



World Library and Information Congress: 71th IFLA General Conference and Council

"Libraries - A voyage of discovery"

August 14th - 18th 2005, Oslo, Norway

Conference Programme:

<http://www.ifla.org/IV/ifla71/Programme.htm>

July 11, 2005

Code Number:

034-F

Meeting:

150 SI - ICABS (IFLA/CDNL Alliance for Bibliographic Standards)

Nestor et Kopal – deux projets nationaux pour garantir un accès à long terme aux documents électroniques en Allemagne

Reinhard Altenhöner

Responsable du Service Informatique

Die Deutsche Bibliothek

Allemagne

Traduction : Nathalie Yakovleff (HFA Documentation Texte, France)

Résumé

Le développement du marché des publications en direction des publications électroniques via des réseaux change de manière significative la définition de la mission des bibliothèques - et tout particulièrement celle des bibliothèques qui ont traditionnellement la charge de l'héritage culturel d'une nation. Cette situation est particulièrement importante en Allemagne car beaucoup de bibliothèques différentes y ont été chargées de la collecte, de l'indexation et de l'archivage des publications. Cette mission doit être redéfinie dans ses aspects organisationnels et techniques pour s'adapter à un univers de publications en changement. Il y a, toutefois, une communauté d'intérêts pour comprendre que toutes ces activités doivent être entreprises dans un effort coopératif.

Quelques uns des principaux efforts actuels de l'Allemagne sont présentés dans cet exposé: alors que, d'un côté, l'accent est mis sur une clarification du cadre légal pour le nouveau spectre de missions, de l'autre, deux projets sont mis en place. Le premier est consacré à la construction d'un réseau de compétences inter-institutions sur les aspects de la conservation à long terme (nestor), le second est consacré au développement coopératif d'un système de dépôt pour les ressources électroniques (kopal). Die Deutsche Bibliothek est impliquée au premier plan dans ces deux directions présentées ici.

1 – Le contexte général

Le processus de publication devient de plus en plus électronique. Et, pour l'instant, nous ne savons pas encore comment les publications électroniques, les pièces d'art, les images et les documents sonores, les fichiers, les données primaires et la musique doivent être archivés pour être lus de manière permanente et ainsi accessibles à tous dans le futur. Ce que nous pouvons dire, c'est quelle que soit la stratégie suivie dans le futur pour donner accès à un contenu numérique, il va dépendre de l'existence d'une suite de bits dont l'intégrité et l'authenticité auront été préservées au cours des années. A partir de la conservation d'un code bit la conversion d'un document devra être rendue possible pour permettre un futur accès aux objets numériques. Plusieurs stratégies sont envisagées et elles peuvent se résumer en deux types d'actions possibles :

- soit migrer les objets numériques dans un environnement contrôlé
- soit essayer de reproduire l'environnement du système au moment de la création de l'objet, y compris l'émulation du matériel informatique et des logiciels

Avec la disponibilité d'un code bit nous avons besoin en même temps d'informations supplémentaires contenues dans des métadonnées concernant en particulier une information technique à propos des objets stockés, leur environnement technique originel et l'historique de leur stockage. La mission pour construire un système et une infrastructure pour une préservation à long terme est complexe et requiert de gros financements et des capacités humaines et nous avons besoin à la fois d'un réseau d'échange et de connaissances - national aussi bien qu'international.

2 - Le contexte de travail de Die Deutsche Bibliothek

Die Deutsche Bibliothek¹ est la bibliothèque nationale et le centre pour la bibliographie nationale de la République Fédérale d'Allemagne. La bibliothèque est responsable de la collecte, du traitement et de l'indexation bibliographique de toutes les publications allemandes et de langue allemande depuis 1913. Cette mission se fonde sur un mandat statutaire pour la collecte, le traitement bibliographique et la conservation à long terme de toutes les publications parues en Allemagne ou publiées à l'étranger en langue allemande. La législation prend aussi en considération les publications numériques distribuées sous un support physique mais elle n'a pas encore prévu le cas des publications en ligne. Aujourd'hui, nous attendons de nouvelles dispositions légales qui étendront notre domaine de responsabilités à tous les types de publications du net et numériques. En conséquence de la nouvelle réglementation, nous espérons

¹ Die Deutsche Bibliothek : <http://www.ddb.de>

que beaucoup de bibliothèques régionales ayant un mandat de dépôt légal pour des régions particulières vont avoir un nouveau cadre légal pour leurs collections numériques.

