

Ar(abes)ques

JANVIER - FÉVRIER - MARS 2015

DOSSIER

Dialogues avec la machine

Concevoir, interagir, analyser

PLEINS FEUX SUR • La bibliothèque de l'École française de Rome

ACTUALITÉS • JPAC 2014 / Istex / En quête d'Arabesques / SGBm...

abes
agence bibliographique
de l'enseignement supérieur



(Dossier) DIALOGUES AVEC LA MACHINE

Concevoir, interagir, analyser

« L'écosystème numérique ne peut que renforcer les interrogations récurrentes sur la place respective occupée par l'homme et la machine dans la production, la circulation et la validation des informations. En écho à l'émergence de nouveaux modèles sociétaux, économiques et techniques, le rôle et les missions des bibliothèques sont revisités à travers ce prisme. À partir de quelques exemples, le dossier de ce numéro illustre les interactions nécessaires, qu'il s'agisse de la gestion de contenus ou encore de la mise en place d'expériences collaboratives de production d'outils. Force est de constater que l'échange et le partage sont, au sein de ces dispositifs, des valeurs indispensables qui fondent la capacité humaine à s'approprier et contrôler les mécanismes liés à l'intelligence artificielle. Le comité de rédaction vous souhaite une année 2015 riche en projets et placée sous le signe du dialogue constructif avec votre environnement numérique, professionnel et personnel. »

22 (Pleins feux sur...)

La bibliothèque de l'École française de Rome



© Ambassade de France en Italie / Phot. Z. Colantoni

24 (Actualités)

28 (Agenda)

- 04 Écosystème numérique, culture et emploi
JOËL DE ROSNAY
- 07 Mobilearn : la réalité augmentée au service de l'apprentissage informel
ANCA AILINCAI / ELHADI DJEBBARI
- 08 Les désarrois de la Machine catalographique
PHILIPPE LE PAPE
- 10 « Faire le lien », un besoin vital
ALINE LE PROVOST
- 12 ezPaarse visite les logs
THOMAS PORQUET ET L'ÉQUIPE EZPAARSE
- 14 Des Fab Labs en bibliothèque ?
THOMAS FOURMEUX
- 16 De la PirateBox à la BiblioBox
- 17 Lorsque le robot se plante...
- 18 CultureWok : de quoi avez-vous envie ?
L'ÉQUIPE DU WoK
- 20 Opinion mining : des vertus de l'hybridation humain/machine
DOMINIQUE BOULLIER

Ar(abes)ques

REVUE TRIMESTRIELLE DE L'AGENCE BIBLIOGRAPHIQUE DE L'ENSEIGNEMENT SUPÉRIEUR
227 avenue Professeur Jean-Louis Viala CS 84308 - 34193 Montpellier cedex 5
T. 04 67 54 84 10 / F. 04 67 54 84 14 - www.abes.fr

Directeur de la publication : Jérôme Kalfon

Coordination éditoriale et secrétariat de rédaction :

Béatrice Pedot (beatricepedot@wanadoo.fr)

Comité de rédaction : Dominique Esmenjaud, Christine Fleury, Marion Grand-Démery, Philippe Le Pape, Béatrice Pedot, Maryse Picard, Marie-Pierre Roux.

Conception graphique : Anne Ladevie (anne-ladevie@orange.fr)

Impression : Pure Impression

Couverture : Thomas Edison face à un de ses projecteurs, 1915.

Wikimedia Commons / Domaine public

http://commons.wikimedia.org/wiki/File:Ed_d21m.jpg

Revue publiée sous licence Creative Commons CC BY-ND 2.0 (Paternité - Pas de modifications) sauf pour les images qui peuvent étre soumises à des licences différentes ou à des copyrights.

ISSN (papier) 1269-0589 / ISSN (web) 2108-7016

Qualité des données *et* *défis de l'automatisation*

Le catalogage partagé, les catalogues collectifs, ainsi qu'un certain nombre de services associés, sont le résultat de la contribution de chacun à un enrichissement global. Alimenter un même silo de métadonnées a permis la création et le développement du Sudoc. Il est le fruit d'une collaboration entre les hommes avec l'aide de la machine. Mais, comme la plupart des catalogues, le Sudoc a une histoire. La fusion de catalogues préexistants – dont il est le résultat – a généré des doublons. Doublons bien trop nombreux pour être résorbés manuellement, d'où la tentation de l'automatisation.

Ainsi, on a confié à la machine la « décision » de fusionner les entrées. Certains rapprochements ont été heureux, réduisant le nombre de doublons pour un même auteur. Mais, dans d'autre cas, ces automatismes sont allés trop loin, fusionnant dans une même notice d'autorité des auteurs différents.

quement. Nous nous engageons donc dans un cercle vertueux de propagation de la qualité au sein d'un réseau et en direction d'autres réseaux interconnectés. On ne quitte pas tout à fait le travail en silos, mais ceux-ci ne fonctionnent plus isolément et s'insèrent graduellement dans des réseaux.

Dans cette démarche, la machine dépasse le rôle de simple « mise en relation ». Elle permet de traiter des quantités de données hors de portée de l'homme seul, elle cesse d'être un simple instrument pour devenir un collaborateur. Outre ces



quelques exemples strictement liés à nos catalogues, le dossier présenté dans ce numéro d'*Arabesques* nous montre à quel point les catalogues cessent d'être un but en soi, pour servir largement au-delà de leur objectif originel. Nous constaterons à quel point la qualité des données est primordiale, dans la mesure où, tout comme la qualité, les défauts aussi se propagent par ricochet.

Nous nous engageons dans un cercle vertueux de propagation de la qualité au sein d'un réseau et en direction d'autres réseaux interconnectés.

Ces premiers algorithmes, par trop imparfaits, nous ont conduits à mettre fin à l'expérience. Car, autant il est possible de fusionner automatiquement, autant scinder des rapprochements malheureux est difficile. De ces erreurs, il demeure encore aujourd'hui quelques scories difficilement rattrapables sans intervention humaine, chose que l'on cherchait justement à éviter.

Depuis ces expériences qui remontent à quelques années, la puissance de calcul des machines, les progrès en matière d'apprentissage, les travaux de recherche linguistique, auxquels l'Abes est parfois associée, ouvrent de nouvelles perspectives. La détection de candidats à la fusion devient plus fine et, pour éviter les erreurs passées, une contribution humaine est jugée indispensable. La machine propose, les catalogueurs et autres intervenants disposent.

Ces améliorations de la qualité de la base peuvent être utiles non seulement au Sudoc, mais aussi, grâce aux possibilités d'alignements entre divers référentiels, servir à d'autres et récipro-

Le travail du réseau sera primordial à l'heure du traitement des métadonnées des ressources électroniques. Il s'agira de récupérer les métadonnées fournies par l'éditeur, de leur faire subir des traitements automatiques avant de les soumettre aux réseaux pour leur amélioration, puis de disséminer ces améliorations. Ici encore nous pouvons parler de cercles vertueux, à condition que les circuits de traitement, de la répartition des rôles et des étapes des *workflows* soient clairement explicités : un chantier dans lequel l'Abes entend s'engager progressivement en commençant par des expérimentations.

Les bibliothèques deviennent laboratoires, lieux d'expérimentation et d'innovation. La machine contribue à la démultiplication de l'efficacité et du pouvoir de chacun au travers de réseaux de cellules interdépendantes. Nous constaterons, à la lecture de ce numéro, l'importance des perspectives qui s'ouvrent devant nous.

JÉRÔME KALFON
Directeur de l'Abes

Écosystème numérique, *culture et emploi*

L'écosystème numérique modifie profondément la perception traditionnelle du monde du travail. Plutôt qu'une approche qui serait centrée spécifiquement sur les métiers du numérique, il s'agit d'appréhender ses répercussions globales sur l'organisation même de la société. Analyse prospective par Joël de Rosnay.

Nos sociétés industrialisées reçoivent de plein fouet le nouveau choc du futur : celui de l'écosystème numérique. La crise de société, qui se manifeste par le ralentissement de la croissance, la montée du chômage et la contestation du leadership traditionnel, traduit en fait la rapidité de la transition que nous vivons entre société industrielle et société informationnelle. Dans cette zone turbulente, entre deux mondes, émerge le pouvoir des groupes face au pouvoir centralisé des élites politiques et industrielles et la diversité des personnes face à l'anonymat des « usagers ». Ces tendances et grands courants porteurs, s'ils sont mieux compris, peuvent aider à construire avec lucidité le monde de demain.

VERS UN NOUVEAU MODÈLE DE SOCIÉTÉ

La société industrielle traditionnelle est caractérisée par la centralisation des moyens de production, la distribution massive d'objets standardisés, la spécialisation des tâches et leur contrôle hiérarchique. Son modèle structural emprunté à la géométrie ou à la mécanique est la pyramide ou l'engrenage. Tandis que les trois piliers qui fondent le contrat de travail dans l'entreprise sont l'unité de lieu, l'unité de temps et l'unité de fonction.

L'émergence de la personne est un des phénomènes les plus profonds et les plus significatifs de la société collaborative.

Avec l'extension mondiale de la numérisation, du smartphone et de l'Internet à haut débit, le modèle traditionnel vole en éclats. Aux trois unités de lieu, de temps et de fonction s'opposent la délocalisation des tâches, la désynchronisation des activités et la dématérialisation des échanges. L'écosystème numérique s'organise en réseaux plutôt qu'en pyramides de pouvoirs, en cellules interdépendantes plutôt qu'en engrenages hiérarchiques ou en filières industrielles linéaires et séquentielles. D'où le désarroi des politiques et des hauts fonctionnaires des grands corps de l'État, habitués aux évolutions quantifiables, proportionnelles et extrapolables, face au foisonnement multidimensionnel ou aux accélérations

brutales des développements par effets d'amplification.

Autre bouleversement : l'émergence des personnes. Aux nœuds de l'écosystème numérique co-évoluent désormais des acteurs diversifiés, communicants et potentiellement créateurs : les « neurones » d'un cerveau planétaire en voie d'émergence. Ce ne sont plus les « usagers » de jadis, passifs utilisateurs de services pensés par d'autres, mais des producteurs/consommateurs (*prosumers* en anglais) utilisant des nouveaux outils interactifs, démultipliant le pouvoir et l'efficacité de chacun. Cette émergence de la personne est un des phénomènes les plus profonds et les plus significatifs de la société collaborative. Désormais, c'est l'abondance de variété et de diversité que le politique va devoir gérer, lui qui s'était habitué à un univers régi par les statistiques, les probabilités et les sondages d'opinions.

DES POLITIQUES PUBLIQUES INADAPTÉES

Les politiques sont désarmées devant cette nouvelle démocratie des réseaux. Ils ont des difficultés à comprendre la culture numérique de la « génération du millénaire » (les 15-35 ans) qui est en train de changer le monde. Ce décalage est la conséquence d'un nouveau paradigme et d'un saut culturel. La pensée cartésienne, analytique, linéaire, séquentielle et proportionnelle, partagée par tant de décideurs politiques et industriels, formés aux mathématiques et au droit, appartient à l'ancien paradigme. La culture de la complexité, qui est partie intégrante du nouveau paradigme, se réfère à la pensée systémique, au multidimensionnel, au non linéaire, et intègre la dynamique due aux effets d'amplification.

L'exemple le plus marqué du décalage culturel lié à la co-évolution entre les hommes et les femmes, « augmentés » par les outils numériques et l'écosystème informationnel, est que les critères de la société industrialiste et de l'économie de marché ne s'appliquent plus. Le « noyau dur » de l'économie marchande traditionnelle assure pour un État la collecte des impôts et des taxes diverses, les prélèvements destinés à la protection sociale et aux caisses de retraites. Les « moteurs » de la croissance sont



Sebastian ter Burg / Flickr (CC BY 2.0)

➤ De nouvelles formes d'échanges se développent au sein de la société informationnelle dont l'accroissement d'ateliers collaboratifs, notamment sous forme d'hackathons. Ici, celui organisé par Wikimedia en mai 2013.

la recherche, l'innovation, le développement industriel et la diversité des marchés assurant le progrès économique et le bien-être social. La compétition et la concurrence stimulent l'économie, tandis que la croissance permet de créer les emplois nécessaires à une saine économie. Ce schéma traditionnel est resté longtemps pertinent dans un univers matériel fondé sur la production et la distribution d'objets manufacturés. Il n'est plus adapté à la société du numérique. On atteint dans l'économie classique la loi des rendements décroissants. Un accroissement considérable des efforts gouvernementaux et des investissements financiers, industriels et humains est nécessaire pour une augmentation marginale des bénéfices, des parts de marchés ou de la compétitivité économique. Telle est la dure loi des dernières années : l'accroissement de productivité due à la robotisation, à l'informatique et à Internet entraîne l'apparition de poches de chômage irréductibles. Les politiques de relance créent des effets pervers. Les « moteurs » sur lesquels agissent les gouvernements ne tournent pas en harmonie et souvent neutralisent leurs forces. On assiste à des mesures spectaculaires, mais à forte résistance sociale et aux résultats limités.

Pourtant la sphère des activités est, elle, en constant développement. La société informationnelle catalyse de nouveaux échanges, d'autres formes de transactions. Les réseaux sociaux de communication interpersonnelle voient leur densité s'accroître, amplifiant les flux des échanges immatériels. Ces activités ne sont pas toujours solvables dans l'économie classique. Elles traduisent pourtant une forte demande sociétale. Comment concilier le noyau dur de l'économie marchande née de la société industrielle et la sphère en expansion des activités

immatérielles liées à l'essor de la société du numérique ? Il faut repenser en profondeur la relation entre le temps et la nature du travail.

UN MONDE DU TRAVAIL EN PROFONDE MUTATION

Les règles traditionnelles de l'unité de lieu, de temps et de fonction bloquent l'essor de l'économie de la société informationnelle. Mais si l'on peut travailler à distance, on peut aussi travailler en temps choisi, effectuer plusieurs tâches de nature différente. On voit ainsi apparaître une nouvelle catégorie socio-économique, les « salariés libéraux » : tantôt salariés à employeurs multiples, tantôt consultants, conférenciers ou enseignants, champions des téléactivités grâce à leurs outils de communication et de traitement de l'information. Les enquêtes indiquent que les salariés sont ouverts à des formes complémentaires de rémunération. À une augmentation de salaire, certains préféreront une amélioration de la qualité de vie, une formation, une réduction du temps de travail, la création d'une « épargne-temps ». Les habitués des réseaux sociaux pratiquent déjà une nouvelle forme d'économie : le « troc d'information ». Une création originale (logiciel, texte, conseils, musique, graphisme, audiovisuel) est mise gratuitement à la disposition des utilisateurs. En retour, les créateurs sont « rémunérés » en information à plus haute valeur ajoutée. La matrice traditionnelle s'élargit. Aux deux cases du contrat de travail, temps et salaire, viennent s'ajouter de nouvelles lignes et colonnes : information, valeurs, reconnaissance... La sphère des activités « rémunérées » par ces nouvelles formes d'échange entre en expansion. C'est une des caractéristiques fondamentales de la société informationnelle. On assiste à de nouvelles formes

de troc, d'échange de marchandises, de partage, à l'essor du volontariat, du bénévolat, de l'assistance humanitaire, des mouvements associatifs. Cette co-économie, ces communautés collaboratives créent du « capital-temps » dont on peut utiliser les « intérêts » et du « capital-information » qui permet, grâce au temps investi, d'accélérer et de rendre plus efficaces des processus de travail en équipe où de nombreux modules fonctionnent en parallèle.

