



La valeur des services documentaires en prise avec le numérique

GHISLAINE CHARTRON

Conservatoire national des arts et métiers
ghislaine.chartron@cnam.fr

Ghislaine Chartron est professeur titulaire de la chaire d'ingénierie documentaire au Conservatoire national des arts et métiers (Cnam). Elle dirige l'Institut national des sciences et techniques de la documentation depuis 2006, elle est membre de l'équipe de recherche Dicen-Cnam. Elle a auparavant été professeur à l'INRP (IFé)-Lyon, en charge de la cellule Veille, et coresponsable de l'Urfist-Paris. Son champ principal d'investigation concerne la documentation et l'édition scientifique numérique, elle publie régulièrement dans ce domaine, <http://dicen.cnam.fr/enseignants-chercheurs/>

Vouloir souligner l'importance du numérique dans l'exercice des métiers de l'information, bibliothécaire, documentaliste ou archiviste, est certainement une évidence... Les contenus circulent massivement sous forme numérique, les procédures et les modes de valorisation ont évolué en parallèle pour répondre aux besoins et à la qualité de services attendue. L'objectif de ce texte est surtout d'insister sur le caractère très évolutif des technologies mobilisées et de pointer certains enjeux actuels bousculant les organisations de travail en place : quel périmètre doit-on considérer aujourd'hui pour certains services ? Pourquoi le nouvel enjeu de l'ouverture des données ? La place du numérique dans les formations ? Mais l'objectif majeur reste de veiller à la qualité des services, des contenus, des pratiques informationnelles et d'intégrer les innovations numériques porteuses d'enrichissements et de sens...

Cycle de vie numérique et évolutivité du web

De sa production à sa lecture, le document est aujourd'hui en prise avec des technologies numériques. Un enjeu majeur des dernières années fut certainement la question des formats à partager pour permettre le développement d'un « écosystème », d'un marché permettant la rencontre d'une offre et d'une demande, le développement de services performants entre les auteurs, les éditeurs, les intermédiaires et les usagers. Nous rappelons dans le tableau en page suivante

quelques technologies accompagnant désormais le cycle de vie des documents.

Mais l'ensemble du cycle de vie est confronté à une très grande évolutivité et il en résulte, de toute évidence, une instabilité des services, des procédures et des organisations de travail. L'histoire des services est intrinsèquement liée aux différentes générations de technologies qui ont permis leur développement.

Le web 1.0, développé au tout début des années 1990 par Tim Berners-Lee au CERN, est celui de l'hypertexte mondial en réseau, avec une logique diffusionniste sur de multiples serveurs : l'utilisateur accède de façon inédite à une profusion de contenus auto-diffusés sans interagir directement. Les bibliothèques vont s'enrichir de sites web donnant accès à une diversité de ressources et penser au mieux leur portail d'accès selon les besoins et les technologies disponibles.

À partir des années 2000, le web 2.0 va introduire des technologies permettant l'interaction des internautes entre eux ; les services blogs, wikis, tags, réseaux sociaux, se multiplient. C'est le web qualifié de « web social » où la légitimité se redistribue et se partage avec les internautes. L'horizontalité du web est introduite et compose désormais avec la logique descendante. Certaines bibliothèques ont choisi d'intégrer ce type de technologies qui viennent compléter l'offre de services existante : Opacs intégrant les commentaires d'utilisateurs, développement de services Netvibes rapprochant des sources de l'édition traditionnelle avec des sources d'inter-

**EXEMPLES DE TECHNOLOGIES ASSOCIÉES
AU CYCLE DE VIE DU DOCUMENT NUMÉRIQUE**

Phase du cycle de vie	Technologies numériques	Fonctionnalités
Production	<ul style="list-style-type: none"> • DTD XML spécifiques • Format du livre numérique : Epub • Techniques de numérisation 	<ul style="list-style-type: none"> • Documents structurés : performance ultérieure des traitements • Affichage multisupport • Automatisation de l'encodage numérique
Identification, description	<ul style="list-style-type: none"> • Identifieurs uniques (DOI, URI...) • Métadonnées structurées • Tags 	<ul style="list-style-type: none"> • Unicité d'identification • Qualité des accès et des exploitations, des liens • Enrichissement social
Diffusion, recherche	<ul style="list-style-type: none"> • Technologies de portails dont les protocoles de dialogue (OAI-PHMH, SRU, connecteurs) • Technologies web 2.0 : Netvibes, blogs... 	<ul style="list-style-type: none"> • Recherche fédérée, outils de découverte • Services d'actualités, services interactifs
Usages	Format COUNTER	Mesure quantitative des usages des ressources numériques
Communication	Technologies de réseaux sociaux	Animation d'une communauté
Archivage	Modèle conceptuel (OAIS)	Pérennité d'accès

