinateurs, télécommunications) qui y voit une stimulation des évolutions des matériels et des développements logiciels, ainsi qu'une diversification de son secteur d'activité, génératrice de profits. Les partenaires de l'industrie de l'électronique sont acteurs et même moteurs de la normalisation.

– Les secteurs prometteurs du marché sont : l'édition et la communication institutionnelle, le secteur financier, les jeux – en particulier ceux qui concernent les enfants –, les documents de référence, le juridique, le secteur scientifique-technique-médical (STM), l'éducation, le secteur voyage / hobby / tourisme.

– Deux marchés peuvent être identifiés : le marché de l'utilisation professionnelle et le marché de l'utilisation domestique. Ces recours seront combinés à deux plates-formes de consultation : les ordinateurs personnels ou le moniteur de télévision.

– Les technologies associées au développement de l'information électronique d'ici l'an 2000 sont : les disques compacts (CD-ROM, CD-XA, CDI, CD Photo...), les cartes à mémoire, les réseaux, les techniques de compression, les périphériques de lecture portables, l'interactivité.

Les grandes tendances de la normalisation

Tout d'abord il faut rappeler pour mémoire que la normalisation de tout ce qui touche à l'information et à l'électronique est exclusivement élaborée au niveau international. Il ne peut plus y avoir dans ce domaine de démarche normative nationale isolée du contexte international. Bien entendu les discussions sont toujours le reflet des intérêts économiques des différentes parties du monde.

Evolutions de la démarche normative

La normalisation est traditionnellement organisée en secteurs verticaux par grands domaines, subdivisés par thèmes et questions de travail, selon les principes d'une classification arborescente. Cette organisation des activités présente les mêmes inconvénients que les classifications du savoir lorsque l'on est en pleine mutation et c'est le cas des technolo-

gies informatiques et de tous les secteurs d'activité qui les utilisent pour la création, le traitement, l'échange et la diffusion des informations. Il faut alors créer provisoirement des structures plus transversales permettant d'attendre qu'un nouvel ordre des choses émerge et se stabilise. Les différents secteurs traditionnels de normalisation d'activités liées à l'informatique sont donc actuellement en train de multiplier les contacts transversaux entre les commissions, et l'on assiste à la création de groupe *ad hoc* ainsi qu'à des réunions internationales regroupant les experts de différents comités, pour aborder les problèmes communs et définir des stratégies globales de normalisation. C'est le cas par exemple de l'harmonisation des données-types entre les dictionnaires de données développés par les groupes les plus divers (médecine, documentation, industrie...) avec la création du Groupe ISO de coordination des données-types (DCG).

La normalisation est le résultat d'accords entre des parties concernées, dans l'intérêt général d'un domaine d'activité. Elle se construit généralement sur un ou plusieurs usages et quelquefois sur des standards de fait. C'est une approche ascendante qui se construit sur la pratique. Cette approche à la fois verticale et ascendante peut conduire au développement de normes au cas par cas, plus ou moins concurrentes dans des domaines connexes et qui deviennent particulièrement gênantes dans les phases de mutation.

L'organisation transversale, qui se met en place actuellement, fonde donc son activité sur une démarche normative descendante, par une approche conceptuelle globale conduisant à des normes-cadres sur des thèmes transversaux. On aboutit ainsi, non plus à des normes techniques appliquées mais à une modélisation conceptuelle d'un domaine, visant à déterminer le cadre stratégique de réalisation des normes techniques plus concrètes. Cette démarche permet de mieux prendre en charge tout un marché en devenir, pour lequel on va définir et planifier le développement cohérent de familles de normes compatibles, qui pourront être utilisées en association.

Produits électroniques en l'an 2000 en Europe	
Domaine	Part de marché
Fiction	1-10 %
Enfant	15-25 %
Education	10-20 %
Référence	15-25 %
STM (Sciences, technique, médecine)	20-30 %
Droit	15-25 %
Autres	1-10 %

C'est sur une telle démarche que se fonde aujourd'hui la normalisation du codage de l'information multimédia.

Dans la pratique, la démarche normative des secteurs stratégiques est le plus souvent une combinaison raisonnable des deux approches, de manière à tenir compte de l'existant, tout en ayant une réflexion prospective.

La systématisation de la seconde démarche n'est pas en effet sans danger. Une norme peut devenir un cadre stratégique de développement de produits qui n'existent pas encore. Elle peut n'être qu'une réflexion théorique destinée à susciter des pratiques et des produits qui n'existeront jamais autrement que sous forme respective de recherche fondamentale et de prototypes qui ne seront jamais commercialisés. Il y a même risque d'aboutir à la création en parallèle de standards, ignorant de fait la norme, à l'image de ces espaces verts possédant deux types de chemins pour les traverser : ceux conçus par l'architecte, déserts, et les chemins sauvages tracés par l'utilisation permanente.