DES INTERACTIONS À DÉPLOYER

Une des clés du développement économique des sociétés collaboratives dans l'écosystème numérique se trouve sans doute en ce point précis de transition : dans l'impossibilité actuelle de faire croître le « noyau dur » de l'économie classique, ne faudrait-il pas tenter de le faire croître de l'extérieur, aspiré par une sphère d'activités en expansion, elle-même créatrice indirecte d'emplois ? La densité des échanges et des contacts dans l'écosystème numérique réalise justement cette relation. C'est ce qu'ont parfaitement compris les grandes plateformes numériques que l'on appelle en abrégé Gafama (Google, Apple, Facebook, Amazon, Microsoft, Alibaba) qui « emploient » leurs clients pour créer de la valeur ajoutée, ensuite revendue par eux sous diverses formes.

L'homme et les contenus des messages qui donnent du sens à son action sont au cœur des réseaux de demain. L'emploi ne résultera pas de mesures ponctuelles incompatibles les unes avec les autres. Mais d'un changement des relations entre le temps, l'espace et le travail. Il convient de favoriser tout ce qui accroît la densité des interactions et des transactions : désynchronisation des tâches par le temps partiel, le temps partagé, l'épargne-temps,

le temps choisi ; délocalisation des activités par la réduction des coûts de communication, la démocratisation et la simplification de l'usage des smartphones et des tablettes, la réorganisation des lieux de travail (bureaux mobiles, entreprises virtuelles) ; diversification des fonctions, par l'aide aux entrepreneurs, la création de pépinières d'entreprises, les avantages accordés aux « salariés libéraux », la réforme du contrat de travail ; valorisation de l'expérience des plus âgés, afin qu'ils puissent transmettre leurs connaissances, grâce à la co-éducation intergénérationnelle, voire aux Moocs (*Massive Online Open Courses*, ou cours de masse en ligne gratuits).

Le défi de la société informationnelle est la remise en cause des politiques traditionnelles d'aide à l'emploi. Le politique ne doit plus craindre la diversité, mais au contraire la favoriser. Il ne peut la contrôler, mais peut en revanche « catalyser » l'émergence des potentialités de chacun. Toute la question de la transition entre la société industrielle et la société informationnelle est contenue dans cette alternative : soit poursuivre l'exercice (parfois solitaire) de l'intelligence *élective*, soit favoriser la pratique solidaire de l'intelligence *collective*. La complexité ne se réduit pas à quelques éléments simples par l'analyse cartésienne. Elle se construit au contraire par l'action simultanée de personnes responsables, informées et créatives. La réussite de la grande transition du XXI^e siècle se fera à ce prix : celui de la responsabilisation des « neurones » du cerveau planétaire.

JOËL DE ROSNAY

Scientifique, prospectiviste, écrivain
Conseiller de la Présidence d'Universcience
Président exécutif de Biotics International

• • • UN FORUM POUR CHANGER D'ÈRE

Créé en 2013 à l'initiative de Véronique Anger-de Friberg, rédactrice en chef de la publication en ligne *Les Di@logues stratégiques*¹, avec le concours de Joël de Rosnay, le **Forum Changer d'Ère** réunit des intellectuels, des scientifiques, des analystes du changement, des décideurs économiques, des entrepreneurs et des chercheurs de la jeune génération, pour aborder autrement les grands défis économiques, technologiques et sociétaux d'aujourd'hui avec l'ambition d'inspirer un nouveau modèle qui permette de faire face, ensemble, aux grands enjeux de la globalisation et de la civilisation numérique.

Une rencontre annuelle, organisée sous forme de sessions collaboratives, se tient à la Cité des sciences et de l'Industrie. La prochaine est d'ores et déjà programmée le **3 juin 2015** autour

du thème « **Au-delà de la révolution numérique (nouveaux modes de pensée, nouvelles technologies, nouvelles visions, nouvelles organisations)** ».

Tout au long de l'année, en marge du forum et en association avec *Les Di@logues stratégiques*, les émissions diffusées en direct sur Internet, interrogent et informent sur les grands défis de la modernité, les effets sociétaux des mutations engendrées par le numérique, avec des regards croisés de penseurs, d'entrepreneurs, de décideurs et de chercheurs.

POUR EN SAVOIR PLUS : www.forumchangerdere.fr

[1] <http://lesdialoguesstrategiques.blogspot.fr>



Visuel du Forum 2014.

● ● ● MOBILEARN : LA RÉALITÉ AUGMENTÉE AU SERVICE DE L'APPRENTISSAGE INFORMEL

Universcience est un pôle culturel majeur œuvrant pour la diffusion de la connaissance scientifique. La Cité des Sciences expérimente des formes de médiation exploitant les nouvelles technologies depuis sa création.

Mobilearn s'inscrit dans cette évolution. Il s'appuie, d'une part, sur les smartphones (Android et iPhone), à la synthèse de technologies multiples, et, d'autre part, sur l'arrivée à maturité de différentes techniques de localisation : wifi, RFID (*Radio Frequency Identification*), NFC (*Near Field Communication*)... L'objectif est de toucher le public le plus large et de proposer une expérience forte au sein de la bibliothèque.

La Bibliothèque des sciences et de l'industrie (BSI) dispose d'une offre riche et diversifiée : le fonds documentaire, la médiation, les services pratiques... À côté de cette offre, nous pouvons considérer les visiteurs comme une ressource potentielle. Chacun dispose en effet de connaissances ou expériences dans différents domaines susceptibles d'être partagées. L'objectif est de valoriser ce potentiel à travers des services et fonctionnalités proposés par Mobilearn.

Comment ça marche ?

Le dispositif exploite la connexion wifi locale, sans ajout de bornes supplémentaires, ainsi que les qr-codes équipant les documents pour leur gestion ou à défaut les codes-barres d'identification. Mobilearn s'adapte aux infrastructures et équipements existants pour limiter son impact écologique. Il repose sur un CMS (*Content Management System*) basé sur des formats et protocoles ouverts (HTML5, JSON...) dans l'objectif d'une adaptabilité et interopérabilité avec des contextes variés.

La BSI propose un territoire ouvert à l'exploration des domaines scientifiques. L'ancrage de la connaissance dans des zones thématiques et des ouvrages offre un contexte privilégié pour l'apprentissage informel. Cette situation appelle la mise en œuvre d'un outil de médiation qui accompagne le visiteur dans son accès à la connaissance. La géolocalisation indoor, basée sur le wifi et le bluetooth, présente une opportunité indéniable dans ce contexte.

Créer une communauté de visiteurs

Mobilearn favorise le développement d'une communauté de visiteurs reposant sur le partage des savoirs et l'entraide. Son équipe, composée à parité de membres d'Orbe¹ et de la BSI, souhaite constituer une expertise pour la transmettre aux autres lieux culturels à travers des publi-

cations, des ateliers ainsi qu'à un accompagnement lors d'une transposition du dispositif.

Cet outil, offert au visiteur par la bibliothèque, l'accompagnera durant chaque étape de l'accès à la connaissance. Il pourra ainsi :

- **chercher** de l'information, un document, un service ;
- **isoler** l'élément recherché ;
- **consulter** le document, accéder aux données associées ;
- **comprendre** le discours, les concepts et échanger ;
- **enregistrer** les références, les données associées ;
- **partager** via un espace de travail collectif, en proposant des conseils, un accompagnement.

Mobilearn permet la collecte de données d'usage riches et diversifiées. Ces données anonymes peuvent être centralisées sur serveur pour croisement et analyse. Elles concernent notamment : les pratiques de l'espace, les cheminements, l'utilisation des espaces de travail ; la consultation du fonds accompagné d'un retour qualitatif (*rating*, commentaires) ; l'utilisation des services du guide (options et fonctionnalités privilégiées).

Du tangible au numérique

Mobilearn est un véhicule vers les territoires de la connaissance que constituent les collections des bibliothèques et des musées. L'objectif est de revaloriser par la réalité augmentée l'inscription tangible des savoirs dans des objets, des

documents et des espaces. La participation du corps à l'apprentissage par la déambulation, la manipulation, l'exploration permet en effet un *ancrage* plus profond des connaissances. Mobilearn hybride le numérique et le tangible et ajoute ainsi une couche d'interface familière sur des espaces complexes.

Les musées et bibliothèques sont des lieux d'apprentissage informel. Ils permettent à tout un chacun d'explorer de nouveaux territoires de la connaissance. Mobilearn accompagne cette démarche d'apprentissage individuel au sein des lieux culturels et après la visite. Le visiteur construit ainsi son territoire de la connaissance qui se formalise progressivement sur une carte. Il peut approfondir ou agrandir ce territoire et projeter de nouvelles conquêtes.

ANCA AILINCAI ET ELHADI DJEBBARI

Cité des sciences et de l'industrie

[1] Orbe réunit des programmeurs, designers, artistes et scientifiques autour de l'expérimentation et la mise en œuvre de nouvelles expériences impliquant le corps et les nouveaux médias.



LES DÉSARROIS DE LA MACHINE CATALOGRAPHIQUE

Rien ne va plus pour la Machine catalographique. Après des décennies de bonheur, la voici mal à l'aise en société, gênée dans son couple...



LA MACHINE SUR LE DIVAN

– Les gens vous comprennent, ils me posent toute sorte de questions. Ils me demandent mettons « dernier livre de modianeau », ou bien « qui a écrit la cousine bete », j'ai eu ça dernièrement. Ça me coupe tous mes moyens des questions comme ça, voyez, je ne dis rien, je suis sans voix. « Pas de réponse » je dis. Mais les gens m'insultent figurez-vous, si vous saviez ce que j'entends des fois ! Ou alors « dvd pasolini ». Alors là oui, là j'en ai des choses à dire ! Ça j'aime, des questions comme ça. Là je déballe tout. « 18 résultats », je dis, et je les sors en moins d'une seconde. Mais là non plus, ils sont pas contents les gens, ils disent qu'ils s'y retrouvent pas, que tout est en vrac. C'est comme Google, ils disent, mais en moins bien. Google il répond « environ 673.000 résultats (0,29 secondes) ». Moi 18 résultats. Je peux pas lutter voyez, c'est pas possible. On joue pas dans la même division comme on dit. Et puis sur Google après ils peuvent naviguer d'un truc à l'autre, ils peuvent y passer la nuit s'ils veulent. Alors qu'avec mes 18 résultats, ils peuvent pas aller plus loin, quoi, ça s'arrête là comme sur un mur. En plus ils disent que c'est du charabia ce que je leur raconte. Moi je voudrais bien leur parler autrement, mais ces choses-là je ne peux pas les inventer que voulez-vous, la plus belle fille du monde, elle peut donner que ce qu'elle a comme on dit hein, moi je peux donner que ce que j'ai. Je donne ce que j'ai, c'est tout.

– Je vous le redis Madame, il faut en parler à votre partenaire. En toute franchise.

LA MACHINE À LA MAISON

– Écoute...

– Tu ne vas pas recommencer...

– Écoute-moi. Cette fois je suis décidée. Non ne me coupe pas la parole. Il faut que tu me parles sur un autre ton désormais, il faut que tu me parles autrement. Les gens, ils comprennent plus ce que tu me fais dire, voilà ce qu'y a. Pour eux c'est du chinois ou je ne sais quoi, quelque chose qui ne les concerne pas, oui voilà, c'est ça. Ils ne me regardent même plus, ils me trouvent ringarde. Tu comprends, tu me dis *Lettres de la religieuse portugaise [Texte imprimé] / calligraphies, Stéphanie Devaux*. Tu me dis que c'est ce qu'il faut que je dise moi aussi. Tu me dis *Paris : Éd. Alternatives, 2004*, et je répète *Paris : Éd.*

Alternatives, 2004 quand on me demande. Je dis comme toi. Tu me dis 77p. : *ill. en noir et en coul., couv. ill. en coul. ; 19 cm*, et je répète ça. Tu me dis (*Pollen*), je dis (*Pollen*). Mais ils comprennent pas ce truc entre parenthèses, c'est quoi, ça veut dire quoi « *Pollen* » entre parenthèses ? Et 77 p point, ça veut dire quoi ? Et ça, « *ill. en coul., couv. ill. en coul.* » ? Ça ne peut plus durer, tu comprends. Et puis « dvd pasolini 18 résultats » ça ne va pas non plus. Tes notices comme tu dis c'est des culs-de-sac, une fois qu'ils y sont les gens ils sont obligés de revenir en arrière, en plus t'as une notice pour *Uccellacci e uccellini* et une autre pour *Des oiseaux petits et gros* alors que c'est le MÊME film, si c'est vrai, y en a DEUX, tu n'as qu'à regarder toi-même !

– Comme tu as changé... Ça fait combien de temps qu'on se connaît toi et moi ? Tu t'en souviens sûrement mieux que moi. Au début je n'ai pas fait grand cas de toi, je dois dire. Je ne t'aimais pas, non. Comment tu marchais, comment tu parlais surtout... En plus, on ne s'adressait à toi qu'avec tes mots à toi. Tu dis que ce sont les miens, mais ils ne font que se plier à la syntaxe de ta langue à toi. Une langue d'une très grande subtilité quant à elle, qu'on acquiert au prix d'un rude et long apprentissage dont peu triomphent. Je suis de ceux-là. J'ai parlé, je parle toujours, ta langue de Machine catalographique à la perfection, sans accent, au point de passer moi-même sans difficulté pour une Machine catalographique. Tout ce que j'ai pu te dire, dans cette langue-là ! Il y a longtemps que je te parle, à toi, rien qu'à toi, dans de longs tête-à-tête quotidiens. Et maintenant, tu me dis que ce sont « les gens » qui t'intéressent. C'est à eux que tu veux plaire. C'est avec eux que tu veux parler, plus avec moi. Tu veux parler leur langue aussi ? Ces mots, illencoul, couvillencoul, d'une telle douceur, tu ne veux plus les entendre de moi. Leur charme s'est épuisé pour toi, je le vois bien ; ils ne t'enchantent plus. Tu as changé. Comme tu as changé !

Au fond je me rends compte d'une chose affreuse. C'est-à-dire que, depuis tout ce temps, je parle devant toi comme devant un hygiaphone, comme les types des guichets pour qui la personne en face n'a pour ainsi dire pas de réalité. Longtemps je t'ai parlé, à toi, sans le moins du monde voir au-delà de toi. Jusqu'à comprendre que tu n'es personne.