nautes (blogs)¹, présence de la bibliothèque sur les réseaux sociaux de leurs usagers de type Facebook...

L'horizon se déplace aujourd'hui vers d'autres technologies prometteuses : le web sémantique et le « *linked data* », l'informatique dans les nuages, l'universalité (indépendance à tout système d'exploitation, à tout matériel, fabricant, marque, logiciel, ou *plugin*...). Le terme web 3.0 voudrait couvrir toutes ces nouvelles ambitions, son identité est encore diffuse. Les enjeux du web sémantique dominent toutefois avec le mouvement de l'ouverture des données dans des formats ré-exploitable qui devraient permettre une interaction accrue entre les données pour élaborer de nouvelles connaissances, concevoir des services innovants par l'agrégation de données. Nous reviendrons en détail sur cet enjeu dans la suite de l'article.

Pour souligner l'évolutivité des technologies web et son impact sur les services d'une bibliothèque, nous

prendrons encore trois exemples : le catalogage, l'accès à l'information et l'architecture technique de support.

Le catalogage des ressources documentaires

Le catalogage des ressources documentaires est confronté désormais aux nouvelles modélisations de données pour intégrer le web sémantique et s'agréger avec d'autres données. RDA (*Resource Description and Access*) est un nouveau cadre pour la description et l'accès aux ressources, au cœur des discussions des réseaux de catalogage. RDA repose sur les modèles conceptuels FRBR (*Functional Requirements for Bibliographic Records*) et FRAD (*Functional Requirements for Authority Data*) et s'appuie sur le formalisme informatique entités-attributs-relations. Le code associé décrit ainsi les entités dans les deux modèles, les attributs qui permettent de les caractériser et les relations établies entre ces entités. L'organisation de l'information bibliographique selon les entités du modèle FRBR permet de regrouper les informations relatives à une même

œuvre (par exemple, les traductions, les révisions, les adaptations...). Cette modélisation est en phase avec les formalismes du web sémantique et avec des outils de recherche plus diversifiés que ceux dédiés exclusivement aux bibliothèques². L'ouverture, l'interopérabilité avec les outils du web sont les enjeux majeurs. Les États-Unis, l'Australie, le Canada, le Royaume-Uni et l'Allemagne ont décidé d'adopter RDA en 2013. La France a, pour le moment, mis en place un groupe stratégique chargé d'orienter les futures décisions³.

L'accès à l'information

Concernant l'accès à l'information, la recherche est confrontée aujourd'hui à des ressources de plus en plus nombreuses et hétérogènes. Les techniques d'interopérabilité visent à gérer cette diversité, à masquer la complexité des stratégies d'interrogation pour présenter des réponses unifiées. La figure 1 empruntée à Philippe Bourdenet illustre cette diversité des ressources et des protocoles qui peuvent être sollicités, c'est une accumulation de standards et de structurations sémantiques différentes au fil des strates du web.

Le rôle du professionnel de l'information est alors de rendre possible cette interopérabilité avec une description sémantique commune (penser les correspondances), d'implémenter ces descriptions dans un langage structuré compréhensible par les machines et de faire dialoguer ces ressources par des protocoles partagés. Le cahier des charges et le travail sémantique sont des tâches centrales, la mise en œuvre se fait avec les équipes informatiques. Le SID (Système d'information documentaire) est le dispositif global rassemblant des ressources hétérogènes et dont le catalogue en ligne du système de gestion de bibliothèque (SGB) n'est qu'un composant.