Enfin, la normalisation est un processus long, jalonné de procédures à caractère administratif, qui sont le garant de la décision démocratique de tous les pays membres des organisations internationales. Ainsi, il n'est pas rare que l'établissement d'une norme prenne cinq ans, entre le moment où elle est introduite comme sujet de travail, et le moment où elle est publiée dans les versions linguistiques officielles. De tels délais sont incompatibles avec l'activité économique des technologies de l'information et plusieurs processus peuvent être adoptés pour les réduire. Ainsi :

– Une norme nationale d'un pays particulièrement en avance peut être introduite selon une procédure accélérée, « *fast track procedure* », qui permet d'arriver directement à la discussion de forme finale du texte normatif adopté dans ce pays. Une telle procédure donne évidemment un poids économique considérable au pays concerné, puisqu'elle consacre par la norme une avance technologique déjà importante. En contrepartie, elle permet aussi la mise en place plus rapide et dans de meilleures conditions de produits clés.

– Beaucoup de normes techniques informatiques ne sont publiées que dans la langue d'élaboration du document : l'anglais. Il est vrai que, dans certains cas, leur traduction serait assez factice, car la terminologie du domaine est uniquement anglaise. Cependant, dans les applications satellites de ces techniques, la question de la version française est un dilemme permanent : faut-il mettre une priorité sur l'élaboration d'une version française, donc mobiliser l'énergie d'experts trop peu nombreux, pour une traduction qui sera souvent effectuée au détriment de leur engagement dans la discussion de fond et de leur participation active à la défense des intérêts économiques

Ce qui est nouveau, c'est la volonté d'avoir une normalisation cohérente et compatible du codage de l'information, de manière à permettre l'utilisation d'informations multimédias sur des plates-formes matérielles uniques.

– Texte

Alors que la normalisation du codage des caractères, pour leur manipulation et stockage en machine, a jusqu'alors été plutôt anarchique, consacrant les pratiques de divers secteurs d'activité, on travaille aujourd'hui à une norme d'échange fondée sur un code universel de tous les systèmes d'écriture, dans tous les domaines (y compris les mathématiques, l'édition, le domaine

La normalisation est un processus long, jalonné de procédures à caractère administratif

nationaux qu'ils représentent ? Faut-il ajouter le délai de traduction à celui de la publication ?

– La normalisation générique du minimum nécessaire au développement du secteur est pratiquée de préférence à la norme très détaillée. En complément, des *Rapports techniques* sont élaborés selon une procédure plus souple. Cette démarche permet de tester l'utilisation de pratiques conseillées, de corriger plus rapidement les problèmes rencontrés et de décider ainsi du niveau de normalisation nécessaire et suffisant dans le domaine.

Les normes clés du codage numérique de l'information électronique

Normes de contenu

Le codage de l'information constituant l'information électronique connaît depuis l'origine une normalisation par type d'information : textuelle, graphique, image fixe, son, image animée, dont l'état d'avancement n'est pas identique dans tous les domaines.

bibliographique...). Ce jeu universel de caractères codés UCS (ISO 10 646), entrepris en 1984, est en train d'aboutir, après des discussions vives entre les différentes parties et en particulier avec le secteur bibliographique.

– Image et son

De gros investissements humains et technologiques viennent d'être consentis pour la normalisation des codages de l'image et du son, dont l'enjeu est l'avènement du multimédia. Les normes, parues ou à paraître en 1993, portent sur la compression des informations numériques, pour permettre une optimisation du rapport qualité / volume-débit d'information sur les supports ou canaux actuels. Trois groupes de travail ont permis l'aboutissement de normes dans différents domaines :

- Groupe codage d'image photographique (JPEG), ISO 10 918, 1992 : norme de définition de méthodes de compression / décompression d'image photographique fixe, couleur ou monochrome. Les méthodes de compression / décompression proposées prennent en compte les possibilités de reconstruction d'image en mode séquentiel ou progressif pour une

consultation facilitée selon les types d'usage.

- Groupe codage d'image binaire (JBIG), ISO DIS 11544 : norme de codage de l'image fixe à deux niveaux. Comme pour JPEG, il est possible de reconstruire l'image en mode séquentiel ou progressif.
- Groupe codage d'image animée (MPEG) : famille de normes de codage d'images animées numériques, ainsi que du son et des données complémentaires associées. Elles sont destinées à satisfaire les besoins de stockage et de restitution d'image et de son, à partir de supports utilisant les technologies courantes (magnétique, optique), avec leurs contraintes de débit.

Le codage de l'image animée est divisé en trois sous-groupes : image vidéo, son et combinaison image / son.

