

Diplôme de conservateur de bibliothèque

Mémoire d'étude / Janvier 2017

**Organisation des données,
organisation du travail en
bibliothèques universitaires à l'heure
du Big Data**

Luc Bellier

Sous la direction de Nathalie Marcerou-Ramel
Directrice des études et des stages – Enssib

Remerciements

Ce travail n'aurait pas été possible sans le soutien et l'appui de Nathalie Marcerou-Ramel dont les conseils et la rigueur m'ont été d'un grand soutien, ni sans Emmanuelle Bermès qui a inspiré l'origine et la conclusion de ce mémoire.

Il doit également beaucoup à tout ceux qui ont pris le temps de me recevoir et de partager leur expérience.

Ce travail s'il a été mené pendant cette année de scolarité à l'Enssib, est alimenté par ces 20 années passées dans de nombreux établissements. J'y ai rencontré de nombreuses personnes de qualité qui m'ont fait aimer ce métier. De cette galaxie de rencontres je voudrais remercier les membres passés et présents d'Orhion. Ce travail leur doit beaucoup. J'espère ne pas avoir trop trahi ici l'esprit qui nous a animés. Ma reconnaissance va également aux nombreux comparses de la TIN7 dont les conversations presque toujours professionnelles pouvaient nous mener tard dans la soirée. Les bureaux de Dominique, Sébastien, Catherine, Lionel, Marie-Elise, Gildas, Isabelle, dont les portes ouvertes éclairaient jusque tard dans la nuit le couloir silencieux se souviennent encore de mes questions tardives. J'adresse mes remerciements et ma gratitude à ceux qui chacun à leur manière m'ont aidé et soutenu ces dernières années : Brigitte, Arnaud, Céline, Isabelle, Emmanuelle, Marianne, Pauline et Sébastien.

Je voudrais également témoigner ma reconnaissance aux équipes croisées à Noisy-Le-Sec. J'y ai appris plus qu'ailleurs ce qu'était le sens d'un service public, grâce à Cyril et quelques autres.

Je voudrais encore remercier ma famille qui m'a apporté son soutien à ce projet professionnel et témoigne encore de beaucoup de patience et de constance. Que Marie, mon épouse, Nicole, ma mère, et mes enfants trouvent ici le témoignage de ma reconnaissance.

Résumé :

Les bibliothèques universitaires sont confrontées à la multiplication des silos de données dont la nature et la structure sont très éloignées de celles du catalogue bibliographique. Organisées depuis de nombreuses années autour du catalogue et de la chaîne de traitement documentaire, les bibliothèques doivent apprendre à se structurer autour de ces nouvelles données. Ce travail étudie les conséquences organisationnelles, et métier qui peuvent s'observer dans un tel contexte.

Descripteurs :

Données massives

Bibliothèques et internet

Sociologie des organisations

Abstract :

University libraries are confronted with the proliferation of data warehouse whose nature and structure are far from the bibliographic catalog. Organized for many years around the catalog and the document processing chain, libraries must learn to structure themselves around this new data. This study is a reflexion on the organizational, and business consequences that can be observed in such a context

Keywords :

Big Data

Libraries and internet

Droits d'auteurs



Cette création est mise à disposition selon le Contrat :
« **Paternité-Pas d'Utilisation Commerciale-Pas de Modification 4.0 France** »
disponible en ligne <http://creativecommons.org/licenses/by-nc-nd/4.0/deed.fr> ou par
courrier postal à Creative Commons, 171 Second Street, Suite 300, San Francisco,
California 94105, USA.

Sommaire

| | |
|--|------------|
| SIGLES ET ABREVIATIONS | 7 |
| INTRODUCTION..... | 9 |
| LES ENJEUX D'UNE CONCEPTION ELARGIE DES DONNEES : OBJET ET METHODE | 13 |
| 1.1. Problématique générale | 13 |
| 1.2. Présentation de l'enquête et des données collectées | 14 |
| 1.3. Etat de l'art et questions identifiées | 15 |
| 1.4. Les révolutions de la data | 17 |
| 1.5. Extension du domaine des données en bibliothèque..... | 20 |
| <i>1.5.1. Les données internes à la bibliothèque.....</i> | <i>21</i> |
| <i>1.5.2. Les données externes à la bibliothèque</i> | <i>23</i> |
| <i>1.5.3. Disruption et position de la bibliothèque par rapport à ses données.....</i> | <i>27</i> |
| ORGANISATIONS ET DONNEES | 30 |
| 2.1. La place des données dans l'organisation..... | 31 |
| <i>2.1.1. Outils, données et maturation des processus de travail</i> | <i>31</i> |
| <i>2.1.2. Maîtrise des données, maîtrise des organisations.....</i> | <i>35</i> |
| <i>2.1.3. La place du SCD et des données</i> | <i>42</i> |
| 2.2. Mises en œuvre de la gestion des données..... | 46 |
| QUELS CHANGEMENTS POUR LE TRAVAIL EN BIBLIOTHEQUE ?.... | 57 |
| 3.1. l'utilisateur au cœur ... ou la stratégie du changement | 57 |
| 3.2. Statuts, culture métier et représentations professionnelles..... | 64 |
| CONCLUSION | 69 |
| BIBLIOGRAPHIE..... | 72 |
| ANNEXES..... | 81 |
| INDEX | 99 |
| TABLE DES MATIERES..... | 103 |

Sigles et abréviations

| | |
|---------------|---|
| ADBU : | Association des Directeurs de Bibliothèques Universitaires |
| ANR : | Agence Nationale pour la Recherche |
| API : | Application Programming Interface / Interface applicative de programmation |
| BnF : | Bibliothèque nationale de France |
| BU : | Bibliothèque Universitaire |
| BSN : | Bibliothèque Scientifique Numérique |
| CDO : | Chief Digital Officer |
| CINES : | Centre Informatique National de l'Enseignement Supérieur |
| CNRS : | Centra National de la Recherche Scientifique |
| COMUE : | COMunauté d'Universités et d'Ecole |
| DMP : | Data Management Plan / Plan de Gestion de données |
| DRM : | Digital Right Management / Système de gestion des droits numériques |
| DSI : | Direction /Département des systèmes d'Information |
| Enssib : | Ecole Nationale Supérieure des Sciences de l'Information et des Bibliothèques |
| ESGBU : | Enquête Statistique Générale des Bibliothèques Universitaires |
| ETP : | Equivalent Temps Plein |
| FMSH : | Fondation de la Maison des Sciences de l'Homme |
| FOAF : | Friend Of A Friend |
| FRBR : | Functional Requirements for Bibliographic Records / Fonctionnalités requises des notices bibliographiques |
| GIS : | Groupement d'Intérêt Scientifique |
| GED : | Gestion Electronique des Documents |
| H2020 : | Horizon 2020 Plan de financement européen de la recherche et de l'innovation |
| Huma-Num : | Très Grande Infrastructure de la Recherche dédiée aux Humanités Numériques |
| IdRef : | Identifiants et référentiels pour l'enseignement supérieur |
| INIST : | Institut National de l'Information Scientifique et Technique |
| INSEE : | Institut National des Statistiques et des Etudes Economiques |
| IoT : | Internet of Things / Internet des Objets |
| ISNI : | International Standard Name Identifier |
| ISO 9001 : | Norme définissant les règles de management de la qualité |
| Jussieu PMC : | Campus de Jussieu de l'Université Pierre et Marie Curie |
| LABEX : | Laboratoire d'Excellence, financé sur projet par les Investissements d'avenir |
| LRU : | Loi relative aux Libertés et Responsabilités des Universités |
| MARC : | MACHine READable Cataloging ; premier format informatique de description bibliographique |
| MCC : | Ministère de la Culture et de la Communication |
| OAI-PMH : | Open Archives Initiative – Protocol for Metadata Harvesting |
| OCLC : | Online Computer Library Center |
| Orcid : | Open Research and Contributor Identifier |
| PNB : | Prêt Numérique en bibliothèque |

| | |
|-----------|---|
| RDF : | Ressource Description Framework |
| RFID : | Radio Frequence Identification |
| SCD : | Service Commun de la Documentation |
| SGBM : | Système de Gestion de Bibliothèque Mutualisé |
| SI : | Système d'Information |
| SIGB : | Système Intégré de Gestion de bibliothèque |
| SUDOC : | Système Universitaire de Documentation |
| TDM : | Text and Data Mining |
| UE : | Union Européenne |
| URFIST : | Unité Régionale de Formation à l'Information Scientifique et Technique |
| URI : | Uniform Resource Identifier |
| URL : | Uniforme Ressource Locator |
| UX : | User eXperience / expérience usager |
| VIAF : | Virtual International Authority File |
| XML : | eXtensible Markup Language |
| Z-39-50 : | protocole d'échanges de données défini par la norme internationale ISO 23950 |

INTRODUCTION

Entre transition bibliographique et données de la recherche, outils de pilotage et révolution des usages, les bibliothèques universitaires sont saisies de différents dossiers dont l'importance est régulièrement rappelée¹. Aussi divers soient-ils, ce sont des données qui les alimentent ces outils, les documentent, et les fondent. Ces données sont également hors du silo du catalogue où pourtant les bibliothèques ont l'habitude de travailler. Ainsi le développement de ces activités se fait-il à coté de la chaîne de traitement du document, en dehors des cadres organisationnels qui structurent le travail depuis plusieurs décennies dans les bibliothèques, posant la question de l'adaptation de son organisation.