PHILIPPE LE PAPE

Chargé de mission Normalisation, Abes
le-pape@abes.fr

● ● ● PETITE BIOGRAPHIE DE LA MACHINE CATALOGRAPHIQUE

Les années 1960 : la création du monde

Les années 1960 sont celles de la création, au niveau international, d'une série de standards qui vont durablement imprégner les catalogues de bibliothèques :

- les *Principes internationaux de catalogage* (les « Principes de Paris »), Ifla, 1961, régissent le fonctionnement général d'un catalogue ;
- les *Règles de catalogage anglo-américaines* (*Anglo-American Cataloguing Rules*, AACR), 1967 ;
- l'ISBD (*International Standard Book Description*, puis *International Standard Bibliographic Description*), limité à la description normalisée, élaboré dans le cadre international de l'Ifla à partir de 1969.

Tous ces standards ont été établis pour des catalogues de livres et portent en eux ce péché originel.

Les formats MARC (vers 1966-....) : un langage de balisage et une norme de contenu

L'ISBD introduit déjà un embryon de « langage-machine » par la segmentation de la notice en « zones » et « sous-zones » ordonnées.

Le format Marc¹, futur USMarc, puis Marc 21, élaboré à la Bibliothèque du Congrès, encode les éléments de la notice normalisée, description et accès, en champs et sous-champs affectés d'indicateurs de traitement. Des jeux de données codées organisées en vocabulaires propres au format permettent de rendre explicites certaines informations implicites de la notice normalisée (la langue, la forme du contenu de la ressource, etc.). En cela le format Marc, en plus d'un langage de balisage, est aussi dans une certaine mesure un standard de contenu.

L'Unimarc : déjà un modèle entité-relation ?

L'Unimarc, conçu dans le cadre de l'Ifla à l'instar de l'ISBD, est publié en 1977. Il présente une granularité plus fine que le Marc de la Bibliothèque du Congrès et fait un large emploi des données codées, ce qui le rend moins dépendant des langues de catalogage. On y voit déjà l'intuition qu'une ressource représentée par une notice Marc est une entité potentiellement liée à d'autres. À mesure que le format s'enrichit de nouveaux types d'entités, le système des relations se complexifie. Les liens sur identifiants se généralisent : à défaut de modèle, on peut indéniablement parler d'un système entité-relation.

Les années 1990-2000 : un mai 1968 tardif ?

Le catalogage à la papa est sérieusement remis en cause. Les anciens standards conviennent mal pour les ressources électroniques notamment. En outre, les agences bibliographiques nationales, dans un but d'économie, aimeraient que soient définis les éléments d'une notice de base. Commandé en 1989, le *Rapport sur les fonctionnalités requises des notices bibliographiques* (FRBR²) est publié en 1998. Fondé sur les besoins des utilisateurs, il contient un modèle entité-relation. Il sera complété en 2009 par les FRAD³ pour les données d'autorité, et en 2010 par les FRSAD⁴ pour la relation de sujet. En 2009 le modèle FRBR est publié dans un formalisme « orienté objet » en tant qu'extension du modèle Cidoc CRM (Cidoc, *Conceptual Reference Model*) conçu pour les musées, sous le nom de FRBRoo.

Les « Principes de Paris » sont mis à jour : publication en 2009 du *Statement of International Cataloguing Principles*⁵ (ICP, ou PIC en français).

Les années 2010 : vers la transition bibliographique

- 2010 : publication du code de catalogage RDA (*Resource Description and Access*), évolution des AACR à la lumière des nouveaux modèles et standards. Début du travail sur RDA en France.
- 2011 : ouverture de l'*Open Metadata Registry* dans lequel sont publiés les nouveaux standards sous forme de schémas de métadonnées dans le cadre du web sémantique⁶.
- 2012 : le modèle de données BibFrame, destiné à remplacer Marc, est lancé à la Bibliothèque du Congrès.

PHILIPPE LE PAPE

[1] Marc : acronyme de *Machine Readable Catalog*, « catalogue lisible par une machine ».

[2] Ifla Study Group on the Functional Requirements for Bibliographic Records, *Functional Requirements for Bibliographic Records : final report*, 1998. Révisé en 2009.

[3] Ifla Working Group on Functional Requirements and Numbering of Authority Records (FRANAR), *Functional Requirements for Authority Data : a conceptual model*, 2009.

[4] Ifla Working Group on the Functional Requirements for Subject Authority Records (FRSAR), *Functional Requirements for Subject Authority Data (FRSAD) : a conceptual model*, 2010.

[5] Ifla Cataloguing Section and Ifla Meetings of Experts on an International Cataloguing Code, *Statement of International Cataloguing Principles*, 2009.

[6] L'OMR est une initiative du DCM1 (*Dublin Core Metadata Initiative*).

Depuis la création du Sudoc, l'Abes et ses réseaux ont compris, comme d'autres, les bienfaits des liens bibliographiques. Un aperçu de la démarche Qualinca pour résoudre les problèmes inhérents au liage dans les grandes bases de données documentaires : une véritable complémentarité entre l'homme et la machine.

« Faire le lien », un besoin vital

Traditionnellement, les liens permettent le contrôle d'autorité et la navigation dynamique de notices en notices. Avec le modèle FRBR et l'ouverture des catalogues sur le web de données, leur importance ne cesse de croître. Pour la FRBRisation, l'exploitation de certains liens facilite la mise au jour des entités œuvre et expression¹. En matière d'ouverture, IdRef avait déjà émancipé les ex-« autorités Sudoc » en leur permettant d'être utilisées et alimentées par d'autres applications. Quant aux alignements à grande échelle, ils se nourrissent des liens bibliographiques sur un mode ricochet : entre deux entités sur le web de données, plusieurs entités peuvent se donner la main. La qualité des liens internes au catalogue est primordiale car elle se répercute bien au-delà d'une seule base de données et d'un seul usage.

DÉFINIR ET MESURER LA QUALITÉ DU LIAGE

Appuyés par des théories et méthodes en intelligence artificielle, les chercheurs du projet Qualinca, parmi lesquels des membres de l'équipe GraphIK (voir encadré), s'aventurent sur le terrain des grandes bases de données documentaires comme le Sudoc pour résoudre les problèmes de liage. La réconciliation d'entités dans les bases de données a fait couler beaucoup d'encre depuis la fin des années 1950². Il s'agit de confier à une machine le soin de dire si des descriptions d'entités (une personne, une collectivité, une œuvre, etc.) décrivent ou non la même entité « réelle ». Il existe plusieurs approches, dont celle à base de connaissances dans laquelle s'intègre le projet. L'objectif est de disposer de connaissances formalisées sur lesquelles appliquer des opérations logiques censées simuler les opé-

rations humaines. *In fine*, on aimerait pouvoir repérer et caractériser les problèmes de qualité des liens et disposer de schémas de réparation.

DES ERREURS DE LIAGE, MAIS POURQUOI ?

Les raisons pour lesquelles les catalogueurs peuvent faire des erreurs de liage sont multiples : doublons d'autorité, ambiguïté des autorités (informations associés insuffisantes), erreurs de liage préexistantes, absence d'autorité. D'autres erreurs ne sont pas humaines, car pendant plusieurs années un algorithme de liage a fonctionné sur la base Sudoc. Cet outil comparait la chaîne de caractères d'une appellation (nom, prénom) dans une notice bibliographique avec les formes retenues des appellations dans les notices d'autorité. Problème : la machine n'intégrait pas qu'une forme retenue pouvait être constituée d'une appellation accompagnée de qualificatifs (dates de vie, fonction). Fatalement, dans les cas d'homonymie, les autorités dont la forme retenue n'avait pas de qualificatif ont été les seules à être liées par cet algorithme. Autre inconvénient : si cet outil ne faisait pas de lien lorsqu'il était confronté à plusieurs possibilités de liage (ce qui est bien), il ne capitalisait pas cette hésitation (ce qui est dommage). Avec un système à base de connaissances et de règles, la collaboration humain/machine, alors absente, sera fortement mise en valeur.

NOURRIR LA MACHINE : TRANSFORMER UN CATALOGUE EN BASE DE FAITS

Le choix des données est le premier pas vers la construction du système. Le travail de l'analyste consiste à identifier dans les notices ce qui est utile à un catalogueur dans une décision de liage. L'environnement normatif (ISBD, Unimarc, etc.) et

● ● ● LE PROJET GRAPHIK

Créé en 2010, GraphIK (*Graphs for Inferences on Knowledge*) est une équipe commune de l'Inria (Sophia Antipolis), de l'université de Montpellier 2, du CNRS et de l'Inra. Leurs travaux reposent sur la représentation des connaissances et les moyens de raisonner à partir de ces dernières via des approches logiques.

Cette équipe de recherche a développé une bonne connaissance des données et des problématiques de l'Abes, avec qui elle a déjà travaillé (projet SudocAD, 2011).

<https://team.inria.fr/graphik>

[1] Dont on peut voir les résultats dans des applications comme data.bnf.fr.

[2] Depuis l'article fondateur : H. Newcombe, et al, « Automatic Linkage of Vital Records », *Science*, 1959, vol. 130, p. 954-959.

l'historique de construction et d'alimentation du Sudoc imposent une connaissance experte du catalogue, afin de ne pas surinterpréter ou mal interpréter le sens des données, c'est-à-dire l'information qu'elles véhiculent. En résumé, qu'une valeur existe dans le catalogue ne signifie pas qu'elle soit exploitable pour n'importe quel usage.

Il faut ensuite se départir du format source (Unimarc) pour transformer le catalogue en base de connaissances. Pour cela, on fait appel à une ontologie, FRBRoo³ pour ce qui concerne le Sudoc⁴. Là encore, c'est l'humain qui, interprétant la grammaire de l'ontologie, modélise les données sélectionnées. L'attribution d'un identifiant à chaque entité nommée mentionnée dans une notice bibliographique achève ce travail de formalisation.

FABRIQUER LE SYSTÈME D'INTERPRÉTATION

À ce stade, la machine peut, en effectuant des requêtes, isoler des « grains de connaissance » et les attribuer à des références⁵ (voir graphique ci-contre), entre lesquelles des liens potentiels existent. Ces attributs constituent le matériau du raisonnement humain : le catalogueur les combine, les questionne, les compare et, en associant cela à ses propres connaissances, décide de lier, de corriger un lien ou de s'abstenir de lier. Décortiquer ces raisonnements pour les reproduire en machine est un défi majeur du projet. Distinguer des attributs permet d'alimenter la grille d'interprétation du système. Celle-ci comprend des critères, qui comparent des attributs ou combinaisons d'attributs, et des règles, qui combinent les résultats des critères en un résultat d'identification (rapprochement de références) ou de différenciation (éloignement de références). Ce modèle permet de disposer d'un système capable de prendre des décisions de liage/non-liage justifiées et contextualisées. Par ailleurs, ce système n'est pas figé : il bénéficie d'améliorations itératives *via* des ajustements de paramétrage.

De l'intuition à la règle

• Intuition métier

> « Un contributeur a tendance à republier avec les mêmes co-contributeurs ».

• Traduction règle formelle

> Si X a publié avec Y, et si X a publié un autre livre avec Z qui a le même nom que Y, alors il est quasiment certain que Z = Y.

• Explication

On compare la référence X avec la référence Y. Les attributs impliqués sont l'appellation et les co-contributeurs. Deux résultats de critères sont combinés dans la règle de liage : « identique » pour l'appellation et « identique » pour co-contributeur. Le système décide, selon cette règle, que le liage entre X et Y est quasiment certain.



➔ Du catalogue au graphe : identifier les références et formaliser les relations qu'elles entretiennent avec les entités (documents, personnes, concepts...) et les valeurs (noms, dates, titres...) du catalogue.

FAIRE CONFIANCE À LA MACHINE : OUI, MAIS PAS AVEUGLÉMENT !

Dans la démarche Qualinca, le diagnostic qualité sollicite fortement la collaboration de l'humain et du système automatique. Ce dernier sera capable de proposer des rapprochements ou des éloignements de référence plus ou moins forts. L'établissement de seuils de confiance est la clé de la réussite. L'analyse humaine intervient encore une fois en ajoutant une heuristique, c'est-à-dire un système de règles contrôlées qui valide ou non la décision brute. Réparer un catalogue de l'intérieur serait ensuite possible, en se fondant sur des combinaisons de décisions fiables de liage/non-liage. Ces scénarios modéliseront des types de problèmes détectables par le système et dont les modalités de corrections auront été décrites. Plus proche du travail quotidien, l'exploitation de décisions de liage pour l'aide à la décision dans une application comme IdRef est également envisagée. À n'en pas douter, la philosophie de la zone 309⁶, qui permet aux mains expertes du réseau de prendre le relais lorsque les traitements de masse se sont épuisés, a de beaux jours devant elle.

ALINE LE PROVOST

Abes
le-provost@abes.fr

[3] FRBRoo est une ontologie formelle destinée à capturer et représenter la sémantique sous-jacente de l'information bibliographique et de faciliter l'intégration, la médiation et l'échange de l'information bibliographique et muséographique : www.cidoc-crm.org/frbr_inro.html

[4] L'Institut national de l'audiovisuel, second partenaire du projet en gestion documentaire, a choisi de développer une ontologie maison pour ses propres données.

[5] Une référence est une entité nommée dans une notice (bibliographique ou d'autorité).

[6] À ce sujet, lire : <http://pункtokomo.abes.fr/2014/05/02/une-zone-309-pour-coordonner-le-travail-collectif-sur-la-qualite-des-donnees-sudoc>

ezPaarse visite les logs

Issu d'une collaboration étroite depuis novembre 2012 entre l'Inist-CNRS, Couperin.org et l'université de Lorraine, ezPaarse est un logiciel libre¹ et multiplateforme qui, à partir de l'exploitation des logs, permet aux bibliothèques de disposer de statistiques d'usage (qu'est-ce qui est consulté et par quelle catégorie d'utilisateurs), de les comparer aux données fournies par les éditeurs et de proposer des outils d'aide à la décision. Décryptage.

ezPaarse se présente sous la forme d'une application web² disposant d'un formulaire et d'une interface de programmation (API) pour l'ingestion manuelle ou automatique de fichiers de logs d'accès aux ressources numériques proposées par les plateformes des éditeurs de littérature scientifique. Cette application est complétée par un portail wiki dénommé AnalogIST³ où, en sus d'informations sur le projet, est publié le travail d'analyse des plateformes⁴ des fournisseurs qui sert de base à la création des parseurs pour enrichir les capacités de reconnaissance d'ezPaarse. Un parseur est un petit programme, spécifique à chaque plateforme, qui se charge de découper une ligne de log pour en extraire des informations (par exemple : l'ISSN de la revue consultée, le type de ressource, etc.). Les contributions extérieures sont encouragées pour augmenter la couverture d'ezPaarse : toutes les personnes intéressées peuvent facilement analyser une nouvelle plateforme en s'appuyant sur les outils⁵ et procédures⁶ mis à disposition. Les plateformes en ligne sont uniques, même s'il peut exister entre elles des régularités ou des ressemblances. Chacune d'elles a donc besoin d'être spécifiquement analysée par un collègue, bon connaisseur de la ressource, qui saura dresser un catalogue complet des actions qu'un usager peut y accomplir dans le cadre de l'abonnement souscrit et faire correspondre des éléments des URL à des informations sur la ressource consultée.

