



68th IFLA Council and General Conference

August 18-24, 2002

Code Number: 066-152-F
Division Number: IV
Professional Group: Bibliography - Workshop
Joint Meeting with: National Libraries
Meeting Number: 152
Simultaneous Interpretation: -

Une stratégie à facettes multiples pour une bibliographie nationale des ressources électroniques

Randi Diget Hansen

Centre Bibliographique Danois
Copenhague, Danemark

La situation au Danemark

Mise en place d'une stratégie nationale au Danemark pour le recensement des publications Internet.
Différents objectifs sont visés :

- élaborer un modèle de coopération entre différents types de services Internet dans le monde des bibliothèques et l'Agence bibliographique nationale - y compris la définition d'une base de données pour le partage des ressources, la création d'un format d'échange, le développement de critères de sélection et d'autres types d'outils et de règles susceptibles de faciliter le travail.
- convaincre les autorités concernées et les professionnels dans le monde des bibliothèques que le contrôle bibliographique d'Internet est nécessaire et réalisable et que la réutilisation et la coopération sont les seules façons d'y arriver.

Bibliographie nationale et dépôt légal des publications Internet au Danemark

En 1996, le Centre bibliographique danois, qui produit la bibliographie nationale danoise, a pris l'initiative de lancer un projet pour déterminer si les publications Internet pouvaient ou devaient faire l'objet d'un contrôle bibliographique de la même façon que les publications imprimées et électroniques ayant un support physique stable. La raison en était que nous avions l'impression que les moteurs de recherche sur Internet souffraient du problème général de la recherche sur des données non "qualifiées" et renvoyaient le plus souvent une trop grande quantité de données. Nous avions aussi le sentiment que les données contenues dans les publications Internet n'étaient pas nécessairement différentes de celles contenues dans les publications ayant un support physique stable, et qu'exclure du contrôle bibliographique les publications Internet c'était prendre le risque que beaucoup de gens aient

du mal à accéder à un nombre croissant d'informations dont les citoyens ont besoin par opposition [à la facilité d'accès] aux informations figurant dans les produits ayant un support physique stable.

A la question de savoir s'il était intéressant de recenser des publications qui peuvent ensuite disparaître, nous avons répondu qu'il était important de se focaliser sur la "qualité" des publications recensées et avons été convaincus que les publications Internet tendront à devenir de plus en plus pérennes.

Parallèlement une nouvelle loi sur le dépôt légal a été préparée et a pris effet au 1^{er} janvier 1998. Cette loi concerne tous les médias d'information y compris les documents numériques présents dans les bases de données - bien que limités aux publications pérennes ("statiques"), c'est-à-dire aux publications qui ressemblent à des livres ou qui se substituent aux livres.

La nouvelle loi de dépôt légal a eu pour conséquence que le Danemark publie depuis 1998 une bibliographie nationale recensant les publications Internet pérennes.

Depuis le commencement nous avons la conviction qu'il ne s'agit pas de recenser des millions et des millions de publications mais que le nombre de publications Internet "qualifiées" susceptibles de répondre aux critères de sélection de la bibliographie nationale n'est pas si important et que la tâche est faisable si elle repose sur la coopération entre les producteurs et les autres acteurs des réseaux au sein de la communauté des bibliothèques.

Nous sommes aussi arrivés à la conclusion que l'idée selon laquelle les publications étaient éphémères - nous savons que la durée moyenne de vie d'un document sur Internet est de 42 jours - ne s'appliquait pas aux publications "qualifiées".

Enfin, nous avons acquit la conviction que la lourde tâche d'organiser l'information "qualifiée" sur Internet allait exiger un processus permanent de développement. Il n'est pas réaliste de penser que la communauté des bibliothèques a trouvé en moins de 6 ou 7 ans les modèles définitifs de gestion des publications Internet.

Que s'est-il passé depuis 1998 - pourquoi avoir défini une stratégie nationale ?

La loi sur le dépôt légal et la bibliographie nationale se limite toujours à la prise en compte des publications "stables". D'un autre côté un certain nombre de recensements sont faits dans la communauté des bibliothèques. Les autorités concernées ont accordé des millions de couronnes danoises à des initiatives qu'on pourrait définir comme autant de projets, mais pas dans le but de recenser les publications "dynamiques" dans la bibliographie nationale.