1. Muriel Amar et Véronique Mesguich, *Le Web 2.0 en bibliothèques : quels services ? Quels usages ?*, Paris, Éditions du Cercle de la Librairie, 2009.

2. Philippe Bourdenet, « Le catalogue à l'épreuve du web : une mise en perspective historique », *Isko-France*, Lille 3, 27-28 juin 2011, Hermès, 2012, p. 119-130.

3. www.bnf.fr/fr/professionnels/rda

L'architecture technique de support

Nous prendrons un troisième exemple pour souligner l'instabilité des technologies avec l'architecture technique de support. Le passé est jalonné d'étapes : banque de données sur serveurs distants, réseau local de cédéroms, serveur web local, hébergé... Aujourd'hui, les questions se focalisent sur le *cloud computing* (informatique dans les nuages) à différents niveaux : pour les données, pour les applications et pour l'archivage.

L'informatique dans les nuages désigne une architecture technique où les applications informatiques d'une entreprise sont transférées à un fournisseur de services sur internet. Ce transfert concerne les serveurs, les systèmes mais aussi les applications. Le prestataire garantit une qualité de service souvent inégalée en interne, notamment au niveau de la continuité du service, de la performance, du suivi de l'évolution technologique et de la sécurité.

Le *cloud computing* est déjà inscrit dans certains de nos usages quotidiens : nous utilisons la suite bureautique de Google, les réseaux sociaux comme Facebook, les services de stockage de fichiers comme Dropbox... La légèreté des solutions techniques, externalisant toute maintenance et permettant un accès distant n'importe où, n'importe quand, est plébiscitée. Elle permet de se consacrer à l'activité principale et à l'innovation, et moins à la gestion. Les risques évoqués sont par contre la confidentialité des données, la perte de contrôle, la dépendance à l'hébergeur.

Pour les bibliothèques, le rapport OCLC⁴ insistait en 2010 sur trois dimensions concernant les avantages du *cloud computing* : l'efficacité technologique, la gestion optimisée des données, le développement de la force d'action commune des bibliothèques :

« Parmi les avantages du *cloud computing* pour les bibliothèques :

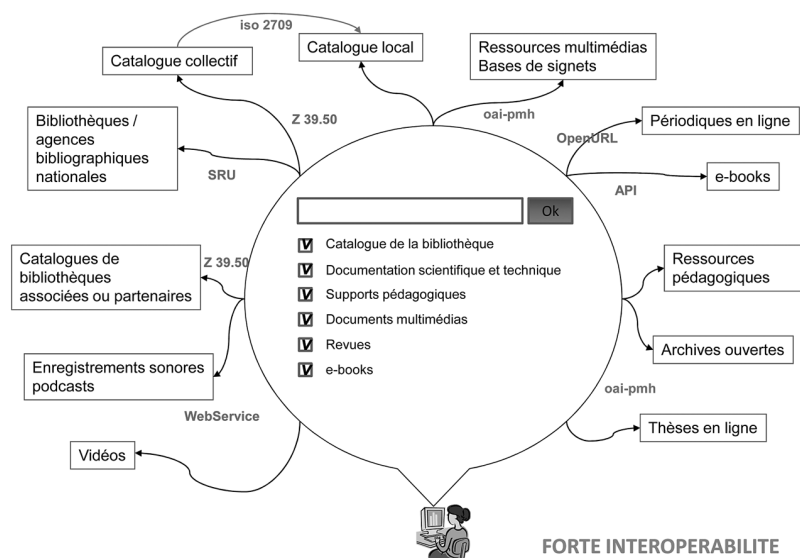


Figure 1 Diversité de l'interopérabilité

(source : Philippe Bourdenet, *Stratégies pour la construction de plateformes interopérables*, journée d'étude 2011 de la Fulbi, www.fulbi.fr/?q=content/2011)

- Bénéficier rapidement des technologies innovantes pour participer pleinement au web de contenus
- Améliorer la visibilité et l'accessibilité des collections
- Réduire la duplication des efforts pour les services techniques et le management des collections
- Rationaliser les flux, bénéficier de la coopération du réseau
- Développement de l'intelligence coopérative, niveaux de service améliorés par l'agrégation à grande échelle des données
- Rendre les bibliothèques plus « vertes » en partageant leurs moyens de calcul, réduisant ainsi l'empreinte carbone. »

Cette orientation vise à libérer les bibliothèques de la gestion technique locale pour les repositionner sur la conception de services, le développement de leurs contenus. Elle vise aussi à optimiser leur visibilité globale sur le web en concentrant leurs données sur un point d'entrée unique et à susciter une émulation commune de développement de services.