ENJEUX ET BASES DE FONCTIONNEMENT

ezPaarse tente de répondre au besoin exprimé par les bibliothèques et leurs tutelles de disposer d'outils pour justifier et évaluer la bonne utilisation des sommes importantes consacrées aux abonnements académiques et scientifiques. Les données fournies par les éditeurs ne suffisent souvent pas à atteindre ces buts parce qu'elles :

- ne sont tout simplement pas disponibles ;
 - sont disponibles, mais non certifiées Counter⁷, ce qui les rend peu fiables et difficiles à comparer d'une ressource à l'autre ;
 - sont disponibles et certifiées Counter, mais doivent être nettoyées et enrichies pour en tirer des informations exploitables, par exemple sur les profils des utilisateurs.
- Pour fonctionner, ezPaarse s'appuie sur un maté-

riau courant, mais encore peu systématiquement exploité : les fichiers logs des serveurs proxy. On y trouve, ligne par ligne, le détail des accès réalisés par les utilisateurs accrédités aux ressources numériques. C'est l'analyse de ces fichiers qui permet de mesurer l'activité de consultation des ressources numériques et de disposer d'informations précises sur les événements de consultation identifiés. Une fois correctement réglé, ezPaarse peut être automatisé et délivrer des rapports complets sur une base régulière. Cette mise en production est d'ores et déjà effectuée à l'Inist-CNRS et à l'université de Lorraine et en cours dans une vingtaine d'autres établissements.

UN FONCTIONNEMENT MUTUALISÉ ET COLLABORATIF

Un fonctionnement agile pour l'implémentation de nouvelles fonctionnalités

Les méthodes agiles sont des groupes de pratiques de développement en informatique (conception de logiciel) pouvant s'appliquer à divers types de projets. La méthode agile Scrum⁸ utilisée pour développer et faire évoluer ezPaarse permet de livrer régulièrement et rapidement, toutes les six semaines en moyenne, des versions fonctionnelles. Les retours des utilisateurs sont immédiatement pris en compte et orientent sans cesse le projet vers les besoins constatés. On peut dresser un parallèle entre les itérations agiles et une série télévisée : les épisodes se succèdent et sont regroupés en « saisons ». L'intérêt des utilisateurs est maintenu et leur implication grandit au fil des développements. Dans le cas d'ezPaarse, la « saison 1 » a duré un an : la version 1.0 du logiciel a été livrée en 10 itérations (versions 0.1 à 1.0) thématiques. La « saison 2 » se terminera en décembre 2014 et sera suivie en 2015 d'une troisième et dernière saison sur la thématique des statistiques de l'*open access* et des archives institutionnelles.

Une mutualisation des efforts et les moyens d'y parvenir

Des outils collaboratifs ouverts sont présents à tous les endroits pertinents, depuis l'analyse des plateformes, à travers :

- un portail wiki : AnalogIST (qui permet la participation sur simple inscription) ;

[1] C'est la licence CeCILL (www.cecill.info/licences/Licence_CeCILL_V2-fr.html), compatible avec la GPL (General Public Licence) en droit français, qui a été choisie. Le code source est disponible en ligne sur : <https://github.com/ezpaarse-project/ezpaarse>

[2] Une instance publique du logiciel est utilisable sur <http://ezpaarse.couperin.org> sur simple inscription.

[3] <http://analogist.couperin.org>

[4] Voir la liste complète des plateformes en cours d'analyse ou analysées sur : <http://analogist.couperin.org/platforms/start>

[5] <https://trello.com/b/wEaLnz8d/ezpaarse-analogist-suivi-des-plateformes> et <http://ezpaarse.couperin.org/doc/tools.html#platform-init>

[6] <http://analogist.couperin.org/platforms/contribute/start>

[7] Voir www.projectcounter.org et <http://couperin.org/groupes-de-travail-et-projets-deap/statistiques-dusage/counter>

[8] [https://fr.wikipedia.org/wiki/Scrum_\(methode\)](https://fr.wikipedia.org/wiki/Scrum_(methode))

- un outil de collaboration en ligne en temps réel : un *board* Trello⁹ pour le suivi fin de l'état d'avancement des analyses ;
- une liste de diffusion publique¹⁰ qui intègre la participation à la création de parseurs, ou même la contribution au code source, grâce à un répertoire sur Github¹¹ qui permet la récupération libre de la totalité du code du logiciel, mais aussi la participation à de nouvelles fonctionnalités¹². En plus des outils décrits ci-dessus, nous tentons d'assurer une présence active sur Twitter¹³, Google+¹⁴ et Facebook¹⁵ pour diffuser rapidement et simplement des informations d'actualité, techniques ou non, sur les avancées du travail et informer sur les événements auxquels nous participons. Nous accordons beaucoup d'importance aux différents événements, formations et ateliers pour faire découvrir ce projet et y impliquer les collègues. L'atelier organisé lors du BibCamp ADBU (juin 2014) aura été une bonne occasion pour faire connaître ezPaarse et l'environnement AnalogIST, présenter l'équipe et la structure mutualisée du travail et, surtout, entamer et parfois finaliser des analyses de plateformes dont les parseurs ont depuis été développés. Sur le même modèle, nous souhaitons organiser des Hackathon/Parsathon/ezCamp pour sensibiliser les professionnels à l'outil et au processus d'analyse des plateformes. Deux journées, dont les détails seront communiqués bientôt, devraient avoir lieu en mars puis en juin 2015.

Des ouvertures internationales

Les préalables techniques pour utiliser ezPaarse sont très basiques et notre logiciel peut être adopté par une large communauté. Nous l'avons donc présenté lors de conférences et de salons et notamment à Liber (Riga), où un « *Liber Award for Library Innovation* » nous a été décerné, à l'Ifla (Lyon), à l'Icolc (Lisbonne) et prochainement lors de la conférence UKSG, en mars 2015.

UNE DYNAMIQUE NATIONALE ET INTERNATIONALE

Il existe des articulations possibles entre ezPaarse, la base de connaissance nationale Bacon et le hub de métadonnées de l'Abes, qui ont été discutées et délimitées lors d'une réunion commune des équipes concernées, début septembre 2014. ezPaarse collecte actuellement des données sur les plateformes éditeurs et met en forme des bases de connaissance qui pourraient être reversées dans Bacon, après nettoyage et enrichissement par le hub. Comme cela apparaît sur le schéma ci-contre, ezPaarse peut progressivement passer d'une position de producteur à celle de consommateur/contributeur en s'abreuvant à Bacon, mais aussi à GoKB (*Global Open Knowledgebase*), rejoignant ainsi une initiative internationale très prometteuse.



phot. Stéphane Gully

➔ L'atelier ezPaarse lors du BibCamp ADBU en juin 2014.

Successes d'un projet interne à l'Inist-CNRS, le logiciel ezPaarse et l'environnement AnalogIST sont ouverts à la communauté, que ce soit pour une simple installation ou une collaboration plus avancée, depuis maintenant deux ans. ezPaarse est utilisé en production dans plusieurs établissements dont l'université de Lorraine est le pilote. L'équipe du projet se charge d'un soutien technique très réactif et se fixe l'objectif de fédérer une communauté d'utilisateurs et de contributeurs pour que le logiciel puisse poursuivre sa route à l'issue de la phase 3, fin 2015.

THOMAS PORQUET
 Chargé de mission, Couperin
 thomas.porquet@couperin.org
 et L'ÉQUIPE EZPAARSE
 ezparse-team@googlegroups.com

[9] cf. note 5.

[10] <https://groupes.renater.fr/sympa/info/ezparse-contribute>

[11] <https://github.com/ezparse-project>

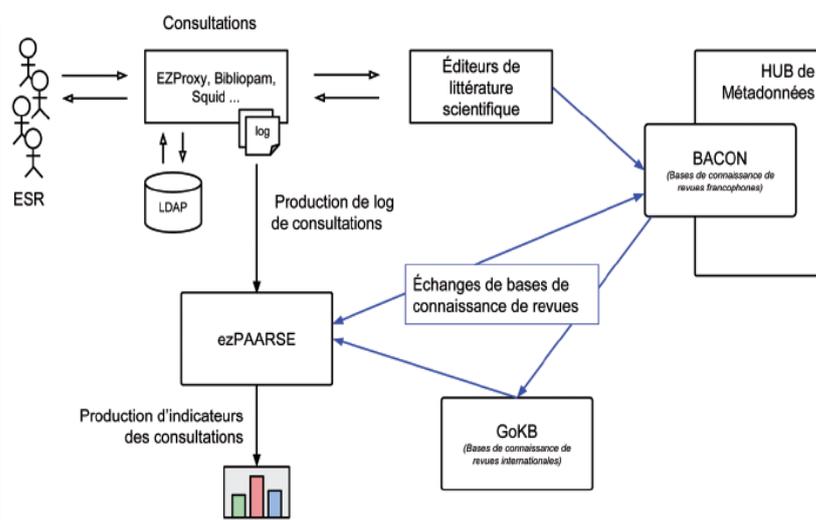
[12] Et cela a été récemment le cas avec une contribution inattendue des bibliothèques de l'université du Minnesota :

<https://github.com/UMNLibraries>

[13] <https://twitter.com/ezparse>

[14] <https://plus.google.com/+ezPAARSE>

[15] www.facebook.com/Ezparse



➔ Articulations entre ezPaarse, Bacon et le hub de métadonnées de l'Abes.

En quoi Fab Labs et bibliothèques partagent-ils des points communs ? Dans quelle mesure leurs missions et les objectifs qu'ils se donnent sont-ils complémentaires ? Approche et enjeu par un bibliothécaire engagé.

Des Fab Labs en bibliothèque ?

Des États-Unis à l'Europe, en passant par l'Afrique, nous assistons à l'émergence de nouveaux lieux ouverts sur leur territoire, accessibles à tous, avec une dimension collaborative et sociale très forte. Si ces éléments font penser aux caractéristiques d'une bibliothèque, ce n'est pas de cela qu'il est question. Il s'agit des nouveaux lieux de fabrication numériques regroupés généralement sous l'appellation « Fab Lab ». Nous reviendrons dans un premier temps sur la notion de Fab Lab. Puis, nous comparerons bibliothèques et Fab Labs du point de vue de leurs missions.

QU'EST-CE QU'UN FAB LAB ?

Si les Fab Labs, abréviation de *Fabrication Laboratory*, bénéficient actuellement d'une visibilité médiatique, leur histoire n'est pourtant pas si récente. En effet, les laboratoires de fabrication sont nés à la fin des années 1990 aux États-Unis. On attribue la paternité de ces lieux au professeur Neil Gershenfeld du Massachusetts Institute of Technology (MIT). Il a mis en place un cours intitulé « *How to make (almost) everything* », durant lequel ses étudiants ont pu s'initier à des machines industrielles pour créer leurs propres objets, qui a rencontré un vif succès. Le premier Fab Lab est ainsi né au sein du MIT. Peu à peu, le concept s'est développé et a acquis une dimension internationale¹, une fondation² a été créée et une charte³ des Fab Labs a été mise au point par le MIT. Le respect impératif des points de la charte est nécessaire pour bénéficier du label « Fab Lab ».

Ces lieux de fabrication numérique se caractérisent par leur ouverture. Ils sont accessibles à tous : étudiants, particuliers, entrepreneurs, designers, bricoleurs... Il n'y a pas de restriction liée à l'âge ou au niveau d'expertise. Au contraire, les personnes novices dans la fabrication numérique peuvent bénéficier de l'expérience et des compétences des autres utilisateurs du Fab Lab. En effet, ces lieux reposent sur l'échange et le partage, en résonance avec le principe du *Do It With Others* (DIWO)⁴. Ces espaces d'apprentissage collaboratif fonctionnent sur un modèle de pair à pair physique. L'horizontalité des rapports entre les utilisateurs et la transmission des savoirs y dominent : j'apprends des autres et j'apprends aux autres.

Enfin, les activités menées dans les Fab Labs sont variées. On y vient pour acquérir des compétences dans le domaine de la conception assistée par ordi-

nateur (CAO), la programmation (Arduino⁵), utiliser des machines-outils pour réparer ou pour réaliser un prototype. Les appareils présents dans ces lieux vont au-delà de l'imprimante 3D. On peut également y trouver des découpeuses laser ou vinyles, fraiseuses à commande numérique...

En théorie, chaque activité doit être documentée. En effet, la notion de documentation⁶ est fondamentale dans ces nouveaux lieux. Elle sert de mémoire collective des projets qui y sont menés. C'est aussi un processus nécessaire dans la construction et la diffusion des connaissances. Dans ces espaces, il est indispensable de publier des contenus sous Creative Commons pour favoriser le partage et la réutilisation. Documenter, à travers un wiki par exemple, est la clé pour apprendre ensemble et partager.

DES BIBLIOTHÈQUES AUX FAB LABS

La réflexion sur le lien entre les espaces de fabrication numérique et les bibliothèques n'est pas nouvelle. Si le lien entre les deux types d'établissement n'est pas toujours admis en France⁷, la réflexion est bien avancée outre-Atlantique. Des bibliothécaires de Montréal⁸ travaillent depuis quelque temps sur cette question. Au regard des missions des bibliothèques et celles des Fab Labs, des passerelles peuvent être faites de façon relativement évidente. À l'image du bibliothécaire, le Fab Manager (celui qui anime l'espace de fabrication numérique) est un facilitateur d'accès aux savoirs et aux savoir-faire. Il met en relation un usager, ayant un besoin identifié, à des dispositifs informationnels. Ces dispositifs recouvrent diverses formes. Cela peut être la documentation générée par les activités antérieures produites par la communauté ou bien l'accompagnement à la maîtrise d'une technique (code) ou d'un outil (imprimante 3D). Les bibliothèques peuvent être complémentaires et fournir à travers leurs collections des documents qui serviront aux usagers à acquérir des connaissances théoriques pour apprendre à utiliser les outils des Fab Labs.