Il y a quantité de projets dans le secteur des bibliothèques publiques aussi bien que dans le secteur des bibliothèques d'Etat ou universitaires. Il existe des portails dédiés aux enfants, à la musique, aux aliments, aux services de santé, etc., mais aussi plusieurs portails généraux.

Le Centre bibliographique danois a procédé à l'enregistrement d'environ 12000 publications danoises, "stables" et "dynamiques", et d'articles et de revues danois. Les autres services Internet dans le secteur des bibliothèques et financés par l'Etat danois, ont recensé 20000 ressources pertinentes en différentes langues en ayant recours à une intervention humaine. Malheureusement, en dépit de la taille du Danemark, les différents projets n'utilisent pas les mêmes techniques et il y a un nombre toujours croissant de doublons dans les recensements.

La raison pour laquelle l'Etat danois n'a pas voulu jusqu'à présent financer une extension de la bibliographie nationale est la crainte d'une explosion des quantités de documents à recenser.

La stratégie nationale porte sur les projets suivants :

- le projet "Statistiques" qui vise à mesurer la quantité de publications Internet danoises susceptibles d'être recensées dans la bibliographie nationale ;
- élaboration d'une base de données "Metanet", en Dublin-Core, où toutes les descriptions produites pour des publications Internet danoises pourraient être stockées ;
- élaboration d'un agent intelligent "Netowl", outil d'aide automatisée aux catalogueurs pour la constitution de la collection. On lui apprend comme à un enfant à découvrir ce qui est utile et à identifier ce qu'on ne veut pas ;
- des alliances stratégiques avec les producteurs, les bibliothèques et les autres institutions concernées ;

- faire des expériences sur différents niveaux de recensement et sur différents formats de description.

Le projet "Statistiques"

Le Centre bibliographique danois a lancé 2 projets pilotes à l'automne 2001 et au printemps 2002 pour mesurer la quantité de publications Internet danoises susceptibles d'être recensées, et tous les deux ont confirmé notre hypothèse selon laquelle la quantité de publications concernées est gérable.

Le projet repose sur un échantillon test de 2000 domaines sur un ensemble de 351792 domaines ".dk". Nous n'avons pris en compte que les domaines ".dk".

Sur les 2000 domaines de l'échantillon, 1469 ont été trouvés mais seulement 732 avaient un contenu. Les autres étaient vides, d'accès réservé, correspondaient à un renvoi automatique, correspondaient à des doublons d'URL, etc. Sur les 732 domaines ayant un réel contenu, seulement 46 avaient un contenu pertinent correspondant aux critères de recensement dans la bibliographie nationale. En tout 76 publications pertinentes ont été trouvées.

Si la tendance se maintient - et les 2 tests la confirme - cela signifie que pour le moment il n'y a dans le domaine ".dk" pas plus de 14000 publications susceptibles d'être recensées dans la bibliographie danoise et l'accroissement annuel est estimé à environ 4000 publications. Depuis l'année 2000, le taux d'accroissement des nouveaux domaines a légèrement diminué.

La base de données "Metanet"

La base de données Metanet, qui est en Dublin Core, est élaborée par le Centre bibliographique danois. Elle est conçue comme un réservoir commun pour tous les recensements danois. Elle est alimentée par le Centre bibliographique danois - en tant que principal fournisseur de données bibliographiques, à la fois pour les notices de la bibliographie nationale et pour le catalogage réalisé pour les bibliothèques - et par toutes les bibliothèques danoises participant au recensement des publications Internet, à la fois bibliothèques publiques et bibliothèques de recherche. Il est important de bien définir les contours de chaque produit, car des recensements en double ou en triple continuent de se faire, ce qui correspond à une mauvaise utilisation de ce qui est déjà disponible. Le but est de faire un recensement d'Internet réutilisable et rationnel. Il a aussi été nécessaire de définir un format d'échange basé sur le Dublin Core avec des composantes administratives compatibles avec les différents formats techniques utilisés par les services Internet. Vous le trouverez sur la page d'accueil du Centre bibliographique danois.

Le "Netowl"

Un des grands défis, sinon le plus grand, est de collecter et conserver les ressources d'Internet dignes d'apparaître dans la bibliographie nationale tout en exerçant une surveillance méthodique. Certains recensements sont fournis par les services du dépôt légal (www.pligtafleivering.dk) mais nous savons que l'habitude de remplir un formulaire de déclaration de dépôt légal n'est pas encore bien établie pour les producteurs de publications Internet. Et seulement les publications "stables" sont soumises au dépôt légal. Le "Netowl" est développé pour fournir une assistance automatisée dans ce travail de surveillance. C'est un moteur de recherche "éduqué" par les catalogueurs. Il repose sur l'algèbre de Boole. L'apprentissage vient juste de commencer, et tout le monde est optimiste.