Dans cette logique, OCLC a initié de nouveaux services pour ses membres, « *OCLC WorldShare Management Services* » plus large que le catalogue partagé qui fonda sa première

activité ; l'offre veut proposer des services de gestion et d'analyse qui permettent de créer collectivement une valeur ajoutée globale, associée aux ressources partagées⁵.

De façon convergente, au niveau national, l'Abes réfléchit au développement d'un système de gestion de bibliothèque mutualisé entre établissements qui mettrait en commun les données de toutes les bibliothèques et de leurs fournisseurs⁶.

La question du périmètre : local, national, international ?

La question du périmètre est donc posée, l'évolutivité des technologies remet en question les modes d'organisation du travail, les procédures, les affectations de moyens, les répartitions entre internalisation et externalisation, la qualité des services rendus, les partenariats à privilégier... L'efficacité à trouver est un équilibre souvent péril-

4. Matt Goldner, *Winds of Change: Libraries ans Cloud Computing*, OCLC report, 2010, www.oclc.org/multimedia/2011/files/IFLA-winds-of-change-paper.pdf

5. www.oclc.org/webscale/origins.htm

6. [Ndlr] Voir, dans ce numéro, l'article de Jean Bernon, « Le projet de système de gestion de bibliothèque mutualisé de l'Abes », p. 61-65.

leux entre les budgets, les mutualisations possibles, les savoir-faire des équipes, la satisfaction des usagers et les contraintes locales.

L'évolution technologique induit à la fois la recherche d'économie d'échelle, la recherche de nouveaux partenariats et l'émergence de nouvelles fonctions locales. La question est donc de savoir quelle sera la redistribution des fonctions et des tâches des bibliothèques entre les niveaux local, national et international dans un contexte où les modes de financement évoluent également.

L'exemple des acquisitions est emblématique : consortiums, licences nationales. Dans le contexte français, émerge également le projet d'un portail commun national pour l'accès aux collections achetées dans le cadre du projet Istex⁷ («*l'excellence documentaire pour tous*») dont le financement est rattaché aux «*investissements d'avenir*» choisis par le gouvernement.

Pour les systèmes de gestion de bibliothèque, les gestionnaires de ressources numériques (ERM) se posent aussi la question du local ou du national. Les architectures de type *cloud computing* permettraient d'optimiser l'investissement et la gestion avec des technologies plus ouvertes que les solutions locales majoritairement propriétaires et captives des fournisseurs. Si certaines données sont exclusivement locales (les prêts, les données usagers), d'autres sont déjà mutualisées au niveau national (catalogage avec le Sudoc), et les bases de connaissances liées à la gestion des collections numériques sont de plus en plus produites au niveau international avec les fournisseurs de contenus (versement des métadonnées des éditeurs et des diffuseurs par exemple). Que restera-t-il au catalogage local dans un avenir qui sera de plus en plus en réseau entre tous les partenaires ?

Parallèlement, la valorisation de la production locale croît dans la mesure où l'auto-édition et l'auto-diffusion sont possibles sur le web et dans la mesure où la visibilité des «*territoires*» devient stratégique. Ces bases locales concernent les archives

ouvertes, les ressources pédagogiques locales, les collections patrimoniales numérisées... Elles doivent s'intégrer dans un environnement plus large (portail national comme Gallica, portail international comme Europeana), nécessitant des échanges nombreux, selon des formats partagés au plus haut niveau.

L'ouverture des données des bibliothèques

L'ouverture des données publiques est un nouveau défi généralisé qui trouve en grande partie son fondement dans le paradigme de *l'innovation ouverte*, telle que définie par l'économiste Chesbrough (2003)⁸ : les idées, les nouveaux savoir-faire sont en partie externes à toute organisation. Pour innover, il faut étendre le périmètre des organisations, intégrer des savoirs externes dans un contexte d'accélération, de spécialisations, de mobilité des compétences et avec l'objectif d'une augmentation du niveau général de connaissances, et de croissance des activités économiques... *L'open data* est un enjeu politique qui irrigue l'ensemble des secteurs et dont les bibliothèques sont également en train de se saisir pour s'associer à l'investissement dans ces nouveaux régimes de croissance.

