



IFLA
2005
OSLO

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>

juillet 27, 2005

Code Number:
Meeting:

035-F
150 SI - ICABS (IFLA/CDNL Alliance for Bibliographic Standards, Alliance IFLA/CDNL pour les standards bibliographiques)

La collaboration en matière d'archivage numérique au Royaume-Uni

Caroline Brazier

Head of Collection Acquisition and Description, The British Library, UK

Traduction par Sandrine AVRIL,
INIST-CNRS,
sandrine.avril@inist.fr

Résumé

Le présent article donne un bref aperçu des travaux collaboratifs actuels au Royaume-Uni concernant la mise en place d'une archive nationale numérique des publications. L'article expose aussi les différentes modifications apportées à la législation qui visent à faciliter la création d'une archive numérique et qui s'inspirent du concept de la National Published Archive pour les publications papier. Les projets pilote collaboratifs destinés à faire progresser la connaissance et l'expérience en la matière et à étudier les défis en présence sont aussi présentés. Nous soulignons enfin combien il est vital de mettre en place un cadre collaboratif pour pouvoir progresser en réunissant non seulement les bibliothèques qui auront la charge de l'archive nationale numérique, mais aussi l'industrie de l'édition.

Introduction

Il existe au Royaume-Uni une longue tradition de collaboration entre les six bibliothèques de dépôt légal britanniques dans le développement et la gestion des archives nationales des

publications ; cette collaboration s'appuie sur une législation sur le dépôt légal qui reste relativement stable depuis 1911. Les bibliothèques concernées ont toujours collaboré de manière très étroite sur les questions de développement des collections ou sur les questions opérationnelles liées à la collecte des documents. La collaboration a parfois pris une tournure plus formelle avec le programme sur le partage du catalogage, "*Shared Cataloguing*", dans lequel les six bibliothèques (British Library, National Library of Scotland, National Library of Wales, Bodleian Library, Oxford, Cambridge University Library et Trinity College Dublin) ont officiellement partagé le catalogage des acquisitions de monographies, et alimenté de ce fait la *British National Bibliography*.

Avec la croissance et le développement rapides de l'édition électronique, qui a commencé au cours des années 1990 avec les formats hors ligne comme les CD, on a craint que les archives nationales des publications ne parviennent plus à refléter le contenu et la nature des publications au Royaume-Uni, à moins d'une révision de la législation sur le dépôt légal qui autoriserait aux bibliothèques la collecte, la préservation et la mise à disposition des formats numériques au même titre que les formats papier. Il suffit de passer brièvement en revue les évolutions législatives, techniques et opérationnelles de ces dix dernières années pour se rendre compte du chemin parcouru pour nous rapprocher de notre objectif de création d'une archive numérique exhaustive pour le Royaume-Uni.

La révision de la loi britannique sur le dépôt légal

La loi sur le droit d'auteur de 1911 a défini les règles qui ont régi le dépôt des publications papier au cours des 94 années qui ont suivi. A mesure que les formats de publication autres que le papier sont apparus au cours du XX^{ème} siècle, et que des enregistrements sonores ou des images devenaient accessibles au public, on a mis en œuvre des programmes volontaires pour collecter ces derniers, sans que la loi sur le dépôt légal ait été amendée pour inclure ces nouveaux supports. Les bibliothèques se sont montrées de plus en plus préoccupées au cours de la décennie 1990 par la prolifération des formats de publication numériques, d'abord sur supports physiques, hors ligne, comme les CD puis sur des formats en ligne. Des discussions avec les représentants de la communauté des éditeurs britanniques ont eu lieu à la fin des années 1990, et ont débouché sur la rédaction d'un Code volontaire de bonnes pratiques pour le dépôt des publications numériques sur support physique (*Voluntary Code of Practice for Deposit for Offline Publications*), qui est entré en vigueur en 2000.

Révision de la législation britannique sur le dépôt légal

- La loi sur le droit d'auteur de 1911 (*Copyright Act 1911*) a défini les règles du dépôt légal pour les publications papier
- Programmes volontaires pour les documents sonores et les films au cours du XX^{ème} siècle
- Code volontaire de bonnes pratiques pour le dépôt des publications numériques sur support physique en vigueur depuis l'année 2000 (*Voluntary Code of Practice for Deposit for Offline Publications*)
- Loi sur les bibliothèques de dépôt légal de 2003 (*Legal Deposit Libraries Act 2003*)
- Création d'un Comité consultatif sur le dépôt légal (*Legal Deposit Advisory Panel*) pour conseiller le gouvernement sur les mesures réglementaires à prendre

- Création du Comité mixte sur le dépôt légal (*Joint Committee on Legal Deposit*) réunissant bibliothèques et éditeurs pour poursuivre la collaboration sur les réglementations à venir.