Evoluer vers d'autres schémas peut cependant s'avérer délicat car en rupture avec des années de représentation des rôles et des métiers qui s'effectuent dans les bibliothèques autour de la chaîne de traitement du document ; d'autant plus que le pyramidage des postes et des statuts est en relation avec cette organisation. N'est-ce pas d'ailleurs pour accompagner ces évolutions que l'IGB a réfléchi à une redéfinition des rôles de chaque corps des bibliothèques²?

Pour autant, contraints à une efficacité tangible par l'intégration complète des BU à l'université depuis la loi LRU³ et la mise en place des compétences élargies⁴, les services documentaires des universités doivent, comme beaucoup d'autres services publics, répondre à une double injonction : adapter leurs services au contexte numérique et aux évolutions des usages et des besoins des utilisateurs tout en se montrant efficaces, ce qui suppose une agilité des organisations, des compétences et une forte capacité à s'évaluer, d'une part, et utiliser des ressources humaines principalement issues de la fonction publique et disposant des compétences acquises durant leur formation⁵, d'autre part. Si ce contexte est connu et identifié depuis quelques années, la multiplication des projets sous la responsabilité de la bibliothèque, consistant à construire de nouveaux silos de données positionnés à côté des précédents, pose la question des compétences, de la culture et de l'identité professionnelle, des organisations de travail et de leur capacité à s'adapter à des contextes mouvants, fluides ou instables selon les points de vue et les contextes. C'est ce double contexte qu'il s'agit de questionner ici :

D'une part, une organisation du travail contrainte à l'efficacité, inscrivant ses missions dans un contexte élargi par le numérique, et par l'inscription des BU dans leur écosystème universitaire ; d'autre part, une organisation qui doit s'adapter en même temps qu'elle s'approprie ces sujets.

Les bibliothèques sont par ailleurs confrontées à des enjeux majeurs autour des données, enjeux traversés par deux mouvements, l'Open Data, d'une part, et le Linked Data, d'autre part.

1 Journées d'études, colloques, congrès d'associations professionnelles, projets européens, textes législatifs et publications diverses en témoignent.

2 Quels emplois pour les bibliothèques? Etat des lieux et perspectives - Livre et Lecture - Ministère de la Culture et de la Communication [en ligne]. [Consulté le 11 juillet 2016]. Disponible à l'adresse : <http://www.culturecommunication.gouv.fr/Politiques-ministerielles/Livre-et-Lecture/Documentation/Publications/Rapports-de-l-IGB/Quels-emplois-pour-les-bibliotheques-Etat-des-lieux-et-perspectives>.

3 LOI n° 2007-1199 du 10 août 2007 relative aux libertés et responsabilités des universités

4 MARTINEZ, Marc, « Les bibliothèques dans l'université après la LRU » in CAVALIER, François et POULAIN, Martine, 2015. Bibliothèques universitaires: nouveaux horizons. Paris : Éditions du Cercle de la librairie. Bibliothèques.

5 MARCEROU-RAMEL, Nathalie, « Nouvelles compétences, nouvelles organisations » in CAVALIER, François et POULAIN, Martine, 2015. Bibliothèques universitaires: nouveaux horizons. Paris : Éditions du Cercle de la librairie. Bibliothèques

Le premier mouvement, qu'on va qualifier ici de politique au sens de politique publique, concerne l'ouverture des données publiques. Sous l'impulsion de la mission Etalab⁶, depuis 2011, mission renforcée par le vote de la loi Pour une République Numérique, notamment l'article L. 321-4 du *Code des relations entre le public et l'Administration*⁷, les services publics et donc les universités, sont incités à rendre publiques et exploitables les données qu'ils produisent. Ce n'est pas l'objet de ce travail que de traiter de l'ouverture des données publiques, mais ce contexte témoigne de l'importance pour les SCD d'aborder la question des données et de créer l'organisation nécessaire à leur publication et exploitabilité.

Le second mouvement traverse et mobilise la profession afin de faire évoluer les données et de les rendre indexables sur le web. C'est ce qu'on appelle le web sémantique, web de données ou Linked Data. Il traverse de nombreux dossiers allant de la FRBÉRisation des catalogues aux services enrichis proposés par les bibliothèques numériques.

Le Big Data en tant que tel est rarement abordé dans les bibliothèques⁸. C'est la raison pour laquelle aussi bien dans le questionnaire en ligne que dans les entretiens en tête à tête, la question du Big Data est rarement abordée de front. Le Big Data est la conséquence d'un processus plus ou moins avancé dans les bibliothèques, mais n'est pas encore un sujet traité en tant que tel. Pour autant, le choix de conserver cette terminologie s'est rapidement imposé au cours de ce travail, car les volumétries dont il est question et les finalités sur lesquelles travaillent les bibliothèques relèvent bien du Big Data, d'une part, et du Linked Data, d'autre part.

L'étude du Big Data en bibliothèque est donc un domaine relativement inexploré en tant que tel, surtout si on l'aborde avec un regard métier et organisation. Cependant si l'on prend une acception large du terme Big Data embrassant le Linked Data, le stockage et l'organisation et l'exploitation de toutes les données en dehors du silo du catalogue, il occupe une part significative des publications professionnelles. Mais ce sont les questions techniques qui sont largement traitées dans ces ouvrages. S'il convient de mobiliser ces publications dans le travail qui suit, il ne s'agit néanmoins pas d'en faire une simple synthèse mais d'essayer de cerner ce qui change avec ces technologies, et que l'informatisation des catalogues n'avait pas changé. Pour cela, nous nous sommes appuyé sur trois éléments :

- Une enquête qualitative a été menée auprès de dix professionnels en situation de responsabilité fonctionnelle sur des sujets liés à la gestion des données, d'une part, ou en situation de responsabilité hiérarchique dans des bibliothèques universitaires ayant pris en charge d'une façon ou d'une autre des silos de données non bibliographiques d'autre part. Données de la recherche, informatique décisionnelle, visualisation de données, archives institutionnelles, Linked Data constituent les principaux thèmes que j'ai pu étudier en lien avec

6 Etalab : Mission créée en 2011 chargée de la politique d'ouverture et de partage des données publiques, au sein du secrétariat général du Gouvernement

7 Code des relations entre le public et l'administration | Legifrance, [en ligne]. [Consulté le 20 décembre 2016]. Disponible à l'adresse : <https://www.legifrance.gouv.fr/affichCode.do?cidTexte=LEGITEXT000031366350>

8 Le récent mémoire Enssib, GILIUM, Johann, Big Data et bibliothèques traitement et analyse informatiques des collections numériques. Mémoire d'étude. Enssib. Mémoire DCB. traite en fait du Text et Data Mining et des enjeux pour les bibliothèques. On peut citer le billet d'Emmanuelle Bermès sur son Blog sur le sujet. BERMÈS, Emmanuelle, 2015. Big Data et bibliothèques., [en ligne]. 2015. Disponible à l'adresse : <https://figoblog.org/2015/01/13/big-data-et-bibliotheques/> consulté le 12/12/2016

On peut également mentionner Affordance et Hubert Guillaud qui alimentent la profession de leurs travaux mais dans un contexte plus large que celui des bibliothèques.

les questions d'organisation et de management. Huit personnes ont été interrogées en présentiel, une par messagerie et une par téléphone. Les entretiens se sont déroulés entre juillet et septembre 2016. Deux entretiens complémentaires ont été menés en octobre et décembre 2016. Un des établissements étudiés dans ce cadre est basé aux USA.