Les missions des bibliothèques et des lieux de fabrication numérique se font écho à travers leur capacité à fournir un accès ouvert aux technologies. En 2012, Lauren Britton, responsable de la bibliothèque de Fayetteville (FLL) aux États-Unis, expliquait : « *Les bibliothèques publiques sont là pour fournir un accès libre et ouvert à l'information, aux*

[1] Carte des fablab : www.fablabs.io/map

[2] Fab Foundation : www.fabfoundation.org/about-us

[3] <http://fab.cba.mit.edu/about/charter>

[4] « Les mots clés de l'univers des fab-labs », *La revue du projet*, 3 novembre 2014, <http://blogs.mediapart.fr/editon/la-revue-du-projet/article/031114/les-mots-clefs-de-l-univers-des-fab-labs>

[5] Carte électronique libre programmable qui permet de mettre en œuvre des robots, des imprimantes 3D, des capteurs, des automatismes et toute sorte d'objets connectés.

[6] Laurence Battais, « La documentation, le graal des Fab Labs et autres lieux de bidouilles », *Carrefour Numérique*, <http://carrefour-numerique.cite-sciences.fr/blog/la-documentation-le-graal-des-fab-labs-et-autres-lieux-de-bidouille>

[7] Sabine Blanc, « Que peuvent bien fabriquer bibliothèque et Fab Lab ensemble? », *Carrefour Numérique*, <http://carrefour-numerique.cite-sciences.fr/blog/que-peuvent-bien-fabriquer-bibliotheque-et-fab-lab-ensemble/>

[8] Gaëlle Bergougnoux, Marie D. Martel, « Un fab lab dans ma bibliothèque », *Espace B*, 14 juin 2013, <http://espaceb.bibliomontreal.com/2013/06/14/un-fab-lab-dans-ma-bibliotheque>

technologies et aux idées. Construire un makerspace (ce que nous appelons un Fabulous Laboratory) à la FFL offrira à notre communauté la possibilité d'avoir un accès à cette technologie qui change le monde»⁹. Les bibliothèques ont été pionnières dans l'accès aux technologies de l'information et de la communication (ordinateur, connexion à Internet...). De ce fait, elles ont une certaine légitimité à proposer un accès libre aux nouveaux outils¹⁰ de fabrication numérique. Plus encore, cela leur permet de renforcer leur utilité sociale et civique¹¹. Les bibliothèques qui intègrent dans leur projet d'établissement une dimension fabrication numérique ne sont pas uniquement des bibliothèques municipales. En effet, à l'heure des projets de Learning Centres, les bibliothèques de l'enseignement supérieur et de la recherche peuvent



➔ Fab Lab de la Cité des Sciences à Paris, en novembre 2013.

aussi inclure dans leur offre de services des espaces de fabrication numérique. Cela peut avoir particulièrement de sens dans certaines disciplines. Les étudiants en architecture peuvent ainsi réaliser des maquettes imprimées en 3D. Les étudiants en histoire de l'art ont la possibilité de reproduire des œuvres à partir de scanners 3D. De même, les étudiants de médecine peuvent recourir à l'impression 3D pour mieux comprendre le fonctionnement du corps humain. Ce ne sont que quelques exemples imaginables, la technique ouvre de nouvelles possibilités aux bibliothèques qui peuvent diversifier

La technique ouvre de nouvelles possibilités aux bibliothèques qui peuvent diversifier l'accès à la connaissance grâce à la fabrication numérique.

l'accès à la connaissance grâce à la fabrication numérique. Cette tendance se développe de plus en plus dans les bibliothèques universitaires américaines¹². En France, les établissements d'enseignement supérieur qui proposent ce type de service sont pour le moment minoritaires. Nous pouvons signaler le FaCLab de l'université de Cergy-Pontoise qui s'illustre par son accessibilité et les ateliers proposés. Pour les autres établissements, il s'agit souvent d'événements ponctuels : Fête de la science, exposition scientifique...

Dans le cadre des bibliothèques municipales, certaines collectivités se sont déjà emparées de cette

problématique comme celles des Ulis¹³, de Locminé, de Quimperlé... Quelques espaces publics numériques (EPN) ont su trouver un dynamisme en renouvelant leurs actions et en proposant des ateliers de fabrication numérique¹⁴.

Pour conclure, les missions, l'ouverture la plus large, la variété des publics, les technologies font des Fab Labs et des bibliothèques des lieux très proches. La mise en commun des compétences apparaît de plus en plus comme une évidence. Des partenariats sont à mettre en œuvre. Certains éta-

blissements qui ne possèdent pas les compétences pour gérer un espace de fabrication numérique en confient

l'animation à des associations dédiées aux Fab Labs. Dans un contexte de questionnement sur l'évolution des bibliothèques, la mise en place d'ateliers de fabrication numérique peut être l'occasion de repenser l'organisation et les projets des établissements qu'ils soient de lecture publique ou dans le champ universitaire.

THOMAS FOURMEUX
Bibliothèque d'Aulnay-sous-Bois
Membre de SavoirsCom1
Auteur du blog Biblio Numericus
thomas.fourmeux@gmail.com

[9] Sabine Blanc, « Made in ma bibliothèque », *OWNI*, 10 juillet 2012, <http://owni.fr/2012/07/10/made-in-my-bibliotheque>
[10] Jason Griffey, « The case for 3D printing », *American libraries magazine*, 16 septembre 2014, www.americanlibrariesmagazine.org/article/case-3d-printing
[11] Carolyn Sun, « Kansas boy gets new hand, created at a library makerspace », *The Digital Shift*, 11 février 2014, www.thedigitalshift.com/2014/02/k-12/library-innovation-leads-new-hand-kansas-boy
[12] *Libraries & Maker Culture : A Resource Guide*, <http://library-maker-culture.weebly.com/makerspaces-in-libraries.html>
[13] FabLab Mobile des Ulis, <http://fabriquesnumeriques.tumblr.com/fablab>
[14] EPN de Folelli, <http://folelli.blogspot.fr/2014/07/impression-3d.html>

● ● ● DE LA PIRATEBOX À LA BIBLIOBOX

Depuis presque deux ans, certaines bibliothèques françaises mettent en place des dispositifs de téléchargement d'œuvres numériques, appelés BiblioBox. C'est une déclinaison d'un autre dispositif développé par le professeur américain David Darts, la PirateBox. Sa démarche consistait à mettre en œuvre un système ouvert d'échange et de partage d'information anonyme qui respecte la confidentialité des données des utilisateurs. Cela s'inscrit dans un contexte de surveillance généralisée des moyens de télécommunications.

Comment ça fonctionne ?

La BiblioBox reprend certains des principes de la PirateBox. Elle fonctionne grâce à un routeur, qui permet de générer un réseau wifi, et d'une clé USB où sont stockées les ressources à télécharger. Une fois l'installation réalisée (remplacement du micrologiciel – *firmware* – d'origine par un autre, installation des fichiers de la BiblioBox), le routeur génère un réseau wifi auquel toute personne équipée d'un smartphone, d'une tablette ou bien d'un ordinateur portable peut se connecter. Ce réseau wifi s'apparente à un circuit fermé qui n'établit aucune communication avec le réseau internet et n'entre pas en conflit avec un réseau wifi pré-existant. Une fois connecté au réseau de la BiblioBox, l'utilisateur peut consulter ses mails ou effectuer une recherche *via* un moteur de recherche.

Un service ouvert

La BiblioBox est un service ouvert auquel tout le monde peut accéder sans identifiant ni mot de passe pour télécharger les fichiers (livres numériques, musique, vidéo, photographie, logiciels libres...) mis à disposition par les bibliothécaires. Par conséquent et par respect du *Code de la propriété intellectuelle*, les fichiers disponibles sont issus du domaine public ou publiés sous des licences libres comme les Creative Commons. Ainsi, la BiblioBox est l'occasion pour les professionnels de l'information de mettre en valeur tout un pan de la création issue de la sphère non marchande. Par ailleurs, contrairement à la PirateBox, les utilisateurs qui se connectent à une BiblioBox n'ont pas la possibilité de déposer des contenus numériques.

Un outil de médiation

Avec la BiblioBox, les bibliothécaires peuvent mettre en place des dispositifs de médiation numérique innovants. Grâce à leurs compétences en matière de sélection, de tri et de mise en valeur de l'information, ils peuvent proposer à leurs usagers des corpus d'œuvres organisés autour de thématiques spécifiques (Mois du film documentaire, Printemps des poètes, Centenaire de la Première Guerre mondiale, festivals locaux...). La BiblioBox est aussi un



Sylvain Naudin / Flickr (CC BY-SA 2.0)

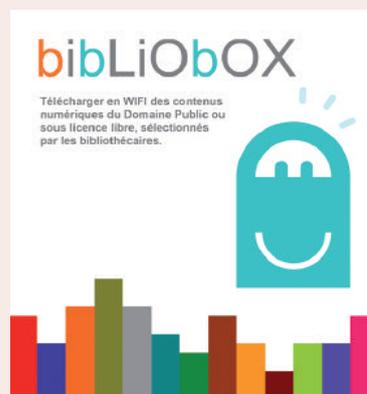
moyen de mettre en valeur les contenus produits par les bibliothécaires : des bibliographies, des critiques d'œuvres...

La BiblioBox a fédéré une communauté de professionnels intéressés par ce dispositif qui s'organise autour de la plateforme BiblioBox.net. Cette plateforme se présente comme un « *site communautaire pour partager et échanger sur l'expérience de BiblioBox en bibliothèque* ». Elle propose des tutoriels pour configurer et personnaliser sa BiblioBox, mais aussi un forum qui permet de poser des questions, de proposer des idées de médiations et de partager des retours d'expérience. Enfin, le site propose également une carte où sont recensées toutes les BiblioBox de France.

THOMAS FOURMEUX



POUR EN SAVOIR PLUS
<http://bibliobox.net>



À Montpellier, 4 médiathèques du réseau de lecture publique sont équipées d'une BiblioBox.

LORSQUE LE ROBOT SE PLANTE...



Simon Pow / Flickr (CC BY-NC-ND 2.0)

« Nos machines, au souffle de feu, aux membres d'acier, infatigables, à la fécondité merveilleuse, inépuisable, accomplissent docilement d'elles-mêmes leur travail sacré » (Paul Lafargue, *Le droit à la paresse*, 1880).

Infatigables assurément, infaillibles sans doute moins : la preuve par l'exemple.

UN ROBOT NOMMÉ ATTRIBUTOR

Depuis 2005, des groupes d'édition font appel à un « drone juridique » – pour reprendre l'expression utilisée par Lionel Maurel alias Calimaq – dénommé Attributor et lancé par une société californienne spécialisée dans la traque de contenus copiés. Il s'agit d'un système d'identification automatisé qui analyse les contenus proposés sur les grandes plateformes de téléchargement.

Hachette Livre fut le premier groupe d'édition français à l'expérimenter à partir de juin 2011, suivi par Editis. Il faut rappeler que le téléchargement de livres n'est pas pour l'instant sous le contrôle de l'Hadopi.

UN TYPOGRAPHE PASSIONNÉ

À l'autre bout de la chaîne, Alain Hurtig, typographe et maquetiste, exploite certaines œuvres du domaine public pour effectuer un travail remarquable de création graphique mis en ligne, sous licence Creative Commons, sur son site « L'outil typographique »^[1]. Il a, entre autres, utilisé un texte de Paul Lafargue (1842-1911), *Le droit à la paresse*, pour une de ses mises en forme typographiques. Ce texte, incontestablement dans le domaine public compte tenu de la date de décès de l'auteur, a été téléchargé sur le site de l'Association des bibliophiles universels. En avril 2013, Alain Hurtig reçoit un premier mail d'Attributor, agissant pour le compte d'Editis et l'enjoignant de retirer ce texte de son site. Malgré ses protestations restées sans réponse, il reçoit un second mail début 2014. Il se trouve que les éditions La Découverte, membres d'Editis, ont réédité en 2010 une version

du *Droit à la paresse*, publiée en 1969 par Maspero et enrichie à l'époque par une présentation de Maurice Dommanget, complétée lors de la réédition par une préface de Gilles Candar et quelques annexes. Bien évidemment, dans ce cas, seul l'appareil critique est sous droits, mais aucunement le texte de Lafargue.

ÉPILOGUE

Il s'agit là d'une erreur manifeste liée à une mauvaise interprétation du robot qui n'a pas pu (su) analyser précisément les métadonnées de l'édition augmentée. À la suite du billet publié par Calimaq sur S.I.Lex et de la mobilisation de quelques professionnels, les éditions La Découverte ont finalement reconnu l'erreur et Attributor a officiellement retiré sa demande le 6 février 2014. Il semblerait toutefois que cette erreur ne soit pas un cas isolé : en février 2013, une vague de demandes de retrait avait frappé le site Feedbooks qui comporte une section de livres numériques gratuits du domaine public ; en juin dernier, nouvelle demande de retrait émanant de Digimarc² pour le compte d'Editis de toute une série d'ePub du domaine public comme, par exemple, *Le Comte de Monte-Cristo* ou *Les voyages de Gulliver*. Et pourtant, Matt Robinson, un des responsables d'Attributor cité par Fabien Soyez sur Owni, affirmait : « notre système est complètement automatisé, l'information que nous trouvons grâce à notre algorithme est quasiment correcte à chaque fois. Nous croisons ce balayage automatisé avec des vérifications humaines, pour être sûrs qu'il s'agit bien d'une copie illégale d'eBook, et non d'une simple citation ou d'un pastiche. Nous ne risquons pas de faire de fausses identifications de piratage, ou pratiquement... voilà où se cache toute la nuance.

B. P.

SOURCES :

• Billet de Calimaq sur S.I.Lex « Message à Editis : laissez *Le droit à la paresse* dans le domaine public ! », 5 février 2014 :

<http://scinfolex.com/2014/02/05/message-a-editis-laissez-le-le-droit-a-la-paresse-dans-le-domaine-public>

• Billet de Fabien Soyez sur Owni « La milice privée d'Hachette Livre », 5 janvier 2012 : <http://owni.fr/2012/01/05/milice-privee-hachette-livre-telechargement>

[1] www.alain.les-hurtig.org

[2] Digimarc, société qui fournit les solutions de tatouage numérique des fichiers permettant le contrôle sur Internet, a racheté Attributor en 2012.

CultureWoK : de quoi avez-vous envie ?

Comment demander à Google ou à ses frères de nous trouver un « roman plutôt drôle et décalé avec un peu d'amour » ? La recherche plein texte ou sur les métadonnées n'est pas adaptée à ce genre d'interrogations et, pour y répondre, il a fallu faire un pas de côté et quitter le monde des mots pour entrer dans celui des sensations. Une démarche originale proposée par CultureWoK¹.

La recherche sensitive change le paradigme de la recherche documentaire et prend le contre-pied des concepts, règles et pratiques de la bibliothéconomie et, en particulier, en ce qui concerne l'administration et l'alimentation des données. Pour autant, il ne s'agit pas de les opposer mais d'ouvrir l'éventail des possibles. Aujourd'hui, la technologie permet de passer d'un mode de recherche à l'autre, ouvrant ainsi de nouvelles portes.