Normes de structure

Les normes de structure permettent de coder la façon dont le contenu est organisé et, s'il s'agit de multimédia, de coder la synchronisation spatiale pour les informations statiques, et temporelles pour les informations qui sont liées à un déroulement dans le temps.

Deux codifications de la structure ont été normalisées, à partir de l'information texte-graphique-image, l'une venant du monde de l'édition et de la photocomposition, l'autre du monde de la bureautique : SGML (*Standard Generalized Markup Language*) et ODA (*Office Document Architecture*). Les deux s'étendent actuellement au multimédia-hypermédia. Les deux normes ont des domaines d'applications différents, bien qu'elles soient souvent présentées comme concurrentes. Ils existe des passerelles entre les deux, *via* leurs formats d'échanges normalisés.

Un travail plus conceptuel est par ailleurs réalisé pour un modèle générique de représentation d'objets multimédias et hypermédias, offrant les possibilités de synchronisation et de navigation entre objets : MHEG.

– SGML / HYTIME

SGML, ISO 8879, est une norme pour la codification de la structure logique de documents révisables. En

association avec toute une famille de normes, SGML permet l'échange de documents électroniques, de la description de leur structure logique, de leur format, des informations sur la structure physique. Sont ainsi autorisés l'affichage ou l'impression, de façon indépendante des logiciels, des matériels, des périphériques et des applications auteur ou utilisateur du document.

Les principales normes associées qui permettent le traitement et l'exploitation des documents codés en SGML sont :

- SDIF (*SGML Document Interchange Format*), ISO 9069,
- DSSSL (*Document Style Semantics and Specifications Language*), ISO DIS 10179,
- SPDL (*Standard Page Description Language*), ISO DIS 10180, Definition and Identification of Glyph Fonts, ISO 9541,
- HYTIME (*Hypermédia / Time based structuring language*), en développement,
- SDML, pour les documents musicaux, en développement.

SGML a été conçue principalement pour le codage de la structure des documents textuels, mais permet le rattachement d'entités externes d'autres types de documents traités dans leur codage propre, par exemple d'images. HYTIME permet l'extension au multimédia et le codage de la synchronisation temporelle. Le lan-

L'emploi du SGML se généralise dans le monde de l'édition

gage SGML n'impose aucune limite de hiérarchie dans la structure, permet de définir les éléments récursifs et traite les exceptions, ce qui explique son succès dans le domaine de la documentation technique, en particulier dans le cadre du projet CALS (*Computer Aided Acquisition and Logistic Support*) du ministère de la Défense aux Etats-Unis. SGML est aussi utilisée par des organismes tels que l'EDF, l'Aérospatiale, l'Office des

publications de la CEE, l'Office européen des brevets, l'ISO pour les normes. Son emploi se généralise dans le monde de l'édition. SGML peut être utilisée non seulement pour la saisie en vue de l'impression, mais aussi pour l'alimentation de bases de données ou la réalisation de produits électroniques tel le CD-ROM.

Les outils permettant de manipuler SGML existent sur le marché : aujourd'hui, on peut choisir entre des outils très professionnels ou des outils plus orientés vers l'utilisateur final offrant une ergonomie de type WYSIWYG (*What You See Is What You Get*).

Les outils SGML conçus initialement pour le monde de l'édition traditionnelle, qui séparent bien les fonctions de structuration logique de l'acquisition de l'information, de composition et de mise en page, sont donc en train de s'étendre aussi vers la bureautique et le multimédia.

– ODA / HYPERODA

ODA, ISO 8613, est une norme plus globale définie dans le cadre de l'OSI. Elle fournit un ensemble de concepts pour décrire les documents (modèle de document) et les processus sur les documents (modèle de traitement de document) pour faciliter la création et l'échange de documents bureautiques entre systèmes hétérogènes. Elle fournit également un format d'échange ODIF (*Office Document Interchange Format*).

Les extensions de la norme sont en cours, au niveau des types de contenu (audio, vidéotex, couleur et niveaux de gris) et des données (tableaux issus de tableurs, graphiques).

La démarche normative ODA est une initiative européenne typiquement descendante, dont le domaine d'application est primitivement la bureautique, avec une conception globale du codage et de l'exploitation du document. Elle a été entreprise en 1981 par le Comité consultatif international télégraphique et téléphonique (CCITT), et la norme a été développée conjointement par le CCITT et l'ISO. ODA a donc particulièrement la faveur de l'industrie des télécommunications. Des produits commencent à voir le jour sur

le marché, grâce notamment au consortium ODAC, fondé par les principaux constructeurs informatiques : IBM, DEC, Bull, Siemens...

Un des apports importants d'ODA est cependant la modélisation conceptuelle du document avec les notions clés de représentation universelle des documents bureautiques : structure logique et structure physique ; structure générique et structure spécifique ; classes de documents ; profil du document.