- Une enquête quantitative menée en ligne entre septembre et novembre 2016 s'est révélée nécessaire pour compléter les données récoltées lors des entretiens. En effet, plusieurs questions portant sur le statut, la position fonctionnelle des personnes en charge de dossiers Data, avaient donné des résultats contrastés dans les établissements rencontrés lors de l'enquête qualitative. Afin de mieux cerner les profils, les parcours et la position des personnes concernées, une approche quantitative complémentaire était nécessaire. 28 réponses de différents établissements ont été collectées donnant un éclairage plus étayé sur ces questions. Le questionnaire s'est également intéressé aux centres d'intérêt des SCD pour les données. Ainsi ces éléments ont-ils pu donner un éclairage quantitatif aux différents domaines traités ou non dans les bibliothèques ayant répondu.
- Une approche par analogie qui impliquait de compléter l'enquête par l'étude de différents travaux. Dans un premier temps, ce sont des travaux techniques (informatiques) qui m'ont permis de cerner le sujet et de justifier une vision extensive du Big Data dans le contexte des bibliothèques allant du Linked Data à l'informatique décisionnelle. Cette étape était également nécessaire à la maîtrise des enjeux métiers et organisationnels soulevés par ces technologies. Des travaux sur la sociologie du travail, la sociologie des organisations et sur les identités professionnelles ont permis de confronter les situations des bibliothèques rencontrées en entretien à des situations connues et étudiées par ailleurs. Enfin, le numérique fait en tant que tel l'objet de travaux sociologiques féconds et qui se sont avérés très utiles pour déterminer les liens possibles entre l'organisation du travail et l'organisation des données.

Dans un premier temps il s'agit de présenter ce qui peut être qualifié de Big Data en bibliothèque. En identifiant les axes où se dessinent déjà ces pratiques, ces technologies, il s'agit aussi de cerner les enjeux stratégiques et métier qu'elles représentent et de dessiner la position dans laquelle se trouvent les bibliothèques vis à vis des données et de leur usage. Il sera alors possible de questionner la place des données dans le travail et son organisation en bibliothèque, pour enfin mesurer les ressorts des changements en cours entre décentrage des stratégies d'établissement, rupture managériale et permanence des cultures métier.

LES ENJEUX D'UNE CONCEPTION ELARGIE DES DONNEES : OBJET ET METHODE

1.1. PROBLEMATIQUE GENERALE

La loi pour une République numérique consacre le Text and Data Mining (TDM)⁹ comme une exception au droit d'auteur afin de favoriser les usages de la recherche. Cette loi en modifiant l'article 211-1 du *Code du Patrimoine* fait des données des archives¹⁰.

Dans ce contexte, il n'est plus nécessaire de démontrer le caractère central de la place des données dans l'activité des bibliothèques. Mais que sont ces données en bibliothèque ?

Si les données de la recherche tendent à devenir un sujet central dont se saisissent une part significative de SCD et d'universités, notamment par le biais d'incitations de plus en plus fortes au niveau européen (H2020¹¹), d'autres champs semblent émerger dans les BU. La collaboration accrue avec les autres services universitaires, sous l'impulsion de la Loi LRU, encourage les SCD à adopter des outils de pilotage en usage dans leur université, à utiliser les données d'autres services pour leurs propres besoins et pour ceux des tutelles¹². Cela témoigne de l'intégration des SCD au fonctionnement des institutions universitaires¹³. D'après l'enquête menée pour ce mémoire, l'informatique décisionnelle n'est pas moins implantée dans les SCD que les données de la recherche. Cela implique d'alimenter un entrepôt de données dédiées au pilotage et d'en assurer la qualité et l'exploitabilité.

D'autres silos se constituent par nécessité afin de mesurer toutes les activités de la BU y compris celles réalisées à distance. De nombreux établissements utilisent des outils tels qu'ezPAARSE¹⁴ et collectent des logs de connexion à leur site web pour mesurer l'audience des ressources électroniques et des services en ligne. Ainsi, ils collectent des données qui sont par essence le cœur, l'origine du Big Data : des données pauvres, faiblement structurées, mais dont la volumétrie constitue la principale richesse.

On peut multiplier ainsi les exemples d'utilisation et de collecte de données dans les SCD de France : données RFID, données liées à la numérisation de collections, données générées par les programmes d'Humanités Numériques dont

9 Loi Lemaire : OpenAccess et TDM, ADBU, consulté le 11:10/2016 sur <http://adbu.fr/loi-lemaire-open-access-et-tdm/>; Loi pour une République numérique définitivement adoptée par le Parlement, Couperin, <http://www.couperin.org/news/item/1277-loi-pour-une-republique-numerique-definitivement-votee>

10 Code du patrimoine - Article L211-1., [en ligne]. [Consulté le 20 décembre 2016]. Disponible à l'adresse : https://www.legifrance.gouv.fr/affichCodeArticle.do?sessionId=6CBC79F21B4F3230BA098B9DA69453BA.tpdila09v_2?idArticle=LEGIARTI000032860025&cidTexte=LEGITEXT000006074236&dateTexte=20161011

11 Horizon 2020 Plan de financement européen de la recherche et de l'innovation

12 Si le Ministère de l'Enseignement Supérieur et de la Recherche demande depuis de nombreuses années des indicateurs, la loi LRU faisant dépendre les SCD des seules universités, ces dernières sont demandeuses d'outils de suivi d'activité moins standardisés, et plus en rapport avec leur fonctionnement.

13 MARTINEZ, Marc, « Les bibliothèques dans l'université après la LRU » in CAVAILIER, François et POULAIN, Martine, 2015. Bibliothèques universitaires: nouveaux horizons. Paris : Éditions du Cercle de la librairie. Bibliothèques.

14 <http://ezpaarse.couperin.org/> (consulté le 12/12/2016) outil de visualisation des données de fréquentation des ressources documentaires en ligne des SCD, à partir des logs de connexions collectés par les serveurs de l'université.

les bibliothèques sont des partenaires réguliers, données sur la fréquentation (service en ligne « Affluences »¹⁵)...

Ces exemples illustrent une part du défi des bibliothèques universitaires pour continuer à servir leur communauté : apprendre à gérer non pas un silo durable dans sa nature et ses modalités d'usage et de gestion, mais un nombre croissant de types de données, qu'on ne sait pas toujours qualifier à leurs débuts, ni comment les gérer, ni comment les utiliser, mais qu'il faut s'approprier, organiser, et adapter pour qu'émerge le cadre de travail où elles seront utilisées par la communauté universitaire.

Le propos de ce travail est d'observer les réponses organisationnelles apportées par ceux qui gèrent de nouveaux types de données dans un cadre technique et procédural encore balbutiant. Entre production de données, traitement de ses données, ou de celles de tiers, et collecte de données pour des tiers, quel(s) rôle(s) se dessinent pour les bibliothèques ?

1.2. PRESENTATION DE L'ENQUETE ET DES DONNEES COLLECTEES

Dix entretiens ont été menés auprès de personnes travaillant dans des bibliothèques universitaires ou de grandes institutions de recherche. La plupart des entretiens ont duré deux heures ce qui a permis un approfondissement des questions abordées¹⁶. Neuf de ces entretiens concernent des établissements basés en France et un aux USA.

Parallèlement, un questionnaire en ligne a permis de collecter des informations sur 28 établissements universitaires ou de recherche en France, soit entre 15 et 20% des établissements. Le mode de collecte sur la base du volontariat a pu favoriser les établissements concernés par le sujet. Cependant, la variété des silos envisagés dans le questionnaire ouvre un large spectre laissant espérer une relative représentativité de la collecte.¹⁷

Le questionnaire propose huit catégories de données ou de projets dont on verra qu'ils ne sont pas étanches et peuvent même se recouper : données de la recherche, archives institutionnelles, données collectées automatiquement par RFID ou autres objets connectés, données de services partenaires, visualisation de données, FRBR, Informatique décisionnelle. Sur 28 établissements,

- 21 ont plusieurs projets en cours sur les sujets considérés
- 2 ont répondu négativement à l'ensemble de ces 8 catégories,
- 5 ont répondu positivement à une seule de ces catégories. Tous travaillent sur des silos de données qui ne concernent pas le catalogue

Donc sur 28 établissements, 26 sont concernés par la multiplication des silos en dehors du catalogue. La multiplication est d'autant plus notable que sur l'échantillon 15 établissements ont au moins trois projets liés à des données en

¹⁵ Service en ligne développé par une société française permettant de connaître le taux d'occupation des places d'une bibliothèques, à partir des données collectées en temps réel dans les bibliothèques clientes, via les portiques de sécurité ou un matériel équivalent mis à disposition par « Affluences »

¹⁶ Les deux questionnaires sont restitués en annexe, ainsi que les données quantitatives anonymisées de l'enquête en ligne. Les entretiens ont été enregistrés mais ne sont pas restitués ici dans leur intégralité, l'anonymat ayant été demandé pour certains aspects des entretiens.