Les déclarations officielles à l'intention du gouvernement se sont poursuivies et ont abouti à l'adoption de la loi sur les bibliothèques de dépôt légal en 2003 (*Legal Deposit Libraries Act 2003*). Cette loi étend le concept de dépôt légal aux supports autres que le papier, tout en confirmant l'applicabilité des programmes volontaires pour les documents sonores et filmés. La loi de 2003 est une législation cadre et ne précise pas dans le détail les dispositifs de dépôt pour les différents formats numériques. Ceux-ci seront précisés dans une série de décrets d'application qui seront adoptés en relation avec la loi. Un Comité consultatif sur le dépôt légal (*Legal Deposit Advisory Panel*) est en cours de constitution ; son rôle sera de conseiller le gouvernement sur la nature de ces décrets d'application. Ce Comité consultatif comprendra à la fois des représentants des bibliothèques de dépôt légal et d'associations représentant différents aspects de l'industrie de l'édition.

Collaboration entre bibliothèques et éditeurs

Actuellement, nous effectuons les opérations de dépôt et de collecte de documents aux formats autres que le papier selon des codes volontaires de bonnes pratiques ou, dans le cas de sites Web, après accord exprès de la partie concernée. Les bibliothèques de dépôt légal et plusieurs associations d'éditeurs britanniques ont formé un Comité mixte sur le dépôt légal (*Joint Committee on Legal Deposit* ou JCLD) pour suivre les codes volontaires de bonnes pratiques et pour apporter leur contribution à l'élaboration de nouvelles réglementations au titre de la nouvelle loi.

Joint Committee on Legal Deposit

8 représentants des éditeurs

- Publishers Association
- Digital Content Forum
- Digital Publishers Association
- Association of Learned and
- Professional Society Publishers
- Association of Online Publishers
- Newspaper Publishers Association
- Periodical Publishers Association
- Scientific and Technical Publishers

8 représentants des bibliothèques

- British Library (3)
- National Library of Scotland
- National Library of Wales
- Bodleian Library, Oxford
- Cambridge University Library
- Trinity College Dublin

Grâce à ses consultations bilatérales, ce Comité mixte étudie les questions auxquelles sont confrontés bibliothèques et éditeurs, notamment sur le dépôt des contenus dans des formats numériques différents ; il vise aussi à définir des recommandations communes pour proposition au Comité consultatif (*Advisory Panel*). Actuellement, le Comité travaille essentiellement sur les programmes volontaires actuels pour les formats hors ligne, puisque le gouvernement devrait commencer par une réglementation sur les formats sur supports physiques dans le cadre de la nouvelle loi.

La création d'un programme volontaire pour le dépôt des revues électroniques est un autre domaine d'activités communes ; un pilote est en cours de réalisation. Dans le cas aussi bien des documents hors ligne que des revues en ligne, le débat et les compromis portent à l'heure actuelle sur les questions d'accès et de la menace que cela représente pour les éditeurs qui craignent des répercussions sur les ventes de leurs publications.

Le Comité travaille aussi sur les questions juridiques de territorialité et tente de définir ce qu'est « l'édition britannique » dans un contexte de mondialisation de l'édition. Une fois traitées les questions relatives aux documents hors ligne et aux revues en ligne, le Comité mixte s'intéressera aux questions de dépôt des bases de données, d'ensembles de données et au moissonnage des sites Web britanniques.

Collaboration entre bibliothèques de dépôt légal

La collaboration entre les bibliothèques de dépôt légal n'est pas un fait nouveau et a déjà fait ses preuves dans le développement des archives nationales des publications (*National Published Archive*). Toutefois, le nouveau cadre législatif ayant mis l'accent sur l'électronique, au détriment du papier, les bibliothèques ont créé de nouveaux groupes de travail pour définir un cadre pour une archive numérique.