La description du contenu est basée sur des objets logiques et des objets physiques.

– MHEG / AVI

L'objectif des travaux MHEG (Groupe codage d'information multimédia et hypermédia) commencés en 1989 est la représentation et le codage d'objets multimédias et hypermédias génériques. Le modèle doit permettre la manipulation de ces objets comme des unités d'informations autonomes par des applications ou des systèmes. La modélisation de leur échange, à l'intérieur d'une application, ou entre des applications dif-

Les services clients particulièrement concernés par le développement de MHEG sont les AVIS (*Audio Visual Interactive Scriptware*). Les AVIS décrivent les liens entre objets dans un script. La norme sur les AVIS est programmée pour fin 1994. Elle est confiée au groupe MHEG pour la partie codage, et à la Commission nationale bureautique pour la description fonctionnelle.

Les normes clés pour l'échange d'informations électroniques

L'échange d'informations électroniques comporte fonctionnellement : la normalisation du support nécessaire à l'échange d'informations entre des systèmes informatiques hétérogènes pour une application donnée, c'est-à-dire nécessaire pour véhiculer tout le contenant de l'information – c'est l'objectif de l'OSI (*Open Systems Interconnection*) –, et la normalisation nécessaire à la compréhension entre les applications privées

venus du monde entier et travaillant sur plus de 50 sujets différents.

Cependant les normes d'interconnexion ne sont qu'une première étape vers l'interfonctionnement total des systèmes et la « portabilité » des applications. OSI est une nouvelle manière de penser, qui a permis de donner l'impulsion vers de nouvelles normes, en particulier dans le domaine du traitement réparti, et la prochaine étape apparaît déjà autour du traitement réparti ouvert (ODP).

Malgré son succès, OSI n'est pas la seule approche viable pour les communications, et les constructeurs n'ont pas attendu pour offrir des architectures privées, répondant aux besoins, et qui continueront à coexister avec l'OSI. Cependant l'OSI ne peut plus être ignoré, et la plupart des fournisseurs de produits de communication ont choisi soit d'intégrer l'OSI, soit d'offrir des passerelles. L'OSI est généralement soutenu par les administrations, par exemple la décision 87 / 95 du Conseil des Communautés européennes, qui exige de faire référence aux normes dans les marchés publics et fournitures d'équipements informatiques, et cite explicitement les normes OSI.

Les premières mises en œuvre importantes sont apparues en 1988, dix ans après la mise en chantier de l'OSI. C'est également à cette date qu'apparaissent les premiers produits commercialisés. Cependant la mise en œuvre d'OSI reste encore assez coûteuse : les produits OSI sont nouveaux et nécessitent un investissement humain de la part de l'utilisateur ; les techniciens OSI sont rares ; les systèmes actuels ne sont pas capables d'exploiter les nouvelles possibilités de l'OSI ; et de nouveaux systèmes doivent être développés. Toutefois, l'investissement peut être payant puisqu'il devrait éliminer les coûts de conversion future et éliminer les passerelles multiples.

– EDI

Le concept d'EDI est apparu en 1986, avec la normalisation EDIFACT (*Electronic Data Interchange for Administration, Commerce and Transport*). Au-delà de cette première démarche normative, on assiste à une

Malgré son succès, OSI n'est pas la seule approche viable pour les communications

férentes, est faite de manière générique, englobant les réseaux de télécommunication ou de diffusion, mais aussi le stockage sur des supports variés tels que CD-ROM et disquettes. Enfin, la norme devra répondre aux besoins des applications interactives.

Elle est prévue en deux parties : objets multimédias d'acquisition et de restitution synchronisés ; objets hypermédias.

Comme pour ODA, le travail normatif est effectué conjointement par l'ISO et le CCITT. Ce travail normatif est stratégique pour l'industrie des télécommunications, mais il associe aussi des constructeurs informatiques, des industriels de l'industrie grand public, des Sociétés de services et de conseils en informatique et des universités.

couvrant les données, les messages, la grammaire des échanges, c'est-à-dire le contenu : c'est l'objectif de l'EDI (Echange de données informatisées).

– OSI

La normalisation OSI a été initiée, en 1978, par une démarche conjointe de l'ISO et du CCITT en vue de définir une architecture normalisée des communications facilitant l'interfonctionnement entre des systèmes différents : constructeurs et technologies différentes, principes d'administration différents, installés dans des pays différents et dans des entreprises différentes.

Cette normalisation représente un investissement considérable. Aujourd'hui, les réunions internationales OSI rassemblent plus de 400 experts

tentative de normalisation globale d'EDI, avec la définition d'un modèle conceptuel des échanges de données informatisés (c'est-à-dire sans aucune intervention manuelle).