Nous devons nous préparer à cette situation. C'est pourquoi ces dernières années Die Deutsche Bibliothek a lancé plusieurs projets pour promouvoir la conservation à long terme des publications numériques en Allemagne. Un certain nombre de principes de base applicables aux collections en ligne ont été définis dans des commissions préliminaires en concertation avec des éditeurs, des experts en bibliothéconomie, des spécialistes en Sciences de l'Information et des représentants du gouvernement et formulés dans un document de principe visé par le comité des éditeurs du Börsenverein des Deutschen Buchhandels en juin 1997 :

- Toutes les publications en ligne doivent être fournies sur demande via un réseau de données ou via un media physique de données
- Les publications en ligne disponibles sous différentes formes doivent être fournies dans le format demandé par la bibliothèque
- Les publications qui sont diffusées sous un support matériel et en ligne et dont le contenu est identique doivent être soumises dans les deux formats
- Les publications en ligne diffusées simultanément par plusieurs fournisseurs de contenu ne doivent être remises qu'une seule fois

A partir de ces règles, Die Deutsche Bibliothek a testé ces dernières années des procédures pour l'enregistrement, la collecte et la préservation à long terme des publications en ligne en coopération avec des éditeurs et des producteurs d'information dans une phase de test. Le groupe de travail "Bibliothèque de Dépôt Electronique" a examiné et déterminé les conditions nécessaires pour que Die Deutsche Bibliothek devienne une bibliothèque de dépôt pour les publications en ligne tout comme :

Depuis 1998, les mémoires et les thèses en lignes (36 000 jusqu'à aujourd'hui) ont été collectés, archivés et rendus disponibles sur un serveur. Depuis 2000 les périodiques électroniques ont été collectés et depuis 2001 Die Deutsche Bibliothek a mis au point une interface d'enregistrement. Pour la procédure d'enregistrement, Die Deutsche Bibliothek a requis des métadonnées de gestion pertinentes à des fins de conservation. Un compromis a dû être fait entre les éditeurs qui veulent baser la fourniture des travaux sur la base du volontariat et les exigences du futur processus de conservation dans le système de dépôt. De plus, Die Deutsche Bibliothek a participé au projet européen Nedlib, au cours duquel le modèle OAIS a été adopté et où des propositions sur la gestion de flux pour l'intégration des procédures des bibliothèques ont été faites.

Autres expériences significatives :

- Die Deutsche Bibliothek a acquis de l'expérience avec son système d'Accès Multimedia (Multimedia - Bereitstellungssystem MMB). MMB rend possible le stockage et l'accès pour des objets numérisés sur des supports matériels. Différents types d'objets (station de travail image, application de kit d'installation, collections de fichiers, présentation d'objets) ont été mis en place pour permettre l'interprétation des objets numériques complexes (applications).

- Une autre partie de son activité couvre le développement d'une infrastructure pour un identifiant permanent en Allemagne. L'utilisation d'identifiants permanents garantit la stabilité et la possibilité sans équivoque d'accéder à un objet numérique sur le net².

3 - La structure fédérale de l'Allemagne.

L'Allemagne a une structure fédérale où les états disposent de larges marges décisionnelles dans leur administration en ce qui concerne en particulier le système éducatif et les secteurs de la science et de la recherche. L'existence de plusieurs bibliothèques régionales responsables du dépôt légal pour leur région est également un autre élément de la structure fédérale. Si l'on considère l'importance du travail de conservation à long terme à l'intérieur de la structure fédérale de l'Allemagne, il est évident que la réussite du projet passe par une approche coopérative. Ceci s'applique d'abord à l'aspect organisationnel, mais il y a beaucoup de raisons pratiques et techniques pour lesquelles nous essayons de répartir la responsabilité de la collecte des objets numériques sur le net.

De là, les partenaires dont nous avons besoin pour la mise en place d'une stratégie de conservation à long terme au niveau de l'infrastructure doivent couvrir le domaine organisationnel tout comme le domaine opérationnel.

4 - Les initiatives nationales en Allemagne

Il y a deux initiatives qui ont trait au problème de la conservation à long terme :

- Sur le plan organisationnel, il y a nestor (Network of Expertise in Long-Term Storage of Digital Ressources : Réseau de compétences dans le stockage à long terme de ressources électroniques) qui a été créé dans le but de construire une plateforme de compétences pour échanger de la connaissance et partager des retours d'expérience dans le domaine de la conservation à long terme et pour échanger les experts et les expertises entre différents types d'institutions gérant un patrimoine culturel.
- Le projet kopal (Kooperativer Aufbau eines Langzeitarchivs digitaler Informationen : Constitution Coopérative d'Archives d'Informations numériques à long terme), sur le plan opérationnel, a été créé pour construire un système d'archives fonctionnant à long terme et basé sur OAIS.