Collaboration entre bibliothèques

- Les 6 bibliothèques de dépôt légal du Royaume-Uni ont une longue tradition de collaboration pour les publications papier ;
- Etudier comment cela peut être transposé à une archive numérique ;
- Création de nouveaux groupes de travail pour étudier les questions de :
 - l'infrastructure commune de stockage et de conservation qui devra être interopérable avec les besoins en local ;
 - un cadre commun de métadonnées, également interopérable avec les dispositifs en local ;
 - politiques d'acquisition partagée des collections ;
 - gestion rentable du cycle de vie des collections et de leur préservation.

Ces groupes de travail sur l'infrastructure, les métadonnées, les politiques d'acquisition et de préservation étudient les questions et problèmes liés à la mise en place et à la gestion d'une archive numérique partagée entièrement opérationnelle. Une telle archive doit pouvoir enregistrer, traiter, stocker, préserver du contenu et permettre un accès adéquat à ce contenu, qui devrait être beaucoup moins restrictif que ne le sont les accords de licence commerciaux habituels.

Groupes de travail des bibliothèques

Infrastructure

Métadonnées

Préservation

Acquisitions

- Interdépendances complexes
- Quelles métadonnées sont requises par l'infrastructure ?
- Quelles sont les métadonnées qui peuvent être moissonnées ? Ou générées automatiquement ?
- Quels sont les coûts liés au cycle de vie des différents modèles d'acquisition ?
- De quelle manière l'infrastructure peut-elle au mieux assurer la préservation des documents ?

Les discussions ont jusqu'ici porté essentiellement sur l'infrastructure technique et la nécessité de mettre en place une infrastructure partagée pour tenir compte des différents objectifs institutionnels et des spécificités techniques locales des 6 bibliothèques. Les discussions sur le choix d'un cadre adéquat pour les métadonnées, couvrant non seulement les métadonnées descriptives, mais aussi les aspects techniques, administratifs, structurels, les aspects liés à la préservation et aux droits numériques montrent également à quel niveau de complexité le nouveau système devra pouvoir répondre. Le cadre de métadonnées devra d'une part servir de base à l'interopérabilité des ressources entre les 6 bibliothèques pour la découverte des ressources et pour la gestion des collections dans la cadre de l'archive numérique et d'autre part prendre en charge l'interopérabilité entre une collection de dépôt légal numérique centralisée et les 6 collections distinctes d'archives numériques locales que chacune des bibliothèques développe en parallèle.

Les questions d'acquisition revêtent également une importance cruciale dans la définition des spécifications d'une future infrastructure partagée. Les bibliothèques doivent définir leurs intentions concernant la nature et l'importance des collections qu'ils envisagent de constituer pour l'archive nationale des publications numériques. Du fait même des volumes et de la nature dynamique de l'édition numérique, en particulier sur le Web, les bibliothèques concernées seront amenées à adopter une approche bien plus sélective qu'elles ne le faisaient jusqu'alors pour les publications papier.

D'un point de vue pragmatique, le degré réel de sélectivité découlera des ressources disponibles pour gérer tout ce qu'implique le cycle de vie complet des documents aux formats autres que le papier, non seulement pour l'alimentation initiale de l'archive et pour la création des métadonnées, mais aussi pour la préservation à long terme. Les travaux des quatre comités sont donc étroitement liés.

Dépôt volontaire des publications électroniques

Dépôt volontaire des publications électroniques

- Code volontaire de bonnes pratiques pour le dépôt des documents aux formats autres que le papier, en vigueur depuis 2000
- S'applique aux publications hors ligne
- À l'heure actuelle, 90 % des dépôts concernent des documents en ligne
- 70 % de l'acquisition des publications en série concerne des documents en ligne

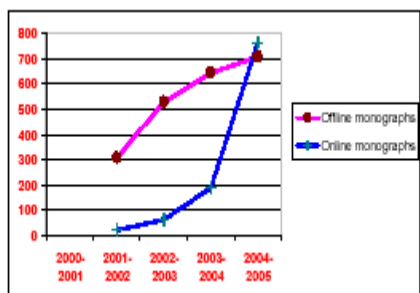
- DigiTool de Ex Libris est utilisé pour l'enregistrement et le stockage des documents
- Les questions d'acquisition devront être continuellement rediscutées pour garantir la qualité à mesure que la quantité augmente

Les dispositifs actuels de dépôt volontaire sont en vigueur depuis l'année 2000. Il n'existe actuellement aucune opération partagée et des dispositifs légèrement différents sont en place dans chacune des 6 bibliothèques. Le présent article aborde plus spécifiquement les pratiques en usage à la British Library.

