



World Library and Information Congress: 70th IFLA General Conference and Council

22-27 August 2004
Buenos Aires, Argentina

Programme: <http://www.ifla.org/IV/ifla70/prog04.htm>

Code Number: 025-F
Meeting: 148. IFLA-CDNL Alliance for Bibliographic Standards (ICABS)
Simultaneous Interpretation: -

Une exploration des stratégies et des développements récents dans le domaine de la préservation numérique : un aperçu d'une enquête internationale menée par la *Koninklijke Bibliotheek*

Hilde van Wijngaarden

Digital Preservation Officer
National Library of the Netherlands
The Hague, Netherlands

Résumé

Cet article décrit l'enquête internationale que la Bibliothèque Nationale des Pays-Bas (Koninklijke Bibliotheek ou KB) réalisera pour l'ICABS. L'enquête s'intéressera aux exigences/conditions pour l'archivage à long terme des ressources électroniques et explorera les stratégies, les méthodes et les normes relatives à la migration et à l'émulation. L'étude des deux sujets portera surtout sur les questions qui sont les plus pertinentes pour le milieu des bibliothèques. À partir d'une recherche documentaire, d'entrevues et de l'analyse des nouveaux développements, un rapport sera rédigé rendant compte des meilleures pratiques, de l'évaluation des normes émergentes et des recommandations pour des recherches et développement futurs. Après plusieurs années à débattre du problème et de ses solutions possibles, les dernières années ont vu l'augmentation de mesures qui ont été prises pour constituer les archives numériques, organiser la gestion des ressources numériques et définir des stratégies pour l'accès permanent. Il est temps désormais de faire l'inventaire et de définir les meilleures pratiques, d'évaluer les normes émergentes et de décrire les questions qui nécessitent davantage d'attention.

Introduction

Dans l'énoncé de la mission du plan stratégique de l'ICABS, il est mentionné que l'alliance ICABS a pour but de « faire progresser la compréhension des questions relatives à l'archivage à long terme des ressources électroniques, incluant la promotion de nouvelles conventions

recommandées pour un tel archivage. »¹ L'intérêt de l'alliance est stratégique, avec une insistance particulière sur les moyens pratiques pour améliorer la coordination internationale et diriger des développements. La Bibliothèque Nationale des Pays-Bas (Koninklijke Bibliotheek ou KB) s'est elle-même impliquée dans les actions suivantes :

- Explorer les exigences/conditions pour l'archivage à long terme des ressources électroniques.
- Explorer et promouvoir les stratégies, les méthodes et les normes pour la migration et l'émulation.

Ces actions seront réalisées à partir d'une recherche documentaire, d'entrevues et de l'analyse des nouveaux développements et résulteront dans un rapport contenant les meilleures pratiques, l'évaluation des normes émergentes et des recommandations pour la recherche et le développement futurs. Les résultats préliminaires et une première version des conclusions seront discutés avec la Bibliothèque Nationale d'Australie [BNA] et les autres partenaires de l'ICABS. Au sein de l'ICABS, la BNA s'intéresse à deux sujets étroitement reliés. Dans un premier temps, ils exploreront et promouvoir des méthodes d'archivage des publications électroniques récoltées sur le web. Ils formuleront également des directives et des normes relatives à la préservation de matériels numériques s'intéressant au rôle du portail thématique PADI et sur les développements dans l'Asie du Sud-Est. La BNA et la KB demeureront étroitement en contact et partageront leurs résultats.

Au moment d'écrire cet article, l'enquête internationale est à ses débuts. Le canevas de l'étude a été produit et la recherche vient à peine de commencer. Je me concentrerai donc sur les questions et le raisonnement à l'origine de la conception de cette enquête.