¹⁷ Voir annexes

dehors du catalogue traditionnel. Ceci traduit donc bien un mouvement de fond que les établissements abordent sous différents angles, selon le contexte local.

14 collectent des données automatiquement à des fins de suivi à partir d'objets connectés divers et 5 prévoient de le faire prochainement.

11 disposent dorénavant déjà d'une archive institutionnelle pour gérer les publications électroniques et 5 prévoient d'en avoir une prochainement¹⁸.

9 travaillent sur la FRBÉRisation du catalogue.

8 collectent des informations auprès des services partenaires de l'université

5 disposent d'outils en informatique décisionnelle

3 font de la data visualisation et 2 prévoient de le faire

Aucun ne dispose de système de gestion des données de la recherche mais 5 prévoient de s'en doter.

1.3. ETAT DE L'ART ET QUESTIONS IDENTIFIEES

Selon ce qu'on considère comme relevant du Big Data, les questions qui se posent relèvent soit de la prospective car les dossiers sont à l'état de projet, soit de réalités bien établies dans les établissements concernés. Mais dans tous les cas on rencontre un changement de perspective ou d'outils qui agit, modifie les pratiques antérieures. Ainsi quand la Bibliothèque du Congrès parle désormais de Big Data, elle parle en fait de ses collections numérisées¹⁹ bien avant d'être considérées comme telles.

Si la mise en place de projets Big Data est encore marginale et limitée notamment à des projets liés aux données de la recherche, ou à un ensemble de données de pilotage, un certain nombre de facteurs laisse à penser que leur développement dans nos structures n'est qu'une question de temps.

En janvier 2015, sur son blog, Emmanuelle Bermès voyait ainsi, d'un point de vue prospectif, en janvier 2015²⁰ trois pistes de développement du Big Data pour les bibliothèques :

- les données comme un élément des collections des bibliothèques, à collecter, à conserver, depuis les publications jusqu'aux données de la recherche
- un outil pour analyser les publics sur le modèle de ce que développent les grands opérateurs commerciaux afin de mieux connaître et mieux servir ses publics
- la mutation en cours des pratiques des chercheurs notamment en sciences sociales qui, s'appropriant les technologies portées par les Humanités Numériques, veulent aborder les collections globalement et non plus simplement les lire

Le Big Data constituerait donc pour les bibliothèques aussi bien une nouvelle dimension de leurs missions de collecte et de conservation, distincte des collections matérielles, qu'une opportunité technologique pour mieux connaître

¹⁸ Il convient de souligner que ce résultat de 11 archives institutionnelles existantes paraît très important. Il semble que la question ait été diversement comprise, malgré un effort de formulation le moins ambiguë possible, et certaines réponses positives doivent correspondre à une gestion dans HAL des collections de l'Université de tutelle. Cependant, même si ce résultat agrège de fait des pratiques différentes, et des contraintes différentes, cela constitue de fait un entrepôt supplémentaire pour les équipes de la bibliothèque et ne modifie pas le fond de l'analyse.

¹⁹ LESLIE Johnston, Digital Collections as Big Data, Library of Congress, Digital preservation 2012, Disponible à l'adresse http://www.digitalpreservation.gov/meetings/documents/ndiipp12/BigData_Johnston_DP12.pdf (consulté le 12/12/2016)

²⁰ BERMÈS, Emmanuelle, 2015. Big Data et bibliothèques., [en ligne]. 2015. Disponible à l'adresse : <https://figoblog.org/2015/01/13/big-data-et-bibliotheques/> consulté le 12/12/2016

leurs publics, et une nécessité pour répondre aux évolutions des besoins des chercheurs.

On voit ici la grande diversité de ce qui peut être qualifié de Big Data pour les seules bibliothèques. Cela provient sans doute du fait que par Big Data on entend en fait non seulement les données elles-mêmes mais également la façon de les interroger et de les faire parler grâce à des algorithmes qui savent exploiter la masse et l'architecture répartie des ressources de calcul et de stockage.

Ainsi Gautier Poupeau²¹, architecte des données à l'INA (Institut National de l'Audiovisuel), distingue deux « stratégies » d'algorithme :

- Une stratégie d'« *Analyse du comportement* », analyse des traces laissées par les usages et les pratiques des individus. On retrouve ici une des trois pistes de développement potentiel du Big Data pour les bibliothèques.

- Une stratégie d'« *Analyse des contenus et comparaison entre les contenus* » qui suggère aux usagers des contenus similaires et favorise la consommation ou la prolongation de la navigation sur le même site. Ces technologies sont celles que mettent en œuvre les Humanités Numériques ; elles sont évoquées dans la troisième piste proposées par E. Bermès pour le Big Data dans les bibliothèques. Certains établissements travaillent directement sur ces questions de visualisation des données, de la collection²² etc. Mais les bibliothèques sont loin d'être les seules institutions support à la recherche à se saisir des Humanités Numériques. Huma-Num, service du CNRS hébergé par la FMSH (Fondation de la Maison des Sciences de l'Homme) est un des principaux acteurs du domaine. Ce service a ainsi mis en ligne des outils et services proposant aux chercheurs ce type de technologies d'analyse sémantique des contenus, de construction de métriques dans des corpus textuels etc. (Voyant Tools, Nakalona, etc.)²³ Les corpus versés peuvent aussi bien être produits par les chercheurs eux-mêmes que issus des collections des bibliothèques. A elles donc de se saisir de ces nouveaux services pour proposer aux chercheurs des solutions ou des dispositifs facilitant leur travail.

Les bibliothèques pour favoriser le travail des chercheurs en Humanités Numériques constituent des collections exploitables par des outils adaptés. Par le biais des projets de recherche, dont les bibliothèques sont parties prenantes, elles apportent leur compétence sur les formats adaptés à telle ou telle recherche, sur la modélisation des données selon les objectifs des chercheurs et doctorants, sur les conditions de récupération des données produites par le projet en vue de leur archivage et de leur éventuelle réutilisation. C'est là un des enjeux du travail que mènent différents établissements sur les données de la recherche.²⁴ Au travers des projets sélectionnés par les LABEX (Laboratoire d'Excellence) différentes bibliothèques peuvent être retenues pour leurs collections mais aussi pour leurs compétences. Ainsi la BDIC est un partenaire régulier des projets du LABEX Passé dans le Présent²⁵.

21 POUPEAU, Gautier, 2016. De la responsabilité humaine des algorithmes | Les petites cases. [en ligne]. 16 octobre 2016. [Consulté le 18 décembre 2016]. Disponible à l'adresse : <http://www.lespetitescases.net/de-la-responsabilite-humaine-des-algorithmes>

22 C'est la démarche du Hackaton organisé par la BnF que de promouvoir l'usage de ces données avec des services ouverts et réutilisables. Mais la dimension de cet établissement et l'ampleur des données qu'elle propose la mettent dans une situation singulière.

23 Outil Voyant Tool déployé sur la plateforme de service du CNRS, Huma-Num <http://www.huma-num.fr/services-et-outils/traiter> et <http://voyant.tools.huma-num.fr> (consulté le 12/12/2016)

24 Lors des entretiens j'ai identifié la bibliothèque de Sciences-Po, et celle de Lyon 1

25 BDIC, LABEX « Les passés dans le présent », consulté le 20/10/2016, <http://www.bdic.fr/activites-de-recherche/labex-et-programmes-de-recherche> (consulté le 12/12/2016)

Les données sont désormais dans une position pivot, structurante de l'activité des bibliothèques, à côté des activités de production que sont la gestion des collections et les services aux publics étudiants. Quelle place reste-t-il à la chaîne de traitement du document dans l'organisation du travail ? Pour les bibliothèques c'est peut être, en effet, l'occasion de repenser la place des données dans la stratégie des établissements, et ainsi la place des BU dans l'écosystème universitaire.