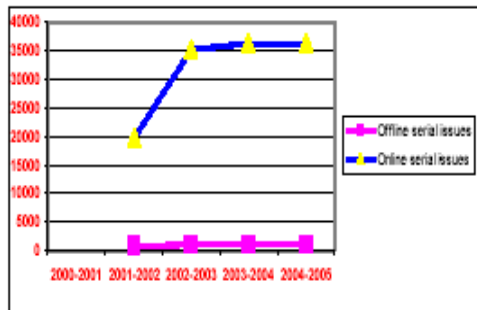
Le Code volontaire de bonnes pratiques pour le dépôt des documents non imprimés¹, en vigueur depuis 2000, est un instrument destiné à couvrir les publications hors ligne. Toutefois, dans la pratique, nous avons observé que de nombreux documents en ligne ont été déposés selon ce code. Le rapport documents en ligne sur documents hors ligne ne cesse de croître et en moyenne 90 % des documents enregistrés par mois en 2005 étaient des documents en ligne. Alors que plus de 70 % des publications en série ont été déposées au format en ligne, ce contenu ne provenait toutefois pas des grands éditeurs scientifiques et universitaires. Les opérations des cinq dernières années ont consisté pour une bonne part à sécuriser le stockage et à définir les procédures de gestion pour ces documents en ligne.

Nombre de nouveaux dépôts volontaires

- Le nombre des monographies en ligne augmente suite aux discussions en amont avec les éditeurs, notamment concernant la littérature grise
- Diminution des formats hors ligne pour les publications en série
- Le dépôt des publications en série en ligne ne progresse pas aussi rapidement en raison de la crainte des éditeurs commerciaux de perdre une part des revenus provenant des licences



¹ <http://www.bl.uk/about/policies/codeprac.html>



Les publications sur support physique, comme les microformes ou les CD, sont traitées conformément à nos procédures existantes pour le papier. Les documents en ligne sont passés par plusieurs phases depuis 2000. Au début, les documents en ligne étaient généralement reçus par courrier électronique, et même si nous avions des interrogations quant à la valeur scientifique de certains contenus, notre politique alors était de tout accepter dans cette phase pilote pour apprendre à gérer et traiter le dépôt en ligne. Tout était donc accepté dans la phase pilote, mais seuls étaient catalogués les contenus jugés présenter une valeur scientifique susceptible de durer dans le temps.

Les besoins en stockage ont vite dépassé les capacités du serveur original de messagerie. Ainsi, dans la phase suivante, le contenu a été gravé provisoirement sur des cédéroms, le temps que la bibliothèque étudie les solutions de stockage et de gestion à long terme des objets numériques. En 2002/2003, la British Library a choisi un système intégré de gestion de bibliothèque pour les opérations de routine de gestion des collections et cela a été l'occasion d'étudier en parallèle les différentes solutions de stockage numérique proposées par les fournisseurs de logiciels de bibliothèques. Le choix a porté sur le produit DigiTool de Ex Libris pour gérer la phase suivante d'enregistrement et de stockage des documents déposés volontairement. Le personnel de la British Library a travaillé en collaboration avec les développeurs de Ex Libris pour personnaliser le logiciel, automatiser en partie certains des traitements d'enregistrement de routine et pour permettre un fonctionnement le plus efficace possible et un partage des métadonnées avec le système intégré Aleph, utilisé pour la gestion du fonds papier.

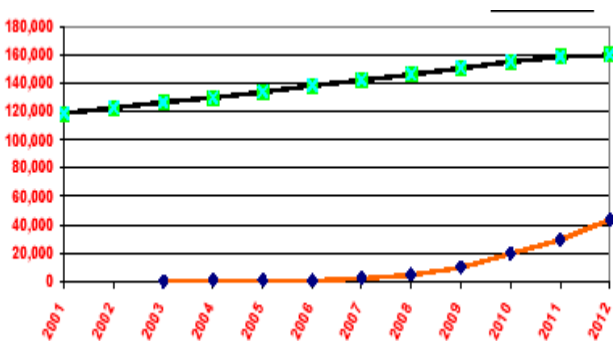
Le dispositif actuel sera revu en fonction des évolutions en cours du système principal de gestion des objets numériques (*Digital Object Management*) et des évolutions générales de l'infrastructure des 6 bibliothèques de dépôt légal.