Les alliances stratégiques

Une autre façon d'assurer une collecte plus large et plus méthodique des publications Internet est de contracter des alliances avec des producteurs, des bibliothèques et des institutions qui ont un intérêt à signaler les publications Internet à leurs usagers dans leur catalogue et qui veulent s'assurer de la pérennité des ressources signalées. Par exemple, la Bibliothèque administrative du Danemark est obligée de fournir et de conserver toutes les publications provenant des ministères, des départements et des agences du gouvernement. Elle vient tout juste de recevoir ce printemps le financement nécessaire à la conservation à long terme des publications. La bibliothèque va transmettre l'URL permanent à la

Bibliographie nationale, et va organiser une campagne méthodique de collecte ; l'Agence bibliographique nationale recensera ces publications. Le projet démarrera en septembre.

Expériences sur différents niveaux de recensement et sur différents formats de description.

Au Danemark on est obligé d'utiliser les Règles de catalogage danoises et le format danMarc2 dans la Bibliographie nationale. C'est stipulé dans la Convention de la Bibliographie nationale. Cela s'applique aux publications Internet. C'est pourquoi nous avons utilisé aussi ces règles et ce format quand nous avons commencé la bibliographie nationale dans ce domaine. Ce sont de bonnes règles et un bon format. Toutefois ils nous incitent à penser d'une manière très traditionnelle, et certains, particulièrement les non catalogueurs, pensent que cela prend beaucoup de temps et que cela revient cher. Par conséquent nous avons agi en toute liberté pour mener l'expérience sur un nouveau type de documents. Nous avons recruté des bibliothécaires sans aucune formation en catalogage et leur avons fait utiliser un niveau minimum du Dublin Core. Cela semble être efficace et la production donne satisfaction. Nous pensons que les gens se concentrent sur les difficultés de la sélection et du traitement matière et non sur la description.

Pourquoi définir une stratégie nationale pour le recensement des publications Internet ?

Pourquoi pensons-nous au Danemark qu'une bibliographie nationale recensant les publications Internet est nécessaire et réalisable ?

La Bibliographie nationale est la mémoire de la nation. C'est un devoir de s'assurer que le savoir publié et les expériences menées au Danemark sont conservés en lieu sûr et recensés pour permettre leur utilisation aujourd'hui et dans le futur. Ceci est une obligation quelle que soit la forme que prend l'édition nationale.

Les publications produites par des auteurs et des artistes danois, des sociétés, des organisations gouvernementales danoises, etc., et qui ne sont publiées que sur Internet, seront conservées en lieu sûr de façon systématique que si le travail est confié ou coordonné par une agence comme l'Agence bibliographique nationale qui a le devoir de procéder à un recensement méthodique et permanent.

Nous ne pouvons pas compter sur le fait que même les meilleurs moteurs de recherche pourraient faire un travail méthodique et prendre en charge l'obligation d'archivage et de mise à disposition. Le but d'un moteur de recherche est autre : il est de donner accès à tout moment aux pages Internet existantes. Le projet "Statistiques" a montré que l'idée de faire un recensement manuel des publications de qualité n'est pas réaliste. Il semble que la quantité de publications à recenser dans la bibliographie nationale soit gérable.

La pérennité des publications recensées semble aussi acceptable. En 2001, le Centre bibliographique danois a vérifié les URL de 4000 ressources recensées dans la Bibliographie nationale et seulement 20 publications avaient complètement disparu. Cela signifie que les publications que nous avons sélectionné sont stables. La coopération avec la Bibliothèque administrative signifie qu'à partir de maintenant toutes les publications gouvernementales sont protégées de la disparition. Un grand nombre de publications stables viennent du dépôt légal mais ne sont malheureusement pas disponibles pour des raisons statutaires. Dans le futur les publications Internet sélectionnées pour être recensées dans la bibliographie nationale, à la fois les publications "stables" et les publications "dynamiques", devraient venir du dépôt légal et être conservées. Cette tâche devrait revenir aux institutions culturelles existantes.