1.4. LES REVOLUTIONS DE LA DATA

Dans le contexte des bibliothèques les enjeux autour des données sont traversés par deux mouvements, l'Open Data d'une part, et le Linked Data d'autre part.

Le Big Data et le Linked Data sont deux phénomènes parallèles qui aujourd'hui tendent à converger ou à se rapprocher.

Le premier, le Big Data, est né de questions d'infrastructures et de temps de réponses posées par la masse des données collectées sur le web par les grands opérateurs internet (Google, Amazon, Facebook et autres). Les solutions d'informatique décisionnelle ne permettaient pas de les traiter dans les formes voulues et avec un délai de réponse satisfaisant. De cette impossibilité ont émergé des technologies permettant de résoudre ces questions.

Le second, le Linked Data, est né de la volonté de mieux structurer le web en s'appuyant sur les ontologies de certains opérateurs (DBpedia²⁶ de Wikimedia²⁷, puis plus récemment Wikidata²⁸, ou encore FOAF²⁹ de Facebook³⁰, VIAF³¹ d'OCLC³² etc.) et sur les mécanismes permis par les URI (Uniform Resource Identifier) qui assure la pérennité et l'unicité des liens entre ontologies.³³ Le monde des bibliothèques s'est impliqué dans le développement du Linked Data. C'est d'ailleurs l'objet de différents projets d'ampleur impliquant des bibliothèques que d'intégrer et de mailler le web de données. On peut évoquer data.bnf.fr³⁴ qui aujourd'hui est largement relayé par Wikidata et DBpedia, mais aussi les identifiants ISNI³⁵ pour lequel la France via la BnF a eu un rôle pilote avec notamment OCLC et la British Library³⁶. Dans l'environnement des bibliothèques universitaires, on peut aussi signaler le SGBM³⁷ (Système de Gestion de Bibliothèque Mutualisé) qui vise à proposer à terme des solutions partagées de FRBérisation des catalogues des BU, c'est à dire de transformation des catalogues MARC en outils ouverts sur les standards du web de données et accessibles aux robots.

26 <http://fr.dbpedia.org/> (consulté le 12/12/2016)

27 <https://fr.wikipedia.org> (consulté le 12/12/2016)

28 <https://www.wikidata.org> (consulté le 12/12/2016)

29 <http://www.foaf-project.org/> (consulté le 12/12/2016)

30 <https://www.facebook.com/> (consulté le 12/12/2016)

31 <https://viaf.org/> (consulté le 12/12/2016)

32 *Ibid.*

33 Il faut souligner que l'URI est un mécanisme inventé par Tim Bernes Lee en même temps que les URL et donc dès les origines du web. Ce dernier craignant que les liens ne changent de dénomination avait imaginé une forme abstraite et non significative désignant une ressource sur le web afin de l'identifier indépendamment de son URL qui peut évoluer avec le temps.

34 <http://www.data.bnf.fr> (consulté le 20/20/2016)

35 Projet international d'immatriculation via le web de données des auteurs de façon non ambiguë (<http://www.isni.org/> consulté le 12/12/2016)

36 https://fr.wikipedia.org/wiki/International_Standard_Name_Identifier (consulté le 12/12/2016)

37 <http://www.abes.fr/Projets-en-cours/SGB-mutualise>

Progressivement le web de données s'élargi³⁸ et les bibliothèques et leurs données structurées y contribuent largement. Or, si la question des données est le plus souvent abordée sous cet angle en bibliothèque, du fait du contexte métier qui favorise une approche ontologique bien naturelle, cela n'exclut pas pour autant le caractère poids des données en jeu que les bibliothèques doivent traiter.

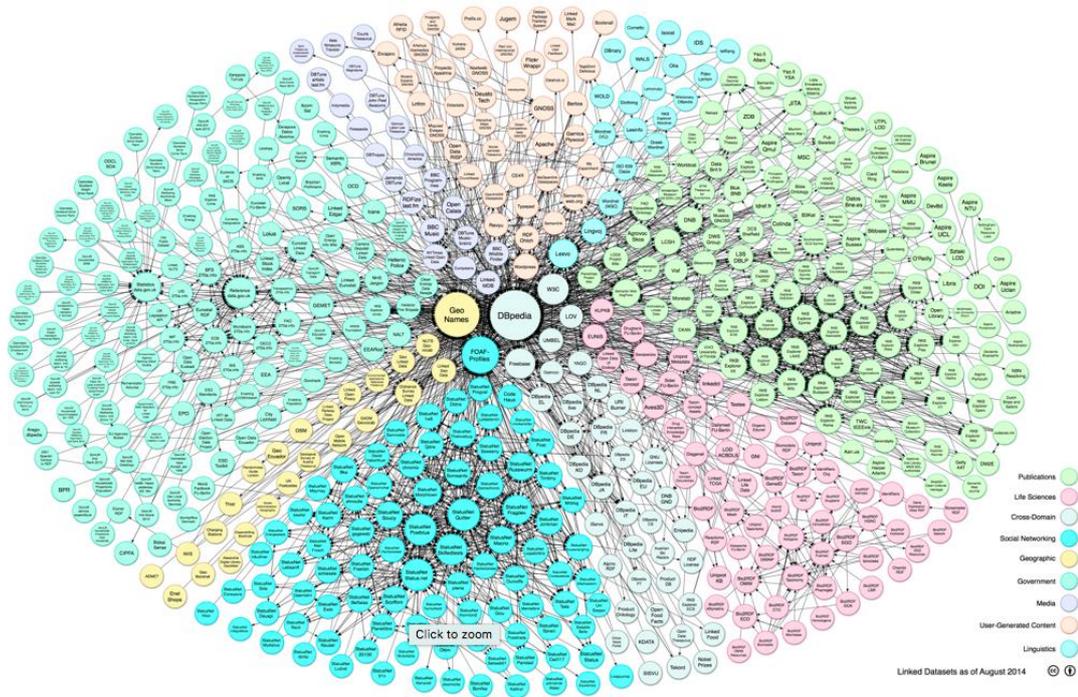


Schéma du web de données <http://lod-cloud.net/>

Le Big Data permet de traiter et d'organiser des informations parmi un grand nombre de données, le Linked Data d'appuyer l'organisation des données sur des ontologies viables et accessibles sur le web. Dans bien des cas, l'un vient en appui de l'autre. Cette présente étude essaie de prendre en compte tout travail sur des données hors du catalogue traditionnel. Il a été considéré dans ce mémoire, toutes formes de projets s'appuyant sur le web sémantique, le Big Data, ou traitant des données hors du SIGB et de leurs silos originels. Ainsi, toutes données échappant à la chaîne de traitement du document concernent l'étude, ainsi que les organisations de travail collectant, ou traitant ces données. On peut également inclure dans cette perspective, les démarches d'informatique décisionnelle qui sans encore véritablement s'appuyer sur des outils Big Data ou Linked Data visent à agglomérer un grand nombre de données éparses, à les associer et à les consolider afin de les faire parler. On peut également à ce stade souligner que dans le monde de l'entreprise, les outils du Big Data, et d'une certaine manière, ceux de l'informatique décisionnelle ont été mis en place ou ont évolué dans le but de mieux répondre aux besoins de leurs clients, de mieux les connaître et donc de mieux documenter leurs usages. C'est en collectant les traces de navigation, les requêtes des usagers, les modes de navigation qu'Amazon, Google, ou encore Facebook, se sont confrontés à deux questions :

³⁸ <http://lod-cloud.net/> (consulté le 12/12/2016)

- Comment gérer la massification des données sans perdre en fluidité ? Les réponses apportées sont celles qui définissent le Big Data dans la littérature informatique, à savoir les 3V : Volume, variété des données et vélocité de l'accès à ces données. D'ailleurs, c'est autant le volume des données que leur variété (image, son, texte etc.) qui a nécessité d'inventer de nouvelles solutions. Le modèle des bases de données relationnelles jusque-là utilisé pour gérer ces questions impose un caractère répétable et structuré de ce que l'on traite. Or la variété des données interdisait de répondre à cette contrainte et donc l'utilisation des bases de données relationnelles. On doit souligner à ce stade que les données de la recherche ont les mêmes contraintes de variété et imposent donc des technologies de type Big Data. Ce sont donc les architectures permettant une répartition des ressources en stockage et en calcul qui se sont ainsi élaborées (Hadoop³⁹ et autre). On pouvait alors passer à la seconde question.