La question du développement des collections devra aussi être posée dans les 6 bibliothèques de dépôt légal. La British Library a déjà dû changer de politique interne de développement des collections pour les dépôts volontaires puisque de nombreux documents ont été jugés ne pas présenter une valeur scientifique à long terme suffisante pour être collectés à grande échelle et pour être conservés à long terme (par exemple les cartes d'anniversaire en ligne, les publications en ligne à compte d'auteur, les documents publicitaires), tandis qu'il était clair que d'autres documents ne provenaient pas du Royaume-Uni ou étaient destinés à des publics internes uniquement (communications internes à une entreprise).

Les prévisions d'échelle pour l'édition numérique semblent indiquer également que les bibliothèques devront être sélectives. Les études portant sur les prévisions d'échelle de l'édition numérique, réalisées pour le compte des bibliothèques de dépôt légal dans le cadre du projet de loi de 2003, montrent une augmentation considérable de l'édition de publications en série et de monographies en ligne. Ces études ne portent pas sur la croissance des modes d'édition en libre accès sur le Web.

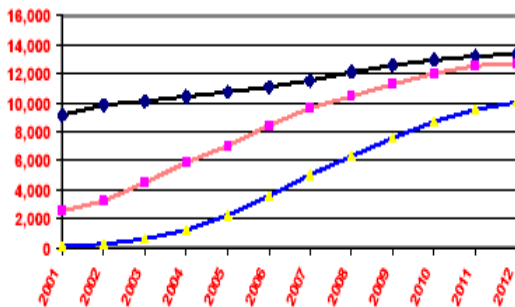
Tendances de la publication de monographies au Royaume-Uni

- On prévoit que la production de monographies devrait augmenter à la fois pour le format papier et pour le format électronique
- 25 % de la production totale des publications sera au format électronique à l'horizon 2012



Tendances de l'édition de publications en série au Royaume-Uni

- La part des publications en série au format électronique devrait augmenter considérablement (par rapport au nombre total de publications)
- On prévoit que 77 % seront uniquement au format électronique à l'horizon 2012



Pilote pour les revues électroniques

La majorité des contenus reçus dans le cadre du dépôt volontaire des publications électroniques sont des publications en série. Toutefois, la plupart des grands éditeurs de revues scientifiques et académiques ne sont pas disposés à déposer leurs articles volontairement en raison des problèmes d'accès et de gestion numérique des droits (DRM). Le Comité mixte a donc fait de cette question et de cette préoccupation une priorité qui doit être examinée dans les meilleurs délais par les éditeurs et les bibliothécaires.

Un pilote volontaire a été mis en place avec environ 75 à 100 titres publiés par près de 20 éditeurs de revues scientifiques. L'objectif est d'étudier les problèmes liés à l'enregistrement, au stockage et à la préservation de toute une variété de formats et de types de contenus provenant de plusieurs fournisseurs pour documenter tous travaux ultérieurs de réglementation sur le dépôt des revues électroniques. Les éditeurs parties au comité et les associations les représentant sont les porte-parole de toute l'industrie de l'édition et rendront compte de ces résultats à l'ensemble de la profession. Il est prévu d'organiser un atelier dès que les premiers résultats du pilote sur l'enregistrement seront disponibles. Une fois encore, un des grands problèmes à résoudre est l'adoption de modèles appropriés pour l'enregistrement des documents et le transfert des métadonnées avec le contenu déposé et/ou l'extraction à partir de ce même contenu, en veillant à ce que la gestion de l'archive entre les différentes bibliothèques de dépôt légal reste rentable.