5 - Nestor

Financé par le ministère fédéral de l'Education et de la Recherche, nestor, l'alliance pour une mémoire numérique de l'Allemagne, a commencé en 2003 et prendra fin en 2006. Sous la

² Ainsi intégrés au projet "EPICUR – « Amélioration des Services d'Identifiant Permanent - Méthode globale pour une Identification non équivoque de Ressource"), les identifiants permanents deviennent une donnée à part entière dans le schéma de métadonnées pour les publications électroniques.

direction de Die Deutsche Bibliothek, les partenaires de nestor viennent du domaine des bibliothèques (la Bibliothèque d'Etat de Bavière, la Bibliothèque d'Etat et de l'Université de Basse-Saxe) d'un centre de media (Service d'Informatique et de media de l'Université d'Humboldt, Berlin) des archives (Archives de l'Etat de Bavière - Direction Générale) et d'un musée (Institut de Science des Musées, Berlin). A ceci s'ajoute un Conseil Consultatif de nestor composé d'éditeurs, de représentants des sciences & technologies, de musées, d'archives, de bibliothèques et d'universités ainsi que de membres d'institutions culturelles & politiques et d'institutions de recherche/centres informatiques.

Le but du projet Nestor est de rassembler la connaissance existante, les acteurs et les compétences sur le stockage à long terme des ressources électroniques afin de constituer une future alliance pour la mémoire électronique de l'Allemagne.

Ainsi, avec ce projet, nous :

- créons un réseau pour l'information et la communication sur les activités de conservation à long terme dans le présent et dans le futur.
- construisons une communauté inter-secteur pour encourager et soutenir les actions de conservation à long terme et pour susciter une prise de conscience dans la société.
- encourageons des synergies entre les activités en cours en Allemagne et coopérons avec des projets de partenaires internationaux.
- développons des stratégies pour la coordination des activités de conservation à long terme en Allemagne.
- proposons un modèle organisationnel à longue échéance pour poursuivre son utilisation comme un réseau d'excellence après la fin du projet nestor en 2006

Le détail des tâches est le suivant :

- Collecter et présenter l'information
- Consolider les domaines d'expertise, les rendre visibles et disponibles
- Promouvoir la coopération et soutenir une solution commune
- Préparer des documents sur les aspects techniques, organisationnels et légaux
- Présenter des modèles, les discuter et encourager la diffusion des meilleures pratiques
- Encourager l'utilisation des standards et représenter l'Allemagne dans les comités internationaux de standardisation
- Mettre au point des lignes directrices pour la collecte et des procédures de sélection pour le stockage des sources numériques
- Coordonner les responsabilités pour les tâches s'étalant sur le long terme
- Susciter la prise de conscience du problème parmi les cercles de spécialistes en documentation et dans le grand public
- Préparer la création d'un organisme permanent coordonnant et représentant les préoccupations du stockage à long terme

Les groupes de travail et la compétence métier représentent une part non négligeable de nestor. A ce jour les groupes de travail suivant sont actifs :

- le groupe de travail nestor Fiabilité des Archives Certification
- le groupe de travail nestor Archivage des objets Multimedia

- le groupe de travail nector Politiques de Conservation et Critères de Sélection (en préparation)

Les rapports des experts traitent des sujets suivants :

1. Le stockage à long terme des journaux électroniques
2. L'analyse des besoins en ce qui concerne le stockage à long terme des objets multimedia
3. L'élaboration d'un schéma de description pour une stratégie de conservation à long terme (politique de conservation)
4. La conservation électronique à long terme et la législation
5. L'étude de l'état des données de recherche existantes et des données brutes des activités scientifiques
6. Une comparaison des systèmes d'archivage existants
7. La numérisation et la conservation des matériaux numérisés dans les musées allemands

Certains de ces rapports ont déjà été publiés et sont consultables sur le web.