L'accélération est appuyée par la directive européenne⁹ de 2003 concernant la réutilisation des informations du secteur public dans le cadre de la «*société de l'information et de la connaissance*», les enjeux majeurs étant à la fois l'égalité d'accès pour les citoyens et le potentiel économique associé à leur réutilisation dans divers domaines : le tourisme, l'environnement, les transports... L'objectif est alors en grande partie l'intégration de la bibliothèque dans son territoire. L'ouverture de ses données peut concerner à la fois la cartographie documentaire, les acquisitions,

les collections, les inscriptions et les prêts, en respectant la protection des données personnelles. Ces données peuvent être agrégées à d'autres pour proposer des services innovants. Imaginons effectivement que la localisation des ouvrages détenus par les bibliothèques publiques d'un territoire puisse être intégrée à des dossiers touristiques locaux, aux activités d'une plateforme associative de cours en ligne, aux intranets des entreprises...

La Bibliothèque nationale de France a ouvert les 12 millions de notices stockées dans ses différents catalogues, «*Nous voulons donc que les ressources de la BnF soient aussi visibles sur le web que la bibliothèque dans la ville*¹⁰.» L'engagement vers le web sémantique est affirmé¹¹. La Bibliothèque nationale et universitaire de Strasbourg fait aussi partie des pionnières, s'engageant davantage encore : son conseil d'administration a décidé en janvier 2012 de placer toutes ses productions numérisées¹² sous la *Licence ouverte/open licence* d'Etalab¹³ : les données bibliographiques mais aussi les fichiers numériques issus de la numérisation d'œuvres du domaine public conservées dans ses collections. Certaines collectivités sont particulièrement dynamiques, comme le conseil général de Saône-et-Loire sous l'impulsion particulière de son directeur¹⁴.

Le mouvement est à son début et il est prématuré aujourd'hui de faire un bilan entre le supplément de dépenses engagées et le bienfait sociétal ou le potentiel économique. Le développement du web sémantique est, en tout cas, pensé comme un moteur d'innovation dans les services et un accélérateur de synergie entre des acteurs pu-

7. www.istex.fr

8. Henry W. Chesbrough, *Open innovation: The new imperative for creating and profiting from technology*, Boston, Harvard Business School Press, 2003.

9. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32003L0098:FR:HTML>

10. <http://data.bnf.fr/docs/databnf-presentation.pdf> et «*Web sémantique et modèle de données*», in : site web de data.bnf.fr, 2011, <http://data.bnf.fr/semanticweb>

11. M. Barre, «*L'ouverture des données publiques à la BnF à travers l'exemple de data.bnf.fr*», *Cahiers du numérique*, à paraître, Lavoisier, 2012.

12. www.bnu.fr/collections/la-bibliotheque-numerique/les-images-de-la-bnu-et-la-licence-ouverte

13. <http://ddata.over-blog.com/xxxxxx/4/37/99/26/licence/Licence-Ouverte-Open-Licence.pdf>

14. www.opendata71.fr

blics et privés. Pour les bibliothèques, OCLC donne déjà une première liste intéressante d'applications innovantes créées avec les données ouvertes des catalogues¹⁵.

Compétences pour le numérique

Les services proches des usagers (accès à l'information, conseils, prêts) comme les services proches de la gestion (*back office* des bibliothèques) croisent pleinement les transformations numériques : le comportement des usagers change avec le web, et l'efficacité de la gestion passe par des solutions technologiques nouvelles. Le médiateur du document doit être familier avec la technique de son domaine pour conduire des choix pertinents et savoir dialoguer avec ses partenaires. L'enjeu est d'acquiescer des compétences numériques pour penser les services, programmer les cahiers des charges, négocier avec ses partenaires internes et externes. L'objectif n'est pas d'être développeur informatique (d'ailleurs, la majorité des entreprises aujourd'hui externalise de nombreux développements qui nécessitent une spécialisation de plus en plus pointue et évolutive) mais d'être le représentant des besoins et l'ingénieur compétent des services d'information numérique en charge de la gestion de projets diversifiés.