- Que peut-on faire de ces données collectées ? Ces technologies nées dans des sociétés qui ne savaient pas forcément alors comment elles dégageraient des profits, ont d'abord répondu à la question du stockage et de l'accès rapide. Cela explique d'ailleurs la variété de langages en cours dans les technologies Big Data indépendamment des architectures existant derrière. Inspirée du modèle économique des médias⁴⁰, la connaissance des publics a été l'axe essentiel du travail sur les données du Big Data dans ces entreprises soit pour favoriser l'acte d'achat ou la consommation comme chez Amazon ou Netflix, soit pour revendre ces informations ou cibler de la publicité comme chez Google ou Facebook. Ce sont les algorithmes évoqués plus haut d'« analyse du comportement » ou d'« analyse des contenus et comparaison des contenus » qui ont permis d'exploiter ainsi les volumes de données et de les transformer en vecteur d'efficacité commerciale. L'expérimentation des modèles économiques guide le caractère expérimental de ces techniques et de ces algorithmes.

Avec des objectifs différents, les bibliothèques partagent avec ces entreprises la nécessité de mieux connaître l'usage qui est fait de leurs services. La question de la mesure de l'audience et de la connaissance des usages est un sujet déjà pris en main.

Les BU en utilisant des plateformes distantes aux ressources électroniques pour assurer l'accès aux périodiques électroniques s'est posée la question de la façon de maintenir la connaissance des publics et de leurs usages des collections. Désormais sur le web, les données collectées sont faiblement structurées et des outils spécifiques sont nécessaires pour les exploiter. C'est l'objectif d'ezPAARSE⁴¹. Ce logiciel libre permet de visualiser les connexions aux ressources électroniques fournies par le SCD, à partir des logs stockés sur les serveurs de l'université. Bien que récent, il est largement utilisé par les BU.

³⁹ <http://hadoop.apache.org/> logiciel développé initialement par Yahoo! qui permet de gérer un archivage distribué sur différentes machines de façon transparente pour les utilisateurs. Il permet également de répartir la puissance de calcul nécessaire à l'interrogation des données stockées sur les machines. Ce fonctionnement est en rupture radicale avec les systèmes de gestion de base de données dont l'architecture ne peut pas être distribuée.

⁴⁰ Voir CARDON, Dominique, 2015. A quoi rêvent les algorithmes: nos vies à l'heure des Big Data. Paris : Seuil. La République des idées.

⁴¹ <http://ezpaarse.couperin.org/> (consulté le 12/12/2016)

Dans le même esprit, la solution « *PREVU*⁴² » portée par Paris 8 propose un outil qui présente de nouveaux modes d'exploitation et de visualisation des données d'usage des collections par les publics de la BU. On retrouve avec PREVU qui est un programme de recherche associant des chercheurs de Paris 8, le caractère expérimental que nous évoquions pour l'usage que font les Gafa des solutions de connaissance de leurs publics.

Ces deux exemples sont des utilisations du Big Data qui se rapprochent de l'informatique décisionnelle pour les données d'usage, et d'audience. Même si ces démarches prolongent les chiffres demandés par l'ESGBU (Enquête de Statistique Générale des BU), ils accompagnent aussi le mouvement qui incite les professionnels des bibliothèques à placer les usagers au cœur de leur stratégie (Design Thinking, UX d'ailleurs sujet principal du dernier congrès de l'ADBU). Cette orientation des bibliothèques ne pourra que confirmer et alimenter la tendance déjà bien entamée, qui consiste à explorer les données pour mieux connaître les pratiques et les besoins.

Cependant, cela nécessite de s'interroger sur la diversité des données disponibles et de tenter de clarifier la position des bibliothèques par rapport à ces données.

1.5. EXTENSION DU DOMAINE DES DONNEES EN BIBLIOTHEQUE

Les différents exemples évoqués plus haut révèlent une grande variété de nature des données. Les établissements ayant déployé des stratégies de présence en ligne assez importantes et multicanal, disposent d'une quantité importante de données, liées aux collections physiques ou numériques, aux usages de ces collections, et à leurs usagers eux-mêmes.

Ces données sont disponibles dans les applications métiers (SIGB principalement mais aussi les applications de gestion financière), sur les serveurs de connexion de l'université, voire dans les systèmes de publication et de modération des sites de la bibliothèque (forum, réponses aux billets de blogs etc.) Elles sont parfois sur des plateformes distantes qui mesurent l'audience du site (Google Analytics ou Xiti) ou la fréquentation d'un lieu (Affluences). Elles peuvent aussi provenir de services partenaires (scolarité de l'université par exemple) ou totalement étranger à l'institution (INSEE qui peut fournir des informations sur les mouvements pendulaires -déplacement quotidien- vers la commune où se trouve la BU, sur les bassins de population des étudiants, leurs profils socio-économique etc.).

Ces données de natures distinctes s'intègrent au travail des bibliothèques de façon différente selon leurs provenances et leurs statuts. De même, elles ne sauraient être exploitées de la même manière pour des raisons techniques, de processus métier ou de statut juridique. Le procédé de collecte, d'exploitation et de visualisation des données varie selon les outils et le cadre juridique. S'en suit une

42 <http://prevu.fr/> (consulté le 12/12/2016) Plus récemment on a vu la BnF se lancer dans semblable projet pour ses publics sur site : <https://blogrecherche.wp.mines-telecom.fr/2016/11/03/bnf-sociologie-big-data-gallica/> (consulté le 12/12/2016)

exploitation plus ou moins contrainte des données concernées. La page Facebook d'une bibliothèque ne permettra de connaître les interactions avec les usagers que dans les outils proposés par Facebook, quand les logs de connexion au site de la bibliothèque laisse une totale liberté d'exploitation sous réserve du respect de la loi informatique et liberté et de disposer d'outils adéquats.

Il convient désormais d'essayer de présenter l'ensemble des données relevant de cette étude en les ordonnant selon leur provenance et donc selon les contraintes qui pèsent sur elles.

1.5.1. Les données internes à la bibliothèque

A ce stade, on peut distinguer les données internes à l'institution, qui relèvent de son activité, de sa responsabilité métier et parfois juridique (sinon via son université de tutelle). Elles se trouvent dans :

- les applications métiers : SIGB, workflow de numérisation, outil de gestion financière quand ces applications sont déployées sur les serveurs de l'université ou de la bibliothèque.
- les systèmes de suivi des objets connectés : compteurs d'entrée, suivi de température et d'hygrométrie des magasins ou salles d'exposition, automates de prêt. Depuis peu, certaines sociétés proposent des solutions intégrées de gestion des espaces publics d'une bibliothèque. Ces solutions pilotées et paramétrables permettent également de gérer la sécurité des accès salle par salle, l'éclairage, les annonces de fermeture etc. Selon les systèmes mis en place un certain nombre de données peuvent être collectées par ces dispositifs et enrichir les données exploitables. Les technologies autour des objets connectés (serrure, éclairage, RFID au-delà de la seule gestion du prêt et des inventaires, capteur de mouvement ou de température etc.) et les scénarios qui peuvent être mis en place avec ces différents éléments représentent probablement à la fois un fort facteur de développement des données disponibles mais aussi un enjeu significatif dans l'élargissement des gammes de services que les bibliothèques peuvent offrir. Ils restent à ce jour très peu développés en France au-delà des usages classiques (portique avec compteur, RFID pour les prêts et les inventaires) mais trouvent dans les systèmes de réservation en ligne de salles de travail une application en plein développement.
- Ce qu'on nomme aujourd'hui les traces, c'est à dire tous les éléments que laisse un utilisateur dans sa navigation, clic, déroulement de page, temps de connexion sur chaque page, commentaire sur un Forum etc. peuvent être collectées sur les serveurs qui gèrent les services en ligne. Ce sont ces traces qu'Amazon exploite en les croisant avec son catalogue marchand pour proposer aux internautes des produits qui pourraient leur convenir. Les bibliothèques peuvent disposer des mêmes données si elles le souhaitent. Ces données peuvent éventuellement être enrichies par d'autres stratégies de collecte. En effet, certains établissements, directement ou par l'intermédiaire de l'université proposent des applications pour terminaux mobiles. Ces applications peuvent servir à collecter des données plus riches que lors des connexions sur les pages web via un navigateur internet. Par prudence, les

services publics s'interdisent de collecter des données (géolocalisation, contact etc.) mais rien ne leur interdit de le faire. La collecte de données sur ces terminaux permettra d'enrichir les données d'usage, même si le succès des quelques applications existantes est très mitigé.⁴³