Deux domaines : l'archivage numérique et l'accès permanent

La préservation numérique se compose de deux étapes : un entreposage sécuritaire et un accès permanent. Pour assurer la conservation sécuritaire des objets numériques, il est nécessaire d'avoir une archive numérique dédiée à la maintenance à long terme. Une telle archive préserve le flot de données ainsi que son format et les métadonnées sont enregistrées pour permettre la description et le repérage. Des procédures détaillées doivent être produites et appliquées pour la sauvegarde et la mise à jour du matériel informatique et des supports. La restitution future des articles numériques entreposés constitue un autre problème. Comment pouvons-nous nous assurer de pouvoir consulter et exécuter des objets numériques d'ici dix, cinquante voire une centaine d'années ? Des stratégies telles que la migration, la conversion, la normalisation et l'émulation doivent être considérées et développées. Nous avons également besoin d'outils permettant d'assurer cet accès permanent.

Depuis ces dernières années le problème de la préservation numérique est reconnu et pris en charge par les institutions de patrimoine culturel et les gouvernements. Après avoir débattu pendant des années du problème et de ses solutions possibles, des mesures ont été prises pour constituer des archives numériques, pour organiser la gestion des ressources numériques et pour définir des stratégies pour l'accès permanent. La coopération, le partage des connaissances et la normalisation ont été appuyés par plusieurs projets innovateurs internationaux.

¹ Notre traduction.

Après une recherche théorique initiale débouchant sur des tests, des études pilotes et des prototypes, le développement d'archives numériques a commencé et plusieurs systèmes dédiés à l'entreposage à long terme des ressources numériques sont maintenant opérationnels. Pendant que le développement de « lieux sécuritaires » pour entreposer les ressources électroniques est bien entamé, R&D dans la technologie à accès permanent est seulement à ses débuts. Les outils et les méthodes pour l'accès permanent requièrent une recherche technologique intensive, et des essais pratiques dans la milieu de l'archivage numérique. Des tests approfondis sont conduits sur les façons de repérer des formats de fichiers obsolètes et de définir des procédures pour prévenir les problèmes futurs, mais en ce qui concerne le développement pratique d'outils, les tests sont plus rares. Le défi technologique est de taille, et jusqu'à présent, les concepteurs de logiciels commerciaux ont été réticents à entreprendre un travail qui ne rapporte pas de profits immédiats.

Les deux questions, l'archivage numérique et le problème de la restitution continue, ont maintenant atteint un niveau qui nous permet de faire l'inventaire. Quels sont les résultats des projets en cours dans ce domaine, quels problèmes apparaissent, en quoi la normalisation a contribué au développement pratique et quels sont les domaines qui nécessitent une attention supplémentaire ? Une évaluation peut faire ressortir les meilleures pratiques et aborder les questions sensibles. Dans cette perspective, l'enquête ICABS de la KB s'intéressera particulièrement à la préservation numérique dans un environnement de bibliothèque.

Publications, vues d'ensemble et sondages

Parallèlement à l'intérêt et aux développements croissants, les publications relatives à la préservation numérique ont augmenté de façon majeure. L'information était habituellement dispersée et difficile à trouver, mais ce n'est désormais plus le cas. Les périodiques consacrés aux sujets reliés aux bibliothèques numériques fournissent une plateforme pour les rapports sur des tests et des développements en cours et sur des analyses de recherche des problèmes techniques ou infrastructurels. Des mémoires et des études ont été publiés relatifs à des sujets spécifiques dans le domaine de la préservation numérique.

En 2002 et 2003 plusieurs vues d'ensemble majeures concernant « l'état de l'art » dans le domaine de la préservation numérique ont été publiées. Les personnes familières avec le sujet connaîtront par exemple les Directives de l'Unesco par Colin Webb et le programme NDIIPP incluant l'étude de Neil Beagrie. [1,2] Des groupes de travail internationaux comme le Groupe de travail sur l'archivage et la préservation numérique NSF-DELOS, le Groupe de travail sur les attributs de l'archive numérique RLG, et le Groupe de travail sur les métadonnées de préservation OCLC-RLG ont débattu et publié sur des sujets spécifiques, fournissant un aperçu international sur les développements majeurs. [3,4,5] D'autres publications importantes destinées aux organisations et aux gestionnaires, fournissent des directives et des recommandations sur la manière de gérer leurs ressources numériques.