INTERNET ET « SENSITIVATION » DES DONNÉES

La recherche sensitive est le fruit d'une réflexion sur nos facultés à faire des choix sur les objets du monde, monde qui nous constitue et que l'on constitue :

- comment trouver de l'information pertinente en ligne sans connaissance préalable alors que tous les moteurs de recherche actuels sont fondés sur des requêtes cognitives ?
- comment trouver de l'information sonore et visuelle pertinente alors que la plupart des moteurs de recherche actuels sont fondés sur des requêtes exclusivement linguistiques ?
- comment trouver ce que l'on ne cherche pas ?

La « sensitivation » des données propose de répondre à ces trois questions fondamentales qui font actuellement l'objet de recherches approfondies, mais sont souvent posées séparément.

Dans ses dernières avancées, le web sémantique propose de favoriser un développement souple et horizontal, fondé sur l'intelligence collective et le partage des données entre tous. Pour autant, les problèmes que suscitent l'indexation d'un nombre exponentiel de données et l'accès à une information pertinente ne sont pas résolus. Cette amplification des échanges horizontaux, de la coopération ouverte, de la construction par collaboration et des réseaux sociaux demande une meilleure intégration des données linguistiques, visuelles et sonores et la prise en compte des dimensions affectives, sensibles, sensorielles, intersubjectives des choix individuels et collectifs.

Les recherches qui sont menées à l'heure actuelle ne prennent souvent en compte que l'une ou l'autre de ces pistes. Le WoK est le premier outil à travailler

au développement parallèle de ces trois pistes. Plus on aura « sensitivé » de données, plus l'outil apportera un accès horizontal et sensuel aux savoirs et simplifiera/favorisera l'accès à des objets complexes en ne leurs déniait aucune de leurs dimensions. C'est l'ensemble des acteurs qui permet l'accès sensitif aux connaissances culturelles ou autres, la recherche sensitive n'est donc pas qu'algorithmique, mais, aussi et surtout, le fait de tous, remettant l'humain au centre de la machine.

Pour autant, il ne s'agit pas de faire un choix entre recherche sensitive et documentaire. Toutes deux répondent à des besoins différents : la première est un outil de médiation culturelle, basée sur la subjectivité, alors que la seconde se doit de décrire l'intégralité d'un fonds et de son contenu le plus objectivement possible. Cette approche différente change la donne pour le travail d'alimentation et d'administration des données.

LA SOBRIÉTÉ HEUREUSE

À l'aune de leur expérience en matière de catalogue, beaucoup s'inquiètent de l'ampleur de la tâche pour mettre en place un WoK. Dans ce cas, l'utilisateur ne recherche pas un document précis mais un objet quelconque, qu'il ne connaît pas encore mais qui répond à ces envies et humeurs du moment. Dans ce contexte, il n'est pas nécessaire, et serait même contreproductif, de vouloir mettre dans un WoK l'intégralité des documents. La condition nécessaire pour « sensitiver » un objet est d'abord qu'il ait ému personnellement le « sensitif » et donné l'envie de partager ce ressenti. Dans une logique de proposition, il suffit de quelques réponses pour satisfaire une demande. On est donc plus dans une logique qualitative que quantitative, l'expérience montre qu'un WoK est efficace avec seulement quelques centaines d'objets.

En recherche documentaire, il faut que l'utilisateur repère rapidement l'objet qu'il recherche dans une liste de résultats. Pour présenter d'abord ceux que l'on pense correspondre le mieux à une demande, on va construire des calculs de pertinence, des mécanismes de tri et des critères de classement sans parler des considérations commerciales mises en place par les grands moteurs de recherche. Les

[1] www.CultureWoK.com

données nécessaires à ces opérations vont devoir être disponibles, organisées et maintenues. Si l'on assume le parti pris de la subjectivité, toutes les réponses se valent. Comment décider que le ressenti d'un individu est plus pertinent qu'un autre ? En se plaçant dans une logique de proposition, il n'est pas nécessaire d'obtenir de nombreuses pages de résultats. Sur CultureWoK, si la requête retourne beaucoup de réponses, on en présente seulement 24 et la requête pourra être rejouée pour en obtenir d'autres. Pour présenter une réponse, on ne peut se contenter de l'image de la couverture du document, il est donc nécessaire de disposer de données bibliographiques minimum : sur CultureWoK, seuls l'image de couverture, le titre, l'auteur et le résumé ou synopsis sont obligatoires. Il faudra donc rentrer ces données, mais il n'est pas nécessaire de disposer de mécanisme de gestion d'autorités. Chaque notice est vue individuellement et à seul titre descriptif.

La recherche documentaire impose que les référentiels soient correctement organisés et respectés, avec un nombre d'entrées pour chaque vocabulaire ou type d'autorité qui se chiffre en centaines de milliers ou millions. La recherche sensitive, elle, se contente d'une dizaine de critères sensitifs et autant de critères catégoriels par type d'objet avec cinq valeurs possibles pour chacun. Cela suffit amplement pour indexer un nombre potentiellement infini d'objets.

UN ESPACE NATURELLEMENT COLLECTIF

Chaque établissement dispose d'un mini site personnalisé représentatif de sa sensibilité et sous-ensemble du site global organisé sous forme arborescente. Les recherches portent sur les objets « sensitivés » du WoK et de ses sous-WoK : une recherche sur CultureWoK se fait sur l'intégralité du site, mais une recherche sur le WoK de la médiathèque de Mérignac, par exemple, sera restreint aux objets de l'espace, fonctionnant ainsi comme un filtre, basé uniquement sur les contributions des professionnels et de lecteurs parrainés par la médiathèque. La géolocalisation des établissements permet de proposer les lieux, médiathèques ou librairies, les plus proches du poste de l'internaute pour les objets trouvés.

Avec les services web et le système de la double publication, les objets sont créés et « sensitivés » sur le système de gestion de l'établissement avant d'être « poussés » sur le site CultureWoK. La recherche peut se faire indifféremment sur l'un ou l'autre site et chacun présente les objets tels qu'ils sont dans leur environnement. Ce fonctionnement de type catalogue collectif se fait naturellement sans travail supplémentaire pour les établissements : catalogage et localisation centralisés puis distribution des informations vers le local sur le modèle du Sudoc

● ● ● UN WOK EN BU ?

Simplicité d'usage ne veut pas dire pauvreté des contenus. Dès sa genèse, notre moteur d'indexation et de recherche a été pensé pour traiter des objets complexes et hétérogènes. Si une de ses expressions, qui est le CultureWoK, est très généraliste dans l'accès aux connaissances, notre outil s'adapte également à des usages spécifiques.

Ainsi, s'agissant des bibliothèques universitaires, le WoK permettrait par exemple :

- pour les chercheurs de créer un outil d'indexation et de recherche, avec leur propres critères et données, de manière simple, leur assurant la qualité des objets proposés à la recherche (tous types : écrits, visuels, sonores...) ;

- le maillage et l'enrichissement entre établissements partageant un domaine de connaissance qui pourraient utiliser des mêmes espaces de critères ;

- la possibilité de créer des « méta-WoK » permettant d'établir des liens transversaux entre les disciplines.

L'outil offrirait ainsi aux étudiants et chercheurs un accès simplifié à des données fiables et un panel plus ample de réponses et d'ouvrages pour une seule requête.

À l'heure où, captée par quelques algorithmes, l'indexation des données échappe à tous, le WoK replace le détenteur/diffuseur des connaissances au centre de la machine.

ou agrégation de catalogues locaux comme cela se pratique dans les établissements de lecture publique. La « sensitivation » se fait au niveau de l'établissement et profite de fait aux WoK plus généraux.

Ce n'est pas parce qu'un WoK est facile à mettre en place et que son administration est réduite qu'il faut en déduire qu'il peut remplacer le catalogue. Les recherches sensibles et documentaires offrent des services complémentaires avec une approche différente. Il serait cependant dommage de se contenter de les juxtaposer, alors qu'en les hybridant elles améliorent leurs mécanismes de recherche respectifs.

L'ÉQUIPE DU WOK

Gladys Cathelain - contact@deuzio.fr
 Renaud Garcia - CultureWoK@gmail.com
 Jacques Kergomard - JK@BiblioTIC.fr



POUR EN SAVOIR PLUS

Rendez-vous sur le site www.CultureWoK.com et son blog pour visiter les WoK de nos partenaires : des bibliothèques municipales (Saint-Médard-en-Jalles, Orthez, Mérignac, Bègles, Pessac, Saint-Jean-de-Luz, Seyssinet Pariset, Saint-Herblain...), la BDP de la Gironde, des librairies (La Machine à lire), BD Fugues Bordeaux, Le Castor Astral éditeur, l'Agence régionale du livre Provence-Alpes-Côte d'Azur... et, comme « Perdu ? Retrouvé ! », quelques WoK privés de passionnés.

Opinion mining : des vertus de l'hybridation humain/machine

L'*opinion mining* consiste en une analyse automatisée des verbatims en ligne (messages, blogs, forums...) afin d'en extraire des informations subjectives (jugement, évaluation, émotion). Une approche qui nécessite toutefois une forte expertise humaine...

On ne peut en vouloir aux fournisseurs de solutions d'*opinion mining* d'avoir mis en avant les vertus de l'automatisation de leurs services puisque l'informatique et la linguistique avaient elles-mêmes adopté un vocable en français « TAL » (traitement automatique du langage, ce qui fait une différence avec « NLP » *Natural Language Processing*) qui laissait entendre qu'il existait des fondements scientifiques et techniques à cette automatisation. Cependant, la demande a augmenté de façon impressionnante depuis que les marques ont considéré qu'il devenait important – pour leurs investisseurs avant tout – de suivre leur réputation sur le web et sur les réseaux sociaux, tant les volumes de verbatims produits dans ces plates-formes étaient massifs et contagieux.

L'AUTOMATISATION ET SES LIMITES

On ne peut comprendre l'explosion de l'offre de l'*opinion mining* (280 services en 2012) si l'on ne renvoie pas à cette anxiété du marketeur en charge du suivi de la réputation d'une marque, de la répercussion des événements qu'il génère, des conversations des communautés qu'il soutient, des risques

mais il se trouve que ces méthodes gagnent même les approches scientifiques malgré leurs limites remarquables.

Cependant, il faut faire la distinction entre ce que les promoteurs de ces méthodes annoncent lorsqu'ils vantent leurs systèmes automatisés tout au long de la chaîne de traitement et leurs pratiques réelles qui s'appuient sur une hybridation beaucoup plus marquée entre approches automatisées, *machine learning* et expertise humaine. Il serait donc plus juste de parler d'« *opinion mining* assisté par ordinateur », comme cela se faisait pour la traduction, mais tout se passe comme si ces formules révélaient une faiblesse des ordinateurs en question et ne suffisaient à vendre l'anxiolytique qu'est l'*opinion mining*.

Selon les objectifs poursuivis, cet équilibre ou cette composition peuvent être différents :

- **centrer le suivi sur des agrégats en masse** autour de quelques mots-clés et suivre leur évolution sur un grand nombre de sources reste nécessairement très grossier et peut donner lieu à une automatisation plus avancée ;

- **suivre à la trace un énoncé qui fait événement** peut aussi être automatisé largement si l'on accepte la réduction de l'étude des dérivations faites à partir des énoncés d'origine ;

- **moniter en permanence un milieu ou des communautés** qui échangent sur des thèmes qui leur sont en partie propres requiert nécessairement des expertises humaines plus importantes parce que les bases de dictionnaires standards n'existent pas toujours et que les énoncés peuvent être fortement contextualisés.

Un simple contrôle de qualité de ces prestations devrait entraîner à examiner la qualité des personnels, des process et des produits. Or, sur toute cette chaîne, de nombreux maillons sont conçus comme des boîtes noires, car ce sont les résultats qui comptent. Ceux-ci sont souvent présentés à un tel niveau d'agrégation à l'aide de tableaux de bords si séduisants que la discussion est rarement possible sur les conditions de leur production. La combinaison avec un choix d'exemples judicieux rend la démonstration toute puissante : « *les clients focalisent leurs critiques sur la vitesse du processeur et n'aiment*

L'automatisation est intéressante si elle s'appuie sur du *machine learning* dans lequel un expert validera les retours de verbatims sélectionnés.

de crise qui ferait dévisser le cours de l'action en quelques minutes. C'est pourquoi toutes les solutions ont trouvé un moment leur place sur le marché en dépit des faiblesses criantes de leurs approches¹. C'est pour contrer cette angoisse que les outils proposés doivent avant tout permettre la réaction et non la réflexivité, qui supposerait modèles sophistiqués ou discussion entre causalités. Le *big data* est convoqué avec ses puissances de calcul et ses approches de *matching* (correspondance) générant des corrélations entre séries de données pour traiter tous ces verbatims, qui remplissent les conditions de volume, de variété mais surtout de vélocité. Tout autre usage de ces techniques peut se faire dans des conditions de scientificité différentes,

[1] Documentées en détail dans : Dominique Boullier, Audrey Lohard, *Opinion mining et sentiment analysis*, Open Editions Press, 2012, <http://books.openedition.org/oepp/198>

(Pleins feux sur...)

La bibliothèque de l'École française de Rome

Travailler à la bibliothèque de l'École française de Rome (EFR), c'est à la fois un bonheur et un défi : d'abord séduit et impressionné, on est vite confronté à l'originalité de cet univers dont il faut patiemment décoder le fonctionnement avant de pouvoir construire une véritable politique documentaire.

Un mot sur l'École, tout d'abord. Fondée en 1873, l'EFR présente une originalité largement due à son implantation romaine : dès sa création, elle a refusé de se laisser enfermer dans le cadre strict de l'archéologie pour s'intéresser à l'histoire dans son ensemble ; on pourrait dire que rien de ce qui touche à l'Italie, de la Préhistoire à nos jours, ne lui est étranger.

UNE BIBLIOTHÈQUE PATRIMONIALE ET SCIENTIFIQUE

Sa bibliothèque en est le reflet : une étude systématique de la constitution des fonds montrerait sans nul doute comment elle en a accompagné les ouvertures chronologiques et disciplinaires (histoire contemporaine, sciences sociales...) au fur et à mesure des décrets qui en précisaient les missions.

Mais elle en est aussi d'une certaine façon la matrice : en témoigne de façon emblématique le « fonds Volterra » (le grand juriste italien Edoardo Volterra, contraint de quitter son poste de professeur suite aux lois raciales de 1938, avait constitué sa vie durant une bibliothèque très importante d'histoire du droit, que ses héritiers léguèrent à l'EFR en 1984, en remerciement de l'accueil et du soutien que lui avait accordés la France). Cette bibliothèque de 7 200 volumes, 16 400 tirés à part et 1 300 livres anciens – dont deux incunables – est comme le microcosme de la bibliothèque de l'École, dont elle incarne trois aspects essentiels.