Plusieurs précisions doivent être apportées. Ces données sont dans la plupart des cas, stockées sur les machines des établissements concernés, bien que rarement exploitées au delà des pratiques habituelles de statistiques SIGB et quelquefois via des solutions d'informatique décisionnelle. Il existe aujourd'hui des solutions de SIGB dans le *cloud*. Elles proposent les mêmes services mais sans déploiement d'applications sur les serveurs et postes informatiques des établissements. Cela permet une plus grande facilité de maintenance pour les services informatiques. Cependant les données sont alors stockées sur les serveurs du fournisseurs et les bibliothèques ne peuvent en disposer avec la même liberté, a fortiori quand les serveurs sont à l'étranger. Cela pose de nombreuses questions, allant de la plateformesisation⁴⁴ des services des bibliothèques, au statut juridique des données collectées par la bibliothèque, sur un serveur qui n'est pas le sien et parfois à l'étranger. Comment la Loi informatique et liberté s'applique aux données stockées sur serveur distant? Quelles garanties sont données quant aux conditions d'accès, de sécurité et d'utilisation des données dans le *cloud* etc. Les données collectées par la bibliothèque relèvent en effet de la loi Informatique et libertés de 1978 qui garantit aux usagers la transparence des informations disponibles, un droit de suppression et de modification et l'anonymat de leur utilisation ultérieure. Comment garantir ces principes quand les données sont stockées sur un serveur qui n'est pas celui de l'université, voire quand il n'est pas en France? Il faut prendre conscience que les données produites par la bibliothèque via son SIGB sont des données externes quand le système s'appuie sur des solutions dans le *cloud*.

Il y a donc là un double enjeu pour les bibliothèques et les universités :

- Pour assurer la maîtrise juridique des données collectées par les bibliothèques et par l'université, les solutions de type plateforme qui déportent le service et les données sur un serveur distant sont un écueil à éviter.
- Pour garantir la capacité à accéder à des données brutes afin de les croiser, et les modéliser pour améliorer la connaissance des usagers la maîtrise totale des données brutes est indispensable dans la durée.

Pour ces deux raisons, les systèmes *cloud* n'offrent pas toujours les garanties nécessaires.

⁴³ BNF, 2016. Les publics - Rapport d'activité 2015 de la BnF [en ligne]. Paris. [Consulté le 18 décembre 2016]. Disponible à l'adresse : <http://webapp.bnf.fr/rapport/publics.html> par le biais d'applications mobiles représentent 3,6 % du total des consultations (consulté le 12/12/2016)

⁴⁴ Voir sur le sujet des plateformes d'édition le billet de blog de Christophe Péralès, Future is now : « uberisation » de la science et économie de plateforme publiée le 2 octobre 2016. <http://blog.educpro.fr/christophe-perales/2016/10/02/future-is-now-uberisation-de-la-science-et-economie-de-plateforme/> (consulté le 12/12/2016)

1.5.2. Les données externes à la bibliothèque

Il existe des données collectées, ou gérées par des tiers pour la bibliothèque. Leur nature, disponibilité, récupération et statuts juridiques varient selon les services et les licences proposées par les prestataires. On trouve ces données chez :

- Des prestataires de service généraliste : (Xiti, Google Analytics, Google Drive, etc.) Les outils identifiés de cette nature peuvent être multiples et il ne s'agit pas ici de lister tous les services et solutions qui peuvent être sous-traités. Quelques services doivent cependant être mis en avant car ils sont au cœur du métier des bibliothèques et les données exploitées sont indispensables à la connaissance des publics. Il s'agit principalement des outils liés à la mesure de l'audience des sites web. Deux solutions émergent et sont signalées dans l'enquête menée pour ce mémoire. Il s'agit principalement de Google Analytics et de Xiti. Le fonctionnement de ces deux services est très différent.
- Véritable plateforme, Google Analytics permet sur un site dédié de suivre l'activité de son site, les traces de navigation etc. Les modalités de fonctionnement, les méthodes d'analyse, la restitution des données sont totalement opaques et contraintes. Le client de Google Analytics est totalement dépendant de la solution proposée et ne peut rien faire en dehors de celle-ci. Le terme de client est d'ailleurs exagéré puisque le service est gratuit. Ce qui est vendu ce sont les AdWords⁴⁵, et liens publicitaires, dont l'impact sera ensuite mesuré par Google Analytics. Dans cette configuration, la bibliothèque dispose d'un service assez qualitatif mais qui l'enferme dans un environnement, hors duquel rien ne peut être récupéré ni exploité. Les données de la fréquentation du site de la bibliothèque appartiennent, dans ce cadre, à Google et elles ne peuvent être exploitées hors de Google. Jusqu'en septembre dernier, les données n'étaient visibles que via les outils standards de Google Analytics. Cependant, Google a ouvert le 26 septembre un outil, DataStudio⁴⁶, permettant de visualiser des données venant de différentes sources dont Google Analytics, offrant un service plus souple, mais toujours dans le même écosystème.
- Xiti offre une logique de service assez différente. Par des lignes de code à poser dans les pages dont on veut connaître l'audience, Xiti est capable de capter les logs de connexion fournis par le serveur hébergeant le site (adresse IP, URL consultée, heure de consultation, durée de consultation) et de fournir un service de visualisation des statistiques en ligne gratuit. Une solution logicielle payante s'appuyant sur cette première technologie qui permet de disposer d'outils de paramétrage, de consolidation, de visualisation plus fins et détaillés. Les données restent à disposition des établissements.

⁴⁵ Google propose ici d'acheter des mots clés pour que le site apparaisse en lien publicitaire sur les pages de résultats utilisant le mot acheté à des fins publicitaires (Advertising Word : Adword)

⁴⁶ <https://datastudio.google.com>

- Des tiers de confiance : on peut en identifier un grand nombre dans le domaine de la publication scientifique et du référencement des publications (des gestionnaires des identifiants chercheurs comme idref⁴⁷ ou Orcid⁴⁸ aux entrepôts de publication eux-mêmes comme HAL ou arXiv⁴⁹, ou aux services « support » comme Huma-num). Qu'il s'agissent d'acteurs nationaux ou internationaux, publics ou para-publics l'enjeu serait plutôt de constater la multiplicité des acteurs et d'observer les stratégies employées pour gérer cette diversité.
- Les réseaux sociaux de recherche sont un autre canal, très utilisés, de référencement et de diffusion des publications. Rachetés, appuyés ou intégrés par des éditeurs scientifiques, ils sont un des dispositifs de plateformes des éditeurs scientifiques renforçant la dépendance des chercheurs à l'égard de ces intermédiaires⁵⁰. Très utilisés, ils sont donc un des nombreux silos de données, externes, que la bibliothèque doit suivre pour assurer ses missions.
- Des plateformes intermédiaires spécialisées : de nombreux services portés par des technologies dans le *cloud* sont disponibles pour les bibliothèques. Ils vont du cœur des collections (signalement avec les SIGB dans les nuages ou, acquisition et politique documentaire avec les plateformes de diffusion des publications électroniques) aux services proposés aux usagers comme la mesure d'audience ou la réservation de salle (Affluences).
- Les logiciels dédiés de lecture numérique :
 - Avec l'accès distant aux ressources électroniques les bibliothèques ont connu de nombreuses difficultés à mesurer l'audience. La logique de Plateforme s'est imposée à plein et a soumis les bibliothèques à la démultiplication de services distincts et payants (accès aux publications courantes, aux archives, interrogation plein texte etc.) Pour effectuer la mesure des audiences des publications en ligne un service payant est imposé sur chaque plateforme. Les tâches de suivi s'en trouvent démultipliées quand pour les ressources papier, une simple requête au SIGB donne le résultat. De plus les modalités de présentation, les modes de calcul diffèrent d'un prestataire à l'autre et ne sont pas toujours documentés. Ce sont ces difficultés qui ont incité Couperin à lancer ezPAARSE. Un semblable schéma se met en place avec la littérature générale. Elle concerne assez largement les collections de sciences sociales des BU. Certaines des solutions proposées dépendent de PNB (Prêt Numérique en Bibliothèque). Dilicom, acteur traditionnel de commande et gestion de stock pour les libraires gère les commandes des bibliothèques en livres numériques et assure ensuite le rôle de tiers de confiance entre les éditeurs et les bibliothèques, pour garantir le respect des contrats sur les usages autorisés sur les livres numériques (nombre de prêts, nombre de prêt simultanés/nombre de livres numériques achetés etc.) C'est une première