La normalisation EDI comporte l'élaboration de la syntaxe normalisée (la grammaire des messages échangés) et l'élaboration de dictionnaires normalisés de données (les données élémentaires, qui sont ensuite regroupées en segments, puis en messages).

Les normes EDIFACT, ISO 9735 pour la syntaxe et ISO 7372 pour les données ou TDED (*Trade Data Element Directory*), sont élaborées dans le cadre des Nations Unies et de l'ISO, au niveau international. Pour la France, elles sont regroupées au sein

de l'AFNOR dans le cadre d'EDIFRANCE, qui rassemble les professionnels intéressés de manière sectorielle.

Les secteurs les plus actifs sont l'automobile, la banque, le bâtiment, les industries chimiques, la distribution. Peu à peu les utilisations d'EDI se multiplient et, par exemple, le service de commandes des libraires « Electre Transmission » va devenir prochainement « Edilectre » pour marquer ses nouvelles orientations.

Les normes existantes, développées dans une approche ascendante, sont cependant trop rigides et marquées par le domaine initial de normalisation, ce qui gêne leur adoption dans d'autres domaines. Une réflexion est

conduite pour mieux prendre en compte, dans un modèle global, les spécificités de domaines particuliers, tel le secteur bibliographique.

La normalisation du CG46 / CN4

Cette commission de normalisation consacre son activité au domaine de l'automatisation en information et documentation. Il existe aujourd'hui deux époques de thèmes de travail, qui correspondent à des normes en fin d'élaboration, sans qu'un nouveau domaine de normalisation émerge vraiment aujourd'hui. Tous les secteurs actifs sont, de plus, confrontés à

Liste des normes ISO et AFNOR

ISO/IEC 646 : 1991

ISO 5426 : 1983

Information technology - ISO 7 bit coded character set for information interchange
Extension du jeu de caractères latins codés employé pour l'échange d'information bibliographique

DIS 5426-2 (DIS = Draft International Standard = Projet de norme internationale)
Information et documentation - Extension du jeu de caractères latins codés employé pour l'échange d'information bibliographique - Partie 2 : caractères latins utilisés dans les langues européennes peu répandues et typographie obsolète

ISO 5427 : 1984

Extension du jeu de caractères cyrilliques codés employé pour l'échange d'information bibliographique

ISO 5428 : 1984

Jeu de caractères grecs codés employé pour l'échange d'information bibliographique

ISO 6438 : 1983

Documentation - Jeu de caractères africains codés

pour l'échange d'informations bibliographiques

ISO 6630 : 1986

Documentation - Caractères de commande bibliographiques

DIS 6861-2

Information et documentation - Jeux de caractères codés des alphabets anciens pour les langues slaves et les langues européennes non slaves écrites en cyrillique, pour l'échange d'information bibliographique

DIS 6862-2

Documentation - Jeux de caractères codés mathématiques pour l'échange d'information bibliographique

DIS 8957

Information et documentation - Jeu de caractères codés de l'alphabet hébreu pour l'échange d'information bibliographique

DIS 10585

Information et documentation - Jeu de caractères codés de l'alphabet arménien pour l'échange d'information bibliographique

DIS 10586

Information et documentation - Jeu de caractères codés de l'alphabet géorgien pour l'échange d'information bibliographique

DIS 10754

Information et documentation - Extension du jeu de caractères codés de l'alphabet cyrillique pour les langues non slaves pour l'échange d'information bibliographique

CD 11822 (CD = Committee Draft = Avant Projet de norme internationale)

Information and documentation - Extension of the Arabic alphabet coded character set for bibliographic information interchange

ISO/IEC 10646-1

Information technology - Universal Multiple - octet coded character set (UCS) - Part 1 : Architecture and Basic Multilingual Plane

ISO 2709 : 1981

Documentation - Disposition des données sur bande magnétique pour l'échange d'informations bibliographiques

NF ISO 2709

NF ISO 8777

à l'édition
Information et documentation - Commandes pour les systèmes interactifs de recherche d'information

EP jusqu'au 20 mai 1992 (EP = Enquête probatoire) (Z.47-050)

DIS 12083

Electronic manuscript preparation and markup

NF EN 28879 : 1990

Systèmes bureautiques - Langage normalisé de balisage généralisé (SGML)

ISO 7154 : 1983

Documentation - Principes de classement bibliographique

Z 44-080

ISO/TR 8393 : 1985

Documentation - Règles de classement bibliographique de l'ISO (Règles standards internationales de classement bibliographique) - Concrétisation des principes de classement bibliographique par des règles modèles

l'examen de leur travail en cours en regard des secteurs généraux et de l'émergence des normes-cadres, qui impliquent de redéfinir les objectifs et les méthodes.