Un aperçu direct du travail qui est conduit à l'échelle mondiale est devenu possible à travers la réalisation de plusieurs sondages et des centres de connaissances approfondis et à jour entretiennent des sites web qui fournissent un accès structuré à diverses publications et projets en cours de réalisation, comme PADI, Erpanet et DPC.

Développements récents

À mesure que l'activité croît, différentes approches apparaissent clairement. Les différences sont relatives à l'architecture des systèmes d'archivage, à la gestion pratique des soumissions, aux formats utilisés pour l'entreposage et aux stratégies pour l'accès permanent. Le modèle de

référence OAIS a été accepté à l'échelle mondiale et fournit un premier départ bien défini pour la conception et la structure des archives numériques. Alors que la conception de beaucoup d'archives numériques est basée sur le modèle de référence OAIS, la façon dont le modèle est interprété varie considérablement. Il y a également des institutions qui n'ont pas conçu leurs systèmes selon l'OAIS : soit parce qu'elles ont commencé le développement avant que la version finale de l'OAIS soit publiée, soit parce qu'elles n'ont pas pris connaissance du modèle, ou soit parce qu'elles ont choisi une autre approche. Cela dépend beaucoup de la mission, du rôle et responsabilité de l'institution d'archivage, des types d'objets à entreposer et de la manière dont les services sont organisés.

Les institutions du patrimoine culturel, avec les bibliothèques et les archives en tête, ont été parmi les premières à constituer des archives numériques pour la conservation à long terme des ressources numériques. Les institutions de recherche scientifique et technique et plus précisément les organisations de recherche spatiale ont également pris très tôt des dispositions à ce sujet. La comparaison des approches montre maintenant des différences entre ces différents types d'organisations. Cela nous permet une analyse plus détaillée des approches et des défis spécifiques aux bibliothèques : le point central de l'étude ICABS de la KB.

En ce qui concerne l'accès permanent, deux approches apparaissent clairement. La première consiste simplement à tout entreposer dans la mesure du possible et à laisser les futurs usagers le soin de s'inquiéter de la restitution. L'autre approche consiste à élaborer et à tester dès maintenant des stratégies et des outils pratiques. Bien entendu, ces approches sont extrêmes. La plupart des institutions d'archivage prennent une position intermédiaire. En choisissant soigneusement les formats qu'elles entreposent, en privilégiant les formats normalisés et dont l'usage est répandu, en produisant des descriptions détaillées des exigences techniques des objets entreposés, en collectant l'information sur les conséquences des développements techniques et en entreposant les exemplaires originaux avec les copies converties, les archives numériques espèrent fournir l'accès aux futurs usagers.

L'enquête internationale de la KB

L'étude KB-ICABS ne sera pas un autre sondage ou vue d'ensemble, mais sera basée sur les publications récentes, l'information sur les projets en cours et les résultats des sondages. Ce qui signifie : faire le résumé des principales conclusions des études, combiner ces dernières avec l'information pratique sur le travail en cours dans les bibliothèques et insister sur les domaines spécifiques tels que les stratégies et la normalisation dans les bibliothèques.

Une sélection sera faite des institutions mondiales de pointe dans le domaine de l'archivage numérique et les responsables seront contactés pour répondre à des questions spécifiques. De cette manière, les développements en cours et les histoires internes des projets de développement pourront être documentés. On demandera aux spécialistes de formuler des recommandations destinées aux bibliothèques planifiant la construction d'entrepôts numériques.