Pilote sur les revues électroniques

- Création d'un pilote pour étudier les questions techniques permettant un enregistrement et un stockage rentables
- Recherche parmi les éditeurs participant à l'opération de volontaires pour faire des essais, en coopération avec les associations d'éditeurs pertinentes
- 20 éditeurs universitaires et scientifiques proposent 75 à 100 titres pour la phase 1
- Les résultats seront communiqués à l'ensemble du secteur par les éditeurs parties à l'opération et par le biais d'ateliers

UK Web Archiving Consortium (UKWAC) – consortium britannique pour l'archivage du Web

Un facteur important limitant à l'heure actuelle les opérations de dépôt volontaire est que nous ne disposons pas de l'infrastructure nécessaire de collecte automatique, qui nous permettrait d'enregistrer tout nouveau site Web ou tout changement dans un site Web que nous signalerait un administrateur de site. Si nous pouvons collecter à la main des documents ou des fichiers isolés, nous avons dû exclure pour des raisons pragmatiques la collecte de sites Web entiers comportant des liens dynamiques. Le consortium britannique pour l'archivage du Web a été créé pour combler cette lacune (*UK Web Archiving Consortium, UKWAC*).

UK Web Archiving Consortium - UKWAC

Membres

The British Library
National Library of Scotland
National Library of Wales
The National Archives
Joint Information Systems Committee
The Wellcome Library

Objectif : moissonner 6 000 sites Web britanniques en 2 ans

Utilisation du logiciel de moissonnage PANDAS de la bibliothèque nationale d'Australie

Le site de l'UKWAC a été lancé en mars 2005 (URL : www.webarchive.org.uk)

Les membres du Consortium UKWAC sont la British Library, la National Library of Scotland, la National Library of Wales, The National Archives, le Joint Information Systems Committee (qui a pour but d'aider les établissements d'enseignement supérieur et de formation continue au Royaume-Uni) et la Wellcome Library. Sa mission est de moissonner 6 000 sites Web

britanniques entre 2005 et 2007, conformément aux objectifs des différentes institutions concernées ; l'ensemble permettra toutefois de constituer une archive globale du Web britannique.

L'infrastructure technique actuelle utilisée pour les opérations de l'UKWAC est basée sur l'outil de moissonnage PANDAS de la bibliothèque nationale d'Australie. Le site de l'UKWAC a été lancé en mars 2005 (Website: www.webarchive.org.uk). Les partenaires partageront les frais afférents à l'infrastructure entre eux, notamment l'hébergement du service et l'assistance technique. Les sites Web britanniques peuvent être définis de différentes manières, mais les critères qui ont été retenus par l'UKWAC sont les suivants :

Il s'agit de sites :

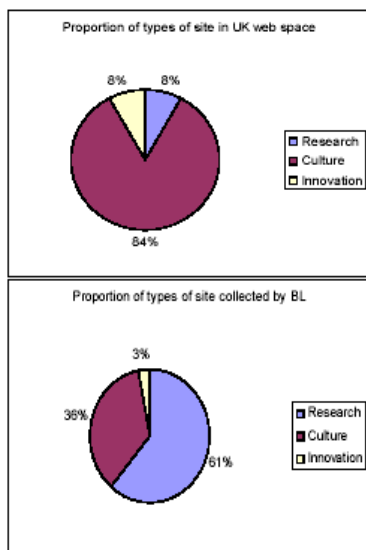
- qui ont un nom de domaine .uk
- qui sont hébergés au Royaume-Uni (.com, .org, etc.)
- qui appartiennent à des entreprises ou organismes britanniques
- qui présentent un contenu intellectuel britannique significatif
- dont le contenu intellectuel présente un intérêt particulier pour le Royaume-Uni.

Ces critères ne correspondent pas nécessairement aux définitions juridiques de la territorialité qui serviront de base aux travaux de réglementation sur le dépôt légal. Toutefois, comme tout moissonnage dans le cadre de l'UKWAC se fait sous conditions d'autorisation expresse des propriétaires des sites concernés, cela ne constitue pas un problème majeur pour ce projet.

On estime à plus de 4 millions le nombre de sites « britanniques », il est donc essentiel que les critères de sélection soient clairement identifiés pour s'assurer que les besoins des chercheurs sont couverts sur le long terme. Dans sa politique de collecte pour l'UKWAC, la British Library a le souci de respecter un équilibre approprié entre des sites portant sur :

- la recherche, c'est-à-dire des sites hébergés par des universités, des associations caritatives, des associations militantes, des administrations ;
- la culture, c'est-à-dire des sites représentant l'importance/la diversité culturelle britannique, les grandes manifestations nationales, l'actualité ;
- l'innovation, c'est-à-dire des sites ayant été primés, des sites illustrant les forces d'information, de communication et de formation du Web.