Thèmes de travail de la décennie 70

C'est l'époque de la normalisation nécessaire à l'émergence de l'automatisation de l'information bibliographique, à son échange et à son interrogation. Certaines des normes, commencées à l'époque de la naissance du format MARC et de l'interrogation des banques de données, entraînent encore une activité de normalisation aujourd'hui, soit par continuation du travail dans différents domaines (codage des caractères), soit par révision périodique de la norme (ISO 2709), soit parce que le sujet a été particulièrement épineux et réorienté maintes fois avant d'aboutir (Langage commun des commandes, CCL).

Codage des caractères à usage bibliographique

Le travail représentait une extension du codage des caractères en général. Les bibliographes ont défendu bec et ongles, et à juste titre, leur besoin de codage riche contre les codages proposés par les constructeurs de matériels informatiques. Cette normalisation a été entreprise à partir de la réflexion de l'American Library Association (ALA), puis à l'initiative des besoins conjugués de la bibliothèque du Congrès et du réseau RILIN (*Research Libraries Information Network*) pour le traitement en machine des systèmes d'écriture non latins. La norme de base que nous connaissons tous est ISO 5426, mais il existe aussi un codage grec bibliographique, cyrillique slave et non slave bibliographique, arabe bibliographique, hébreux bibliographique, chinois-japonais-coréen bibliographique... Cette normalisation a continué à son rythme tranquille jusqu'à l'arrivée du travail sur le code universel où les arguments de départ ont été réexaminés et défendus. Il a été

décidé, en 1992, qu'aucune nouvelle norme ne serait commencée, et que l'on se contenterait de terminer les normes en cours. Il a aussi été convenu qu'un travail d'appropriation du code universel serait nécessaire, dans notre domaine, pour les échanges bibliographiques, travail qui pourra être pris en charge par des associations professionnelles qui ne justifieraient pas nécessairement d'une activité normative spécifique. Notre problème en ce domaine est le suivant : la nécessité du codage riche, à une époque où la solution technique était difficile et proposée par un pays de langue anglaise qui n'utilisait pas dans sa propre culture de caractères accentués, a conduit à fausser notre vision du codage des caractères. Nous mêlons codage interne, codage d'échange et ergonomie de saisie, chose qui n'est plus nécessaire aujourd'hui avec les stations de travail. La question est importante économiquement : quand déciderons-nous de passer au code universel, et comment, alors que beaucoup de nos systèmes de catalogage et tous les systèmes de diffusion des informations bibliographiques fonctionnent sur la base d'un codage qui nous est propre, avec parfois encore des solutions clavier fermées ?

Que ferons-nous lorsque nous aurons à mettre à disposition de nos lecteurs sur un OPAC étendu des documents électroniques reçus des éditeurs en codage universel ? Cette question risque malheureusement de nous occuper dans les années à venir. Il

La norme ISO est basée sur le langage commun de commande

faudra cependant rompre un jour l'isolement dans lequel nous nous sommes enfermés. Quand ? Comment ? A quel coût ? Voilà un beau sujet de réflexion pour la profession, d'autant plus qu'il existe aussi un standard de fait, concurrent de la norme ISO, développé par le consortium nord américain UNICODE, et

qui semble répondre aux attentes des professionnels de la bibliographie aux Etats-Unis.

ISO 2709

La norme qui définit le format logique des données pour l'échange vient d'être révisée, et reconduite pour une période de cinq ans avec pour seule modification l'extension du support physique de la bande magnétique à tout support, ce qui d'ailleurs se pratiquait déjà. Il semble donc que notre format d'échange préféré ait de longues années devant lui... Et pourtant... des esprits modernes regardent du côté de la famille SGML comme langage plus générique de codage d'une information structurée associé à un format d'échange. Je ne pense pas, personnellement, que la question vaille aujourd'hui économiquement la peine d'être débattue et explorée. Il serait certainement préférable que les bibliothécaires investissent dans la connaissance et la manipulation de SGML sur le document complet lui-même et sur les services qui peuvent lui être associés, plutôt que de peaufiner le catalogage en SGML.

Le langage commun de commande (CCL), ISO 8777

Cette normalisation concernait initialement le dialogue entre l'utilisateur et les serveurs des grandes banques de données, utilisant des terminaux passifs. La norme ISO est basée sur le langage commun de commande, développé en 1977 pour Euronet, et on continue à la désigner dans la pratique par CCL, bien que la norme porte un titre différent. Elle s'est faite avec difficulté car la normalisation détaillée du dialogue fruste revenait à imposer un seul modèle de recherche documentaire et les serveurs se plaignaient qu'elle réduisait l'innovation. La position française était de faire une normalisation minimum, qui aurait permis que les grandes fonctionnalités offertes par les serveurs soient les mêmes, en laissant plus de place à la diversité pour le détail. La position française n'a fait que se renforcer avec l'introduction de l'usage du Minitel comme terminal de

recherche documentaire. Avec l'arrivée des banques de données sur CD-ROM, les Américains ont eux aussi revu leur position vers une simplification de la norme, qui a pu enfin être terminée et publiée.