Le virage est pris dans de nombreuses institutions de formation, mettant en perspective le statut de l'information dans la société du XXI^e siècle. Un mouvement emblématique est celui de la reconfiguration initiée par les *i-schools* nord-américaines¹⁶ qui fondent leur cursus sur les interactions entre information, technologie et humain. L'objectif structurant est simple : garantir efficacement aux usagers l'information dont ils ont besoin pour prendre des décisions et faire avancer au mieux leurs

objectifs personnels et professionnels. Les compétences numériques sont renforcées mais toujours en équilibre avec les compétences liées aux sciences sociales pour la compréhension des interactions, des comportements et des besoins.

Un second point important à souligner concerne l'intrication renforcée entre la gestion de l'information et les autres activités d'une organisation aujourd'hui. L'information irrigue toutes les activités et accompagne leur développement, elle est centrale et justifie sa valeur ajoutée par une interaction permanente avec les autres services. Le « retour sur investissement » des moyens affectés s'observe en partie par les effets de cette interaction. Ainsi, pour une université, le renforcement du lien de la bibliothèque avec l'enseignement et avec la recherche (innovation) construit sa valeur ajoutée majeure. Pour une entreprise, le lien entre la gestion de l'information et l'innovation, la qualité de la production et la gestion des risques devient prioritaire. La gestion de l'information s'inscrit plus que jamais dans l'accomplissement des objectifs principaux de l'organisation. Pour une bibliothèque universitaire, le contexte numérique lui donne des opportunités inédites de penser avec les enseignants d'autres façons de transmettre des savoirs. La mise à disposition individualisée de lectures et de corpus de plus en plus riches permet de considérer le cours présentiel comme un espace de débat et de construction du raisonnement. Augmenter la qualité de la formation et l'adéquation à un public qui ne se satisfait plus uniquement d'une posture passive de réception sont des enjeux majeurs dans le contexte éducatif du XXI^e siècle.

En conclusion

L'objectif n'était pas d'encenser le numérique mais de montrer qu'il est, de toute évidence, devenu indispensable à la réalisation des missions de médiation documentaire dans tous les contextes. La compétence numérique doit être développée et régulièrement actualisée afin de remplir les missions fondatrices d'accompagnement pour

l'acquisition des connaissances et de répondre à des besoins d'informations variés avec une exigence de qualité et de diversité.

Le numérique nous appelle aussi à rester vigilants sur les transformations qu'il induit : favorise-t-il l'interdisciplinarité tant prônée aujourd'hui dans nos institutions¹⁷ ? Peut-il apporter une contribution significative à la qualité des processus d'expertise¹⁸ ? Au contraire, peut-il conduire également à des logiques de réduction, à des espaces appauvris pour les lecteurs... et si le mur Facebook devenait le principal prescripteur de lecture ?

Séparer le bon grain de l'ivraie... Le bibliothécaire, comme tous les autres professionnels de l'information, doit rester un éclaireur et un conseiller averti sur le chemin de l'acquisition des connaissances. ●

Septembre 2012

15. www.oclc.org/developer/applications

16. Myriam Hérigault, Massinissa Nait Mouloud et Kristell Roser, *Les i-schools d'Amérique du Nord*, janvier 2012. http://intd.cnam.fr/medias/fichier/les-ischools-en-amerique-du-nord_1338302039920.pdf

17. Hans Dillaerts, « Libre accès », « Accès sans barrière », vecteur de la créativité, de la pluridisciplinarité et de l'interdisciplinarité, conférence « Open access, services, interdisciplinarité et expertise » (Oasie), Cnam, 28 mars 2012, http://dicen.cnam.fr/medias/fichier/cnam-iscc-dillaerts_1334131676530.pdf

18. Emma Bester, « Les services pour les archives ouvertes : de la référence à l'expertise », *Documentaliste – Sciences de l'information*, 2010, vol. 47, n° 4, p. 4-15.