• **Dimension patrimoniale** : les enjeux de conservation y sont particulièrement importants, mais pas seulement dans une optique de préservation matérielle ; ils obligent à questionner la notion même de patrimoine dans une bibliothèque de recherche. Le cabinet de travail de Volterra, reconstitué à l'identique, c'est l'itinéraire d'un chercheur, avec sa relation quasi-maniaque avec les livres, ses échanges de publications avec ses collègues, son environnement de travail... on n'est pas si loin des probléma-



➔ Le *studiolo*, salle de travail des boursiers.

tiques actuelles de traitement des archives et des données de la recherche.

• **Dimension scientifique** : depuis son ouverture au public en 1989, la bibliothèque Volterra, dont les collections juridiques n'étaient pas *a priori* au centre de la politique scientifique de l'École, suscite de plus en plus de recherches. De même, la constitution patiente des collections de l'École n'est pas seulement le reflet des besoins des chercheurs : d'une certaine manière, elle les oriente aussi ; on peut dire que l'EFR exprime au plus haut point le continuum de la recherche, depuis les sources fournies par la bibliothèque jusqu'à la publication de ses résultats par le service des publications.

• **Ouverture vers l'extérieur** : une des conditions du legs Volterra était l'ouverture au public. Cette préoccupation rencontre le souci croissant de l'École de favoriser la découverte de ses locaux, en accueillant des visites régulières du grand public, ou des visites occasionnelles d'élèves, d'étudiants, d'universitaires, de bibliothécaires... français ou étrangers.

UN LIEU, ENTRE ATOUTS ET CONTRAINTES

Car cette bibliothèque, c'est aussi – et d'abord ? – un lieu : sans doute n'aurait-elle pas le même attrait si elle n'était pas aussi idéalement située, non seulement en plein centre historique de Rome, pas très loin de la bibliothèque et des archives du Vatican, de la bibliothèque du Sénat, de la Casanatense... mais, de surcroît, dans le Palais Farnèse qu'elle partage avec l'Ambassade de France et dont elle a patiemment « grignoté » les espaces.

• C'est le **lieu des livres** (210 000 volumes, 2 000 titres de périodiques dont 1 000 vivants) en accès libre (80 % des collections), 72 h par semaine – du lundi au samedi inclus – plus de 290 jours par an.

• C'est aussi « **la maison des chercheurs** », en particulier de ces lecteurs privilégiés que sont les membres (18 chercheurs ayant la possibilité de travailler pendant 3 ans à Rome) et les boursiers (plus de 150 par an, qui disent souvent progresser dans leur thèse davantage en un mois qu'en un an dans leur université d'origine) : ils y ont leur place de travail attitrée, peuvent y laisser les ouvrages dont ils ont besoin, obtenir des lettres d'introduction pour telle bibliothèque ou tel centre d'archives... Mais le service et ses quelque 170 places assises attirent aussi plus de 1 300 autres lecteurs actifs.

• C'est encore et surtout un **lieu d'échanges** : entre chercheurs bien sûr, des néophytes aux plus confirmés, dans toutes les langues (si les Italiens et les Français sont évidemment les plus nombreux, 26 nationalités sont représentées...), mais aussi entre chercheurs et bibliothécaires : rarement j'ai rencontré une telle proximité, un dialogue aussi facile sur les questions documentaires (opportunité d'acheter tel ou tel document, difficulté d'indexation, expertise sur une ressource électronique...).

• C'est enfin un **lieu paradoxal** : très fréquenté, mais difficile d'accès (une fois franchi

© EFR / Phot. F. Bruni, 2014



↳ La grande salle de lecture de la bibliothèque.

le sas de l'Ambassade, un ascenseur dessert les deux étages du Palais occupés par la bibliothèque, mais elle se développe en réalité sur cinq niveaux sans compter les galeries : autant d'escaliers à gravir, autant de rangement à faire manuellement, et autant de limites à l'accès des personnes handicapées). Les salles de lecture (15) sont aussi nombreuses que le personnel de la bibliothèque : toute velléité de surveillance systématique est donc vouée à l'échec. La réserve Volterra et ses incunables se trouvent sous les toits, tandis que le magasin est aménagé dans les sous-sols du Palais, connus dès l'Antiquité pour être inondables...

Quant au poids des livres, il tourne à l'obsession : ni Sangallo ni Michel Ange, en construisant le Palais, ne pouvaient imaginer la contrainte physique que poserait un jour cette accumulation d'ouvrages.

DES RÉSEAUX EN CONVERGENCE

Dans ce type de bibliothèque, tout est donc à penser « sur mesure » : la sécurité (des collections et des lecteurs), la conservation des documents, l'aménagement des espaces, la gestion du personnel (13 agents sur 15 en contrat local de droit italien), mais aussi la constitution des collections, leur signalement, leur accessibilité, leur mise en valeur. C'est alors que se pose la question des réseaux : vers où faire pencher le curseur, quelles sont les échelles pertinentes pour un établissement français (établissement public

à caractère scientifique, culturel et professionnel, mais sans étudiants, sans diplôme, sans autonomie) à l'étranger, dans un environnement plus riche en archives et bibliothèques qu'aucun autre ?

- En tant que bibliothèque française, nous avons un rôle à jouer dans la politique documentaire nationale : d'où notre participation au Sudoc, à Couperin, et de façon plus spécifique au réseau des bibliothèques des Écoles françaises à l'étranger. Nous y apportons des collections originales (4 384 unicas pour l'EFR, 9 103 si l'on y ajoute la Villa Médicis et le Centre Jean Bérard de Naples avec lesquels nous partageons le catalogue collectif « Farnèse ») et une expertise en matière d'indexation qui peuvent être un atout dans divers projets Collex ; concernant la valorisation des productions des chercheurs, nous avons adopté le principe d'une archive ouverte institutionnelle dans HAL, dans laquelle la bibliothèque est maître d'œuvre. Mais tout est à définir : que doit-on considérer comme un « chercheur de l'École », puisque celle-ci n'a pas de laboratoire ? De même, pour les ressources électroniques, comment calculer un forfait, puisque nous n'avons pas d'usagers au sens universitaire du terme ?

- La dimension nationale touche vite ses limites : nos « unicas » identifiés dans le Sudoc ne le sont pas nécessairement en Italie ; il s'agit donc, avant d'entreprendre par exemple une politique de numérisation ou des actions de conservation partagée, d'analyser de façon

précise nos points de force, mais les outils manquent pour les repérer facilement. Là encore, l'identification des partenaires est cruciale et se fonde à la fois sur les affinités géographiques et disciplinaires. C'est ainsi que nous travaillons de plus en plus avec les bibliothèques d'instituts étrangers à Rome, plus qu'avec les bibliothèques italiennes proprement dites, pour dresser notre carte documentaire, faciliter la circulation de nos publics, confronter nos états de collections de périodiques et en répartir l'acquisition et la conservation entre supports papier et numérique ; ou pour essayer de contrer les pratiques monopolistiques de certain(s) éditeur(s) italien(s) ; ou encore pour mettre en place un « *Discovery Tool* » partagé.

« Ni tout à fait la même, ni tout à fait une autre... ». Les quelques réflexions ci-dessus ont simplement voulu montrer que la bibliothèque de l'EFR, loin des clichés qu'elle pourrait susciter, est pleinement impliquée dans les problématiques documentaires actuelles, même si elle doit les explorer et les repenser dans l'histoire et le territoire qui sont les siens.

ANNIE COISY

Directrice de la bibliothèque de l'EFR
 bibliotheque.direction@efrome.it



POUR EN SAVOIR PLUS

VISITEZ LE SITE DE L'ÉCOLE FRANÇAISE DE ROME : www.efrome.it

JPAC 2014 : Nouveaux regards sur le patrimoine

Depuis 1997, les Journées des Pôles Associés et de la Coopération (JPAC) sont aux partenaires de la Bibliothèque nationale de France (BNF) ce que les Journées Abes sont aux réseaux documentaires de l'Enseignement supérieur : un rendez-vous annuel propice aux échanges, non seulement sur les produits et services de l'institution organisatrice, mais plus largement sur les évolutions des métiers documentaires et culturels.

Cette 15^e édition des JPAC, co-organisée par les quatre institutions constitutives du Pôle associé régional Languedoc-Roussillon¹ à l'Opéra Comédie de Montpellier, avait pour fil directeur les « nouveaux visages du patrimoine », à travers ses usages ludiques, créatifs, intimes, régionaux et savants.

Des coopérations étendues...

Le renforcement de la coopération avec les territoires et de Gallica comme outil de partage fait partie des 16 grands axes inscrits dans le contrat de performance 2014-2016 de la BNF². Le nombre de partenaires moissonnés et référencés par Gallica n'a cessé de s'accroître ces dernières années. La logique d'un écosystème numérique



➤ Numistral, fruit d'un partenariat entre la BNU et la BNF, est la première réalisation inscrite dans « Gallica Marque Blanche ». Source : gallica.bnf.fr / BNU Strasbourg

réparti (encore marquée par les manques inhérents à l'OAI-PMH qui ne permet de disséminer ni plein texte ni images) cohabite désormais avec des fonctionnalités de bibliothèque numérique collective, dont le développement connaît des rythmes différenciés. Un tiers des 8,5 millions de pages produites dans le cadre du marché de numérisation 2011-2014 provient de partenaires auxquels la possibilité d'abonder les lots a été offerte³, et le dépôt de fichiers déjà numérisés et répondant à certains critères est également opérationnel⁴. L'expérience « Gallica Marque Blanche », menée depuis 2013 avec la Bibliothèque nationale et universitaire de Strasbourg pour la mise en œuvre de sa bibliothèque numérique Numistral⁵ à l'aide des briques du produit Gallica, n'est en revanche pas arrivée au stade d'une prestation récurrente. Bien qu'une seconde expérimentation soit envisagée dans le cadre de la « Grande Collecte » d'archives privées de la guerre

1914-1918⁶, les dispositifs existants restent à améliorer, le modèle économique à inventer. L'extranet développé dernièrement à l'usage des pôles associés et éditeurs partenaires, ou « Espace Coopération »⁷, tend à devenir un outil central du dispositif numérique mutualisé que coordonne la BNF.

... pour des données ouvertes et partagées

L'ouverture et le partage des données en tant que nouveaux vecteurs de coopération ont trouvé pour illustrations le Catalogue collectif de France (CCFr) – auquel le quart de ses usagers accède désormais directement *via* les moteurs de recherche ayant moissonné ses 5 millions de pages – et l'entrepôt sémantique data.bnf.fr qui devrait couvrir l'intégralité des données BNF d'ici fin 2015. Sa réutilisation par un éventail d'acteurs dessine les prémisses d'une nouvelle forme de service public s'appliquant à des données libres, ouvertes et



➤ L'interné caricaturiste, dessin d'un prisonnier. Source : www.lagrandcollecte.fr / Archives départementales de la Charente-Maritime.

liées. Les nombreuses initiatives de médiation numérique : conception de jeux vidéos inspirés d'archives de la Grande Guerre, commandes littéraires entourant les « Chroniques retrouvées du midi »⁸, traduisent un changement de regard vis-à-vis du numérique patrimonial, aussi bien qu'un renouvellement des services exploitant les données culturelles.

Ces journées professionnelles ont aussi été l'occasion d'avant-premières : quelques semaines après le début de l'opération Prêt numérique en bibliothèques⁹, dont Montpellier Agglomération est l'un des trois sites pilotes, la première démonstration de l'application BCWeb (« BnF Collecte du web »¹⁰) a eu lieu le 2 octobre à la médiathèque centrale Émile Zola, l'un des premiers sites choisis pour

diffuser les collectes du dépôt légal de l'Internet.

En conclusion, permettons-nous de filer une métaphore culinaire pour résumer les propos de Gildas Illien (BNF) sur le processus d'ouverture des données. Les bibliothécaires ont d'abord considéré que le web, ressenti comme adversaire des catalogues, allait compromettre leurs recettes bibliographiques. Progressivement, la mayonnaise entre web et catalogues a pris. Désormais des ingrédients exogènes sont incorporables et les données elles-mêmes gagnent à assaisonner directement le web.

JEAN-MARIE FEURTET
ISABELLE MARTIN
Abes

À NOTER

Les bibliothèques de dépôt légal imprimeur et autres bibliothèques territoriales ont été bénéficiaires en 2013 de plus des deux tiers des subventions allouées par la BNF à ses partenaires (16 % au bénéfice d'établissements d'enseignement supérieur et recherche et 17 % pour les autres types de partenaires). Le taux de cofinancement appliqué par la BNF est de 50 % au maximum, à l'exception de plusieurs programmes de rétroconversions en bibliothèques municipales classées prioritaires et cofinancés jusqu'à 70 %. À partir de 2015, les dispositifs contractuels seront simplifiés et les durées de conventionnements étendues de trois à cinq ans.

[1] Le premier pôle associé régional a été signé en 2008 entre la BNF, la Région Languedoc-Roussillon, Montpellier Agglomération, Languedoc-Roussillon Livre & Lecture et la Drac. En octobre 2014, l'action régionale de la BNF était relayée par 15 pôles associés régionaux.

[2] www.bnf.fr/documents/contrat_performances_2014.pdf

[3] Les partenaires concernés par cette ouverture des marchés de numérisation sont déjà au nombre

d'une centaine (INHA, BHVP, Assemblée nationale... ont été d'importants contributeurs). L'archivage pérenne des fichiers masters dans SPAR est compris dans l'opération.

[4] Seuls des fichiers TIFF peuvent être déposés, mais une extension au JPEG ainsi qu'aux résultats d'OCR (fichiers Alto) devrait bientôt intervenir. La description préalable des ressources au sein du Catalogue général de la BNF est indispensable.

[5] www.numistral.fr

[6] www.lagrandecollecte.fr

[7] <http://espacecooperation.bnf.fr>

[8] Projet mené par Languedoc-Roussillon Livre & Lecture, pour lequel une trentaine d'écrivains ont été invités à revisiter, à travers feuillets littéraires et réclames, la presse régionale des XIX^e-XX^e s. dont 600 000 pages ont déjà été numérisées (www.chroniquesdumidi.fr).

La publication de cette exposition

virtuelle s'échelonne jusqu'à l'été 2015.

[9] Projet inscrit dans le prolongement du hub Dilicom en charge des transactions relatives aux livres numériques.

[10] <https://collecteweb.bnf.fr>

licences
nationales
• FR

ISTEX : QUATRE NOUVELLES LICENCES NATIONALES SIGNÉES

Toujours dans le cadre du projet d'investissement d'avenir Istex, initié par le ministère de l'Enseignement supérieur et de la Recherche, quatre nouveaux corpus ont été acquis depuis le printemps dernier.

Wiley

Une licence signée avec l'éditeur américain Wiley permet d'accéder à près de 2 200 titres de revues scientifiques, de 1791 à 2011. Le contrat exige également la livraison du texte intégral et des métadonnées des cinq années suivantes : les volumes de 2012, 2013, 2014, 2015 et 2016 seront ainsi livrés respectivement en 2015, 2016, 2017, 2018 et 2019. Ces périodiques regroupent des champs de recherche divers parmi lesquels l'écologie, les mathématiques, la chimie, la physique et les sciences de la santé.