Aujourd'hui le CCL attend donc, jusqu'à sa révision qui interviendra en 1997, l'épreuve de l'utilisation et de l'évolution des techniques permettant de concevoir des interfaces utilisateurs bien éloignées de l'ergonomie de l'interrogation en ligne des années 70.

Il existe aujourd'hui des défenseurs de CCL très convaincus et un certain nombre d'applications l'utilisent : le serveur européen ECHO offre l'interrogation de toutes ses applications au travers de CCL. Celui-ci est considéré aux Etats-Unis comme le standard des OPAC pour la recherche de type expert. C'est une norme nationale au Danemark.

D'autres considèrent qu'il doit devenir un standard des grandes fonctionnalités et des commandes internes aux systèmes, éventuellement échangeables entre deux systèmes, mais qui ne devrait plus rien avoir à faire avec l'utilisateur final du produit documentaire.

Les thèmes de travail des années 85

Ces thèmes correspondent à une seconde phase de l'informatisation des bibliothèques, qui voient une volonté de mettre en œuvre l'interconnexion des systèmes de bibliothèques, entre eux, et avec les systèmes de leurs partenaires. Ils correspondent aussi à une volonté d'utiliser la norme SGML et de normaliser en même temps que le milieu professionnel réfléchit à la question.

1985 est la date à laquelle les Etats-Unis ont pris la présidence internationale du sous-comité ISO, avec la volonté de faire aboutir ces deux thèmes principaux.

Les applications OSI

Le monde des bibliothèques avait conduit des projets de recherche depuis 1981 pour une utilisation appropriée de l'OSI. Les projets concernaient le prêt entre biblio-

thèques expérimenté au Canada, et la recherche d'informations bibliographiques expérimentée aux Etats-Unis et aux Pays-Bas. Ce sont donc ces deux applications particulières qui font encore actuellement l'objet de la normalisation de la mise en œuvre des protocoles ILL (*Inter Library Loan*), ISO 10160, ISO 10161 et SR (*Search and Retrieve*), ISO 10162, ISO 10163.

Ces protocoles concernent des activités traditionnelles de bibliothèques : les fonctions ILL couvrent la gestion de la commande et de la fourniture du document, les fonctions SR couvrent la recherche en ligne d'informations bibliographiques. Dans SR, la forme de la question peut être transmise au travers d'un format normalisé dans le protocole, reproduisant la structure des index des OPAC traditionnels de bibliothèques, ou au travers du CCL, ou au travers d'un autre format non défini et non normalisé. SR est actuellement défini pour la recherche d'informations bibliographiques.

Les textes proposés en norme, et modifiés dans le processus de normalisation, datent d'une époque où les protocoles OSI généraux étaient encore largement en développement. De plus, ceux qui sont disponibles ne correspondent pas exactement aux fonctionnalités attendues pour des modules traditionnels de bibliothèque. Une grande partie des débats a cependant porté sur la nécessité de concevoir du spécifique bibliothèque ou d'adopter des protocoles OSI plus généraux, pour conclure, dans tous les cas, à la nécessité de concevoir du particulier, car les besoins des bibliothèques étaient particuliers.

Les normes de base sont aujourd'hui publiées, et la normalisation continue avec celle des profils et des tests de conformité et d'interopérabilité.

Cependant, la révolution électronique va interpeller de nouveau les bibliothécaires sur leurs spécificités. La perturbation est déjà là, avec le souhait de prendre en compte le document électronique complet : doit-il être conçu comme une extension de ILL au sens de la fourniture électronique du document, ou bien appartient-il conceptuellement au domaine

de SR avec la recherche sur le texte intégral et le poste de consultation assisté par ordinateur ? De plus, les deux protocoles sont conçus pour un modèle d'interconnexion de serveurs en ligne : comment le CD-ROM, par exemple, et l'utilisation couplée CD-ROM / en ligne peuvent-ils être pris en compte ? En fait, les modèles conceptuels qui servent de base aux deux protocoles correspondent bien à l'informatisation de bibliothèques de 1980, mais sont très étroits par rapport à ce que sera le modèle à définir de la bibliothèque électronique et des services d'information électronique. Il serait d'ailleurs important que pour cette définition, les bibliothèques et centres de documentation aient une démarche descendante, à partir de la normalisation générale sur l'information électronique, et non pas ascendante, à partir du catalogue et de l'information bibliographique.

La maintenance de ces protocoles sera donc nécessaire tant que la bibliothéconomie classique constituera un marché particulier. De toute façon, ceux qui ont investi dans ce domaine seront mieux à même d'entrer rapidement dans le futur, voire de participer à sa construction.