Les sujets qui seront traités peuvent être groupés en quatre catégories :

1. L'architecture des archives numériques

Le développement des archives numériques a montré à la fois des différences et des similitudes dans l'architecture. Différents modèles seront comparés, les

projets seront évalués et les expériences seront décrites. Il en résultera une définition des meilleures pratiques.

2. *Les métadonnées de préservation et l'enregistrement du format*

Plusieurs nouvelles initiatives s'intéressent à la définition et à la création des métadonnées de préservation. L'enregistrement de l'information sera crucial et la coopération internationale devrait conduire à des schémas qui faciliteront la création des métadonnées de préservation. Les initiatives pour réaliser les enregistrements de format contenant de l'information technique sur la maintenance et le restitution des formats numériques sont d'une grande importance. L'enquête ne discutera pas en détail des métadonnées de préservation mais présentera les initiatives en cours qui sont pertinentes à l'architecture des entrepôts numériques et à l'enregistrement technique du format de l'information permettant une restitution future.

3. *Les normes émergentes*

La coopération internationale au travers des groupes de travail et des initiatives de certification a travaillé sur la définition de modèles et de normes. De quelle façon ces efforts ont contribué au développement et à l'architecture des archives numériques ? En tout cas, cette coopération a conduit à un vocabulaire commun qui nous permet de communiquer au niveau mondial et de comparer différentes approches. À l'aide d'entrevues nous espérons fournir un aperçu sur la façon dont les normes et protocoles sont utilisés en pratique et quels seraient les domaines qui profiteraient de ces nouveaux efforts de normalisation.

4. *La recherche et le développement d'outils d'accès permanent*

Le développement d'outils et de méthodes qui soutiennent les stratégies d'accès permanent est souvent effectué dans les instituts de recherche technique, même si les projets peuvent être initiés par les bibliothèques ou d'autres institutions du patrimoine culturel. Le développement d'outils requiert quelque fois une programmation intensive et un aperçu détaillé des applications et des technologies de l'information. Une connaissance à jour des innovations technologiques les plus récentes et celles à venir est nécessaire. Pour les bibliothèques, il est important de travailler étroitement avec les ingénieurs techniques pour se concentrer sur les développements de la maintenance à long terme et plus spécialement sur les questions d'accès permanent ; l'expertise des professionnels de l'information. Durant l'évaluation des activités d'archivage numérique, les organisations qui travaillent également sur la question de l'accès permanent seront sélectionnées. Une étude sélective des expériences et des attentes concernant les stratégies de restitution suivra. Puisqu'il s'agit d'un sujet qui implique une connaissance technique spécifique qui dépasserait les limites de cette étude, l'intérêt pour les recherches sur la technologie d'accès permanent se limitera à l'utilisation et aux possibilités pratiques pour les bibliothèques.

Les résultats de l'enquête seront présentés dans un rapport visant à servir de document d'aide à la compréhension des stratégies de préservation numériques et de support à l'implantation de politiques. Une première version des résultats de l'enquête sera envoyée aux partenaires de l'ICABS. Après un processus de lecture on déterminera sur la façon dont ce rapport sera rendu accessible.

Références

- [1] *Guidelines for the preservation of digital heritage*, UNESCO, March 2003.
- [2] Neil Beagrie, National Digital Preservation Initiatives: An overview of developments in Australia, France, the Netherlands and the United Kingdom and related international activity, Appendix 5, *Preserving our digital heritage: Plan for the National Digital Information Infrastructure and Preservation Program and the NDIIPP Plan Appendix*, Library of Congress, October 2002.
- [3] *Invest to Save*, Report and Recommendations of the NSF-DELOS Working Group on Digital Archiving and Preservation, 2003.
- [4] *Preservation Metadata for Digital Objects: A Review of the State of the Art*, White Paper, OCLC-RLG preservation metadata working group, January 2001.
- [5] *Trusted Digital Repositories: Attributes and Responsibilities*, RLG/OCLC, Mountain View, VS, May 2002.