Stratégie de l'UKWAC en matière de collections



- Stratégie actuelle sélective portant en priorité sur des sites présentant un contenu scientifique significatif
- Ciblage spécifique des sites d'actualités (p. ex. les élections)
- L'obtention d'une photographie complète du Web dépendra des développements de la technologie des robots automatiques moissonnant le Web (*crawlers*)

En dehors des sites sélectionnés pour un moissonnage régulier, le projet de l'UKWAC s'intéresse aussi aux sites qui couvrent les aspects de la vie britannique présentant un fort potentiel d'actualité, mais de durée de vie limitée (p. ex. les sites sur le tsunami, les sites sur les élections générales de 2005).

Ce travail en est à sa phase de démarrage et nous en tirerons des enseignements sur le plan technique et opérationnel. L'obtention des autorisations nécessaires peut s'avérer être plus longue que prévu. La question sur les contenus couverts par ces autorisations (p. ex. moissonnage des contenus liés) se pose également. La fréquence à laquelle les sites peuvent être moissonnés à l'heure actuelle pose également le problème de la couverture suffisante du Web profond, notamment des bases de données, des ensembles de données et des documents dynamiques.

Ce modèle de moissonnage et d'archivage très sélectif ne pourra pas être étendu à une échelle suffisante pour pouvoir constituer une archive complète du Web. Nous travaillons également en collaboration avec le consortium International Internet Preservation Consortium (site Web : <http://netpreserve.org>) sur le développement d'un robot intelligent (*smart crawler*), qui aura la capacité de moissonner périodiquement des instantanés de la présence britannique sur le Web dans son intégralité. Pour obtenir le contenu déjà périmé (la durée de vie moyenne d'un site Web a été estimée à 44 jours) nous envisageons d'acheter les fichiers d'antériorité des espaces Web pertinents du Royaume-Uni auprès de l'Internet Archive (site Web : <http://www.archive.org>), lorsque les autorisations et les réglementations en la matière auront été clairement établies au Royaume-Uni.

L'UKWAC est un projet sur deux ans ; à l'issue de cette période, il sera évalué et des recommandations seront proposées sur les processus d'archivage évolutifs du Web à long terme. En attendant, nous espérons avoir mis en place une archive des documents disponibles sur le Web

qui constituera une composante essentielle de la future archive nationale numérique ; nous espérons avoir aussi montré que la collaboration est une manière efficace de partager la tâche de constitution d'une archive des documents disponibles sur le Web.

Perspectives

Au-delà du dépôt légal et perspectives d'avenir

- Collaboration au-delà du dépôt légal
 - Archives des universités et de la recherche
 - Infrastructure nationale pour les thèses en ligne
 - Numérisation à grande échelle, p. ex de la presse écrite, des enregistrements sonores
- Priorités d'avenir
 - Infrastructure
 - Métadonnées
 - Droits numériques et nouveaux services aux éditeurs.

Cet article a abordé un large spectre d'activités et de développements liés à différents aspects de l'archivage numérique au Royaume-Uni. Je me suis concentré sur les travaux effectués dans le cadre des évolutions du dépôt légal numérique, mais des travaux sont également en cours sur la création d'archives institutionnelles pour la recherche et les universités, sur la création d'une infrastructure nationale pour les thèses en ligne et sur la numérisation à grande échelle de la presse écrite et des enregistrements sonores, autant de ressources qui seront intégrées au plan national d'archivage numérique du Royaume-Uni. La route sera de toute évidence longue. Mais pour l'instant nous continuons à privilégier trois grands domaines. Il s'agit premièrement du développement d'outils et d'une infrastructure pour l'enregistrement, le stockage, la préservation et la découverte de ressources des contenus numériques destinés au dépôt légal.

Deuxièmement, nous poursuivons, avec un souci de rentabilité, nos travaux sur la création, l'extraction ou la génération de métadonnées qui viendront étayer notre capacité à gérer le contenu de la future archive nationale numérique.

Et, *last but not least*, nous tentons de mettre en place des directives sur la gestion numérique des droits et sur l'accès au contenu archivé. La collaboration continue avec les éditeurs et avec d'autres producteurs de contenus est une composante essentielle des opérations à venir. Nous devons étudier les modèles garantissant qu'une archive nationale numérique ne constitue pas une menace pour les intérêts des éditeurs, mais leur offre la possibilité de développer de nouveaux services essentiels pour la communauté de la recherche. 16 avril 2005