BMJ Publishing Group Limited

Dans le domaine médical, un contrat a également été passé avec la société BMJ Publishing Group Limited (*BMJ = British Medical Journal) afin de proposer aux bénéficiaires du programme 64 titres depuis le début des publications jusqu'à 2013 inclus.

Walter De Gruyter

L'acquisition d'un peu plus de 500 titres couvrant la période de 1826 à 2012 a par ailleurs été négociée auprès de l'éditeur allemand Walter De Gruyter. Ces archives de revues relèvent de domaines de recherche variés, incluant notamment des ressources en sciences humaines et sociales.

Sage

Un autre contrat, conclu avec l'éditeur américain Sage, permet l'accès à plus de 600 revues couvrant l'ensemble des disciplines de recherche, pour une période allant de 1847 jusqu'à l'année 2009.

À l'instar des douze autres corpus acquis depuis le début du programme d'achat en licence nationale, ces contenus restent accessibles à partir des plateformes des éditeurs, après déclaration des adresses IP des établissements *via* le site des licences nationales :

<http://accs.licencesnationales.fr/>

En quête d'Arabesques

Merci chers lecteurs ! L'enquête de satisfaction lancée cet été pour recueillir vos avis au sujet de la revue *Arabesques* – qui, à tout juste 20 ans¹, n'avait jamais encore connu d'évaluation de ce type – a conforté le comité de rédaction de la revue dans son action et a constitué un point de départ favorable pour étayer l'avis du conseil scientifique de l'Abes quant à son avenir.

Une appréciation générale positive

Avec 390 répondants, dont 96 % de professionnels des bibliothèques et de la documentation issus des universités membres des réseaux de l'Abes (merci les collègues !), le taux de satisfaction global (95 %) a dépassé nos attentes, avec une mention particulière pour les 92 % qui estiment la revue « utile dans leurs activités professionnelles ». On a apprécié le

satisfecit particulier (95 % de satisfaits à très satisfaits) pour la rubrique « dossier thématique », mais aussi relevé dans vos commentaires le léger désaveu concernant l'agenda. Reflet des interrogations de la profession, le classement concernant les thématiques abordées depuis la refonte de la revue place en trio gagnant les dossiers consacrés au système de gestion de bibliothèque mutualisé, n° 70 (57 %), aux ressources numériques et cadre juridique, n° 72 (52 %) et à la valorisation des données de la recherche, n° 73 (51 %).

Des suggestions bienvenues

Au chapitre des suggestions, les aspects techniques – normes et formats principalement – et juridiques restent vos préoccupations majeures dans le contexte en pleine évolution des métiers de l'information et de la documentation. Selon vous, ces questions doivent être encore plus largement développées dans la revue. C'est toutefois dans une perspective nouvelle que certains souhaiteraient les voir abordées, dans un « esprit plus polémique » ou « tourné vers l'international ». On a par ailleurs noté



quelques suggestions de rubriques intéressantes : à l'instar de la très estimée rubrique « Pleins feux » (taux de satisfaction de plus de 90 %) qui met en lumière un établissement particulier, la création d'une rubrique « Portrait » permettrait de présenter une personnalité du monde de la documentation, sous la forme d'une interview par exemple. Preuve du besoin d'interactions dans un univers professionnel habitué au collaboratif, une rubrique « courrier des lecteurs » a également été appelée de vos vœux.

Papier et/ou numérique

Mais au-delà des contenus, le comité de rédaction souhaitait connaître l'opinion des lecteurs sur la forme mais aussi – et surtout – sur les formats de diffusion actuels de la revue. Tout d'abord, une bonne nouvelle : la mise en page et la ligne graphique de la revue dont la maquette, réalisée par Anne Ladevie, a été complètement révisée en 2013 (n° 69) sont très prisées (97 % de satisfaits à très satisfaits). S'agissant des formats, les avis sont plus partagés. Si la version imprimée reste de loin la plus appréciée (64 % des votes en sa

faveur), la version en ligne est jugée complémentaire... et largement perfectible. En effet, surtout pour une revue qui consacre régulièrement ses colonnes à la question du numérique, il est primordial de mettre en place une « véritable revue en ligne ». Recherche plein-texte, indexation des articles, citabilité, fonctionnalités de partage amélioreraient grandement sa visibilité et son accessibilité. Une orientation

qu'a d'ailleurs validée le conseil scientifique réuni le 13 octobre dernier. Pour l'heure tout au moins, le maintien des deux versions reste indispensable (82 %).

C'est donc un bilan très encourageant pour *Arabesques*, à qui l'on souhaite encore de beaux jours, malgré le départ prochain de Béatrice Pedot qui assure depuis plus de trois ans avec son enthousiasme bien connu la coordination éditoriale et le secrétariat de rédaction de la revue.

CHRISTINE FLEURY

MARION GRAND-DÉMERY

Pour le comité de rédaction d'Arabesques

[1] Le numéro 0 de la revue a été publié en janvier 1995.



POUR EN SAVOIR PLUS

RETROUVEZ LES RÉSULTATS COMPLETS DE L'ENQUÊTE EN LIGNE SUR LE SITE DE L'ABES : www.abes.fr/Media/Fichiers/enquete-arabesques

SGBM : LA PHASE PILOTE EST LANCÉE

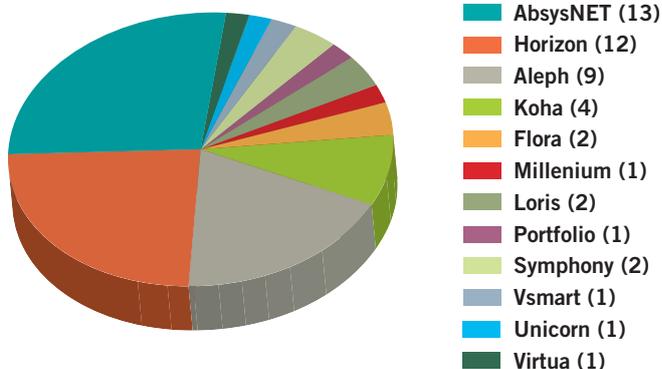
Le groupement de commande pour le système de gestion de bibliothèque mutualisé (SGBM) est constitué. Il compte plus de la moitié des bibliothèques déployées dans le réseau Sudoc.

Le démarrage de la phase pilote du projet SGBM, qui a succédé à la phase d'étude, s'est concrétisé, côté établissements, par la constitution d'un groupement de commande. Il regroupe l'ensemble des établissements qui entendent procéder à la migration de leur système intégré de gestion de bibliothèque (SIGB) vers un SGB de nouvelle génération, opérant en mode services.

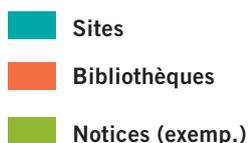
L'Abes, porteur naturel de ce projet de mutualisation, a ainsi proposé à l'ensemble des établissements déployés dans le réseau Sudoc de déclarer leur intention d'adhérer au groupement de commande, puis d'en signer la convention constitutive.

Figurent ci-contre et ci-dessous la liste des 50 établissements signataires, dont les 9 sites pilotes, la répartition des différents SIGB représentés au sein du groupement ainsi qu'un tableau comparatif sites SGBM/sites Sudoc selon les dimensions caractéristiques : ILN (*International Library Number*), RCR (numéro d'identification Sudoc) et notices d'exemplarisation. Ces informations montrent, d'une part, la diversité et la représentativité des sites signataires du groupement de commande au plan national et, d'autre part, que le groupement représente plus de la moitié des bibliothèques déployées dans le Sudoc.

LES 12 SIGB REPRÉSENTÉS AU SEIN DU GROUPEMENT DE COMMANDE



PART RELATIVE DU GROUPEMENT DE COMMANDE SGBM VIS-À-VIS DU RÉSEAU SUDOC



CHIFFRES

ILN : 48 sur 164, soit environ 30 %
 RCR : 835 sur 1 451, soit environ 58 %
 Notices : 17,6 M sur 38,7 M, soit environ 45 %

LISTE DES ÉTABLISSEMENTS SIGNATAIRES

Sites pilotes membres du groupement de commande SGBM

BIU Montpellier
 PRES Clermont Université
 Universcience
 Université Charles-de-Gaulle Lille 3
 Université de Bordeaux
 Université de Toulouse - SICD
 Université du Havre
 Université du Littoral Côte d'Opale
 Université Paris Descartes

Autres établissements membres du groupement de commande SGBM

Campus Condorcet
 Casa de Velazquez
 Comue Sorbonne Paris Cité
 École française d'Athènes
 École nationale supérieure d'arts et métiers
 École normale supérieure de Lyon
 Inria
 Institut français du Proche-Orient - IFPO
 Institut national d'histoire de l'art - INHA
 Muséum national d'histoire naturelle
 * Sciences Po Paris
 Université Aix-Marseille
 Université de Bourgogne
 Université de Bretagne Occidentale
 Université de Caen Basse-Normandie
 Université de Franche-Comté
 Université de la Polynésie française
 Université de La Réunion
 Université de La Rochelle
 Université de Lorraine
 Université de Pau et des Pays de l'Adour
 Université de Perpignan Via Domitia
 Université de Poitiers
 Université de Rouen
 Université de Savoie
 Université de technologie Belfort-Montbéliard
 Université des Antilles et de la Guyane
 Université d'Évry-Val-d'Essonne
 Université d'Orléans
 Université du Sud-Toulon-Var
 Université Grenoble 1 INPG
 Universités Grenoble 2 et 3
 Université Lille 2 Droit et Santé
 Université Nice Sophia Antipolis
 Université Panthéon Assas
 Université Paris 13 Nord
 * Université Paris Diderot Paris 7
 Université Paris Est Marne-la-Vallée
 Université Paris Sorbonne
 Université Pierre et Marie Curie
 * Université Sorbonne Nouvelle Paris 3

* Au titre de la Comue Sorbonne Paris Cité

(Agenda)

Janvier

LE 29

ARCHITECTURE ET BIBLIOTHÈQUE : FLUX ET FORMES DES ESPACES DE LA CULTURE

Lyon, Amphithéâtre de la Bibliothèque municipale

Cette journée d'étude, organisée par l'Enssib et la Bibliothèque municipale de Lyon, se penchera sur les questions de flux et de formes au sein des espaces de la culture (bibliothèques, mais aussi musées, théâtres, archives...). Ces derniers ne sont pas redevables d'une seule logique de « boîtes » : de multiples flux les traversent (hommes, choses, signes), qu'il convient d'interroger. Comment amener les flux dans la « boîte », comment les réguler, comment les faire circuler entre les boîtes ? Mais aussi, de manière symétrique, comment fluidifier la circulation grâce aux formes et à l'information ?
www.enssib.fr/journee-architecture-et-bibliotheque-2015

Février

DU 9 AU 12
CODE4LIB

Portland, Oregon (USA)

Le congrès Code4lib rassemble, depuis 2006, les acteurs de l'informatique documentaire à l'échelon international. Il facilite l'échange sur les différents pratiques et la mise en place de collaborations ou de partenariats.
<http://code4lib.org>



DERNIÈRES PARUTIONS AUX PRESSES DE L'ENSSIB

Produire des contenus documentaires en ligne

Coordonné par Christelle Di Pietro, « La Boîte à outils », novembre 2014

Les professionnels des bibliothèques sont devenus médiateurs, producteurs et éditeurs de matières documentaires numériques. Une quinzaine de professionnels abordent ces différents aspects, en s'appuyant sur des réalisations concrètes menées dans les bibliothèques, de lecture publique comme des universités, pour évoquer les nouvelles compétences relationnelles et techniques à acquérir.

Le plan de l'ouvrage s'articule autour de quatre parties : exploiter les collections ; la curation et la production de contenus ; produire en réseau, et, enfin, connaître le contexte juridique et écrire pour le web.

La documentation dans le numérique

Olivier Le Deuff, « Papiers », décembre 2014

Les bibliothécaires, les documentalistes ne disparaîtront pas avec l'avènement du document numérique : leurs métiers vont changer, et sans doute s'enrichir. Olivier Le Deuff définit ce qu'est une culture de l'information, pour les professionnels ainsi que pour les usagers. Il développe longuement les rôles et missions des bibliothécaires et documentalistes dans ce monde technique (informatique) et professionnel (le classement et les architectures de l'information). Il précise la place nouvelle des lecteurs, usagers, internautes, et les fonctions indispensables de médiation qui sont attachées à ces nouveaux supports. Il replace la formation de chacun comme préalable à ces nouvelles compétences.

→ www.enssib.fr/lecole/preses-de-lenssib

Mars

LE 16

RENCONTRES HENRI-JEAN MARTIN

Villeurbanne, Amphithéâtre de l'Enssib

Depuis neuf ans, les rencontres Henri-Jean Martin proposent un lieu de rencontre, de partage d'expériences et d'expertises pour les bibliothécaires travaillant sur des fonds patrimoniaux de toutes sortes. Si les objectifs initiaux n'ont pas changé, ils prendront corps dans une formule renouvelée : un programme d'une journée ouverte par un forum. Intitulées « *Le patrimoine #tropdelaballe* », ces rencontres s'intéresseront à la relation entre le patrimoine et les jeunes publics (11-15 ans), tant du point de vue de l'offre des professionnels que de celui de la réception par les jeunes.

www.enssib.fr/rencontres-henri-jean-martin-2015

DU 18 AU 19
DOCUMATION 2015

Paris, CNIT La Défense

Rendez-vous annuel dédié aux solutions proposées pour résoudre les problèmes liés à la gestion de contenus numériques, Documation accueille plus de 150 exposants de ce secteur et programme de nombreuses conférences et ateliers de démonstration.

www.documation-mis.fr

DU 20 AU 23
SALON DU LIVRE DE PARIS

Paris, Porte de Versailles

Alors que les lettres brésiliennes seront à l'honneur de cette 35^e édition du Salon du livre de Paris, les organisateurs inaugurent Bibliodoc#2015 « *Enjeux, culture & prospective* », un espace réservé aux bibliothécaires autour de quatre jours de conférences pratiques et théoriques pour aborder les nombreuses questions qui traversent la profession.

www.salondulivreparis.com

DU 30 MARS AU 1^{ER} AVRIL
UKSG 2015

Glasgow (Royaume-Uni)

UKSG, organisme international ouvert aux professionnels concernés par la fourniture d'information scientifique et la recherche électronique, organise son 38^e congrès annuel qui réunira des bibliothécaires, des éditeurs, des intermédiaires et des fournisseurs de solutions technologiques.

www.uksg.org/event/conference15



À NOTER DÈS À PRÉSENT

LES PROCHAINES JOURNÉES ABES se tiendront au Corum de Montpellier le mercredi 27 et le jeudi 28 mai. Le programme complet et le formulaire d'inscription seront accessibles à partir de la mi-mars sur le site web de l'Agence.