Les répertoires de données

Cette normalisation a été initialisée en parallèle à celle des applications OSI, de manière à être cohérente avec elle en ce qui concerne les applications communes. Elle a commencé par la fonction prêt entre bibliothèques, puis la fonction acquisitions, puis la fonction recherche. Au fur et à mesure de leur avancement, les travaux ont été de plus en plus influencés par les développements de l'EDI et, après avoir été suspendus dans l'attente des résultats de la réflexion générale, ils devraient reprendre avec des méthodologies nouvelles, en appliquant la philosophie générale.

L'édition électronique

Le sujet a été abordé dès 1985 sur la base des informations provenant du projet « *Electronic Manuscript project* » mené aux Etats-Unis, pour aboutir à l'introduction du standard de l'AAP (*American Association of Publi-*

shers) comme norme internationale, en « *fast track procedure* », en 1987.

Cette norme d'application de SGML propose trois définitions de type de document (DTD) normalisées : une DTD pour les livres, une DTD pour les périodiques et une DTD pour les articles de périodiques. La norme devrait être publiée en 1993 (ISO 10 283). Son objectif est de normaliser la structure logique générique des grands types de documents, avec le balisage associé, de manière à faciliter les échanges de documents codés en SGML, et à offrir une base pour le développement de DTD particulières, afin d'aider les utilisateurs peu familiarisés avec SGML.

Bien que le sujet ait été introduit très tôt, il n'aboutira qu'en 1992, et on peut contester la procédure d'urgence utilisée, alors que le processus aura pris six ans – ce qui aurait été le délai normal en procédure ordinaire. En fait, l'utilisation possible n'était pas mûre, et la nécessité de normaliser des DTD pas évidente.

Il est vraisemblable que le thème va rester d'actualité, puisque les éditeurs européens scientifiques-techniques-médicaux ont souhaité introduire une nouvelle question de travail concernant la normalisation d'une DTD d'en-tête d'article de périodique et d'une DTD pour le corps de l'article. La première DTD, qui a par ailleurs été publiée par Springer Verlag sous le nom de MAJOUR, est destinée à assurer la fourniture aux utilisateurs de la référence bibliographique de l'article, comprenant le résumé d'auteur. La seconde, qui est toujours en cours d'élaboration, est destinée à

plus long terme à assurer la fourniture du document électronique lui-même. Elle pourrait être disponible fin 1993. Cette question a été introduite en normalisation internationale, sans lien avec la normalisation déjà en cours, et a donc été rejetée dans un premier temps. Il est cependant évident que rien n'est terminé, dans la mesure où l'enjeu est la diffusion des articles de périodiques scientifiques sous forme électronique, et que la normalisation ou l'élaboration de rapports techniques sont nécessaires à l'émergence du marché des produits électroniques, dans le domaine où l'on attend la plus grosse part de marché concurrent du papier pour l'an 2000.

Les nouvelles données de la bibliothéconomie vont donc être influencées par la révolution électronique qui touche l'essence même du document, donc la nature de sa diffusion et de sa consultation. Alors que la normalisation qui accompagne cette révolution se déroule de manière transversale aux secteurs traditionnels, il semble donc particulièrement important que les bibliothèques, centres de documentation, services bibliographiques, banques de données participent activement à cette démarche générale et profitent des méthodologies de modélisation mises en œuvre, sans se limiter à leur domaine traditionnel de normalisation. Le partage des rôles qui est en train de se faire passer, entre autres, par la participation active à la normalisation dans un contexte global.

Mai 1993

BIBLIOGRAPHIE

1. Commission des Communautés européennes, DG XIII, *New opportunities for publishers in the information services market : executive summary*, Janv. 1993, 28 p. (Report EUR 14925 EN).

2. *Le Document électronique* / Cours organisé par l'Institut national de la recherche en informatique et en automatique, 11 au 15 juin 1990, Châtelailon, Le Chesnay, INRIA, 1990, 216 p., ISBN 2-7261-0619-6.

3. MacKinnon (Dennis), Mc Crum (William), Sheppard (Donald), *Introduction à l'OSI : concepts, normes et applications*, Paris, AFNOR, 1991, 254 p., ISBN 2-12-480411-1.

4. *Multimédia et normalisation : conférence internationale sur les normes de codage de l'image et des sons numérisés et leurs applications*, 28-29 janvier 1992, Paris, AFNOR, 1992.

5. *Normative : la normalisation des technologies de l'information*, Paris, AFNOR, Mensuel, ISSN 0299-8521.

6. *SGML-ODA : Présentation des concepts et comparaison fonctionnelle*, AFNOR, 1991, ISBN 2-12-488011-X.