6 - kopal

Kopal traite de la conservation à long terme avec une approche orientée vers la pratique, dans un esprit collaboratif, d'un point de vue technique et organisationnel : soutenu financièrement par le Ministère Fédéral de l'Education et de la Recherche, kopal a pour objectif de développer une solution techniquement innovante qui consisterait en un système d'archivage pérenne à long terme pour des données électroniques. La solution est basée sur DIAS (Digital Information and Archiving System : Information Numérique et Système d'Archivage) développé conjointement par IBM et la Bibliothèque Nationale des Pays-Bas à La Hague.

Dans le projet qui a démarré en juillet 2004, deux des partenaires, à savoir Die Deutsche Bibliothek et la Bibliothèque d'Etat et de l'Université de Basse Saxe (SUB Göttingen) vont enregistrer des matériaux de toutes sortes via des processus de masse dans le système d'archivage à long terme. Cela va de documents électroniques aux formats PDF, TIFF ou TeX à des objets complexes comme des vidéos numériques.

Le système va être mis en place selon les standards internationaux de l'archivage à long terme et avec des métadonnées selon le modèle OAIS (Systèmes Ouverts d'Archivage). La transparence de l'intégration dans les bibliothèques existantes et les systèmes d'information est garantie et est un objectif fondamental du projet.

Afin d'aboutir à une solution ouverte, pouvant être mise à jour et appropriée, plusieurs types de partenaires ont été intégrés au projet. Partant du point de vue que les motivations sont différentes, il a été décidé de séparer les aspects fonctionnement et développement. Par exemple, notre partenaire pour le système d'exploitation (le centre informatique partenaire Gesellschaft für Wissenschaftliche Datenverarbeitung Göttingen (GWDG)) veut acquérir de l'expérience dans un système fiable et évolutif, afin de conquérir de nouveaux partenaires qui bénéficieront des apports du système pour leurs propres besoins. D'un autre côté, il y a une plateforme neutre où les partenaires chargés du développement (Die Deutsche Bibliothek et SUB Göttingen) ont à trouver des solutions communes en matière de logiciels et de procédures de conservation. De plus, le

système doit être capable de traiter avec de multiples utilisateurs tout en garantissant l'indépendance de chaque partenaire.

En ce qui concerne l'architecture logicielle, une séparation a été effectuée entre les fonctionnalités qui sont au coeur du système d'archivage et les outils connexes qui traitent de l'homogénéisation et du transfert des objets numériques à l'intérieur du système. L'isolation d'une fonctionnalité centrale (DIAS Core) requiert des interfaces bien définies et une libre utilisation pour les futurs partenaires : le but est de construire des formulaires d'enregistrement pour des paquets d'information (SIPs) et de les importer dans le système. Ceci va être développé par les partenaires avec une méthode utilisant une licence open source (GPL). Pour le système de présentation des objets nous suivons la même règle : nous voulons devenir indépendants des solutions propriétaires et préférons des interfaces bien maîtrisées, ouvertes pour d'autres partenaires et systèmes.

Le partenaire pour le DIAS Core est IBM Deutschland GmbH. Ceci va garantir un choix professionnel de composants logiciels et va assurer un support permanent sur le long terme.

Ainsi, dès le départ, kopal a intégré plusieurs partenaires en différents endroits. De cette manière, dans le futur, cette archive à long terme pour l'information électronique va donner à d'autres institutions la possibilité de garder leurs données disponibles sur le long terme. En conséquence, kopal donne l'occasion à des universitaires, des entreprises et des administrations de profiter de son utilisation au-delà de celui réservé aux bibliothèques.

Il y a d'importants éléments complémentaires au système DIAS tel qu'il existe aujourd'hui, à savoir :

- l'élaboration de fonctionnalités de contrôle pour préparer la conservation à long terme des documents électroniques.
- des fonctionnalités souples d'import/export de données basées sur le modèle de description des objets (METS Metadata Encoding and Transmission Standard : codage de métadonnées et standard d'échange de la bibliothèque du Congrès, USA) et LMER (métadonnées d'archivage à long terme pour les ressources électroniques : Langzeitarchivierungsmetadaten für elektronische Ressourcen)³.
- la possibilité de réutilisation pour les institutions intéressées.

Pour les raisons énoncées ci-dessus, beaucoup d'efforts sont déployés pour définir un modèle de données et une architecture qui soient puissants dans ce système. Tout au long du projet, plusieurs conférences d'information et ateliers sont prévus à un niveau national comme international.

³ <http://www.loc.gov/standards/mets> et <http://www.ddb.de/standards/pdf/lmer12.pdf>