47 <https://www.idref.fr> (consulté le 12/12/2016)

48 <http://orcid.org/> (consulté le 12/12/2016)

49 <https://arxiv.org/> (consulté le 12/12/2016)

50 De nombreux ouvrages traitent de la tendance de l'économie numérique à favoriser les plateformes sous différents angles (disruption, désintermédiation ou plateformes). Je retiendrais RIFKIN, Jeremy et SAINT-UPÉRY, Marc, 2005. L'âge de l'accès. Paris : La Découverte.

plateforme. Dilicom récolte de nombreuses informations auprès des bibliothèques et les transmet aux éditeurs. Le cadre juridique étant soumis à de nombreuses négociations entre le MCC (Ministère de la Culture et de la Communication) et différents acteurs, la question de l'utilisation des données collectées auprès des bibliothèques est soumise à la négociation avec chaque éditeur. La connaissance de l'usage des collections numériques par la bibliothèque reste acquis à la plateforme entraînant les mêmes contraintes pour les établissements concernés. Il s'agit en quelque sorte d'une tentative de plateformes des audiences qui crée une dépendance des bibliothèques vis à vis des fournisseurs de collections numériques. Cette désintermédiation devrait être en partie jugulée grâce à ezPAARSE pour les BU.

- Une autre plateforme située plus en aval peut déjà s'observer. En effet, les solutions de lectures pour PNB passent par des logiciels de lecture numérique gérant les DRM. Les plus gros opérateurs du marché soutenus par un système de diffusion des eBooks requièrent des solutions de lecture telles que Kobo, Ibooks, Kindle. Ces logiciels collectent des informations que ni les éditeurs, ni les libraires, ni les bibliothèques ne connaissent vraiment et qui pourtant ne sont pas sans intérêt : les pratiques de lecture (temps de lecture, annotations, exhaustivité de la lecture etc.) Ces données seront dans les années qui viennent un élément important de connaissance des pratiques réelles de lecture numérique, objets de négociations et de transactions futures.

Les données externes sont parfois des données brutes disponibles en ligne (INSEE, data.gouv.fr, etc.), parfois des données fournies par des services partenaires par échange éventuellement via des applications (scolarité). Le plus souvent les données transitent par des plateformes de services de partenaires publics (CINES par exemple pour la conservation pérenne des ressources numériques, HAL pour les publications etc.), des sous-traitants (éditeurs électroniques comme CAIRN ou Reed-Elsevier). En s'intéressant aux données, et donc à un des fondements du numérique, on ne doit pas être surpris d'observer cette logique de plateformes. Elle pose cependant des questions de responsabilité et de stratégie aux établissements universitaires :

- Responsabilité sur les données transmises, et sur la façon dont elles sont exploitées par les plateformes, à quelles fins et dans quel cadre légal
- Stratégie quant aux choix qui sont faits de travailler soit avec des partenaires privés soit avec des partenaires publics selon les risques différents de « disruption » dans les missions et les priorités de l'établissement. Encore faut-il que les facteurs de sélection des partenaires prennent en compte ces risques identifiables de « disruption ».

Dans les différents entretiens que nous avons menés sur les données de la recherche, il a été constaté que le principal facteur incitant les institutions à travailler sur un entrepôt de données dans l'établissement réside dans le fait que les solutions existantes sont toutes basées à l'étranger et donc relèvent d'un régime

juridique mouvant par rapport aux obligations légales françaises (Loi informatique et liberté, Loi pour une République Numérique) et laisse l'université française relativement démunie quant aux conditions de réutilisation possible des données. Le constat est celui d'une certaine prise de conscience de l'enjeu stratégique du sujet et de la nécessité de s'en saisir. Les témoignages font apparaître assez clairement pour les données de la recherche le regret d'une absence de solution nationale de type HAL pour le signalement des publications. Certains établissements souhaitent s'orienter vers la mutualisation selon des modalités à définir (mutualisation des développements, ou de la plateforme) D'autres mieux dotés, tout en regrettant une faible visibilité de la politique nationale semblent s'engager à court terme dans la création d'un entrepôt de données de la recherche en lien avec leur archive institutionnelle existante. La recherche de synergie et de mutualisation reste une préoccupation dans les différents établissements rencontrés et ayant pris en charge le sujet.

A l'inverse, la rapide expansion d'« Affluences » témoigne d'un moindre intérêt des bibliothèques pour les données de fréquentation en temps réel. Affluences offre un service que les bibliothèques donnaient uniquement sur site, dans le meilleur des cas. Ces établissements ont permis à « Affluences » de collecter en temps réel des informations que seules les bibliothèques relevaient quotidiennement. Cette granularité pouvait paraître secondaire dans les indicateurs des bibliothèques. Les différentes statistiques nationales (ESGBU) ne demandant que des informations consolidées. Les bibliothèques en disposaient donc via les compteurs placés aux portiques de sécurité. Ces données n'existaient que pour un usage statistique.

On a substitué à un système de collecte publique via les portiques et un relevé quotidien et souvent manuel, un système extérieur, automatisé transmettant les informations en temps réel au service « Affluences ». Ce changement de granularité permet de fournir de nouveaux services aux lecteurs mais faute de redondance des données brutes dans la bibliothèque, on permet ainsi à la société de se constituer en plateforme exclusive de connaissance de la fréquentation des bibliothèques. Une stratégie de mise à disposition des données de compteurs en Open Data auraient permis d'offrir un service identique, suscité une émulation entre les sociétés, des initiatives libres sans pour autant créer de lien de dépendance.

Ainsi la position de la bibliothèque par rapport à ses données change. Elle implique des choix stratégiques à court et moyen terme, d'externalisation, d'internalisation ou de mutualisation et de plateformes. Ces choix diffèrent selon le caractère stratégique et prioritaire ou selon le rôle des acteurs privés dans l'écosystème concerné.

Cette stratégie pragmatique des bibliothèques est très différente de celle que rapporte la littérature sur le Big Data, où la collecte des données est systématique et extensive, parfois avant même de savoir ce que l'on peut en faire, ce qui évite le développement de plateformes et les facteurs de disruption. On mesure combien la priorisation et la sélection dans la collecte de données en interne sont facteur de risques, parmi lesquels on compte la disruption.

1.5.3. Disruption et position de la bibliothèque par rapport à ses données

Dominique Cardon s'interrogeant dans un essai⁵¹ sur le Big Data et les algorithmes comme porteur de vision de la société, propose quatre catégories d'algorithmes :

- Les données à côté du web qui mesurent l'audience, la popularité, la fréquentation du web (médiamétrie)
- Les données au-dessus du web qui hiérarchisent le web, distinguent les sites les plus pertinents par leur influence sur les autres publications du web (PageRank de Google)
- Les données dans le web qui mesurent la popularité qualifiée, le benchmark du web, via les réseaux sociaux. (Facebook, Twitter)
- Les données sous le web, les traces que l'on laisse par notre navigation, qui agglomérées sont porteuses de sens via des calculs aux ambitions prédictives. Ce sont souvent ces données et les technologies associées qui sont qualifiées de Big Data.

Ces technologies sont apparues au fur et à mesure du développement du web et de ses usages. Chaque étape a permis de constituer des modèles autour desquels des plateformes se sont structurées. Ces différences de modèles sont aussi constitutives des tensions qui existent dans le web et son organisation. Ils sont des paradigmes en concurrence, en compétition pour des raisons technologiques et économiques mais aussi sans doute politiques, dans ce sens qu'ils disent un certain rapport au monde, et à la liberté individuelle

Cette taxonomie des données au-dessus, à côté, dans et sous le web semble pouvoir être adaptée au contexte des bibliothèques universitaires ou des bibliothèques dans leur ensemble, puisqu'elles organisent les données et sont donc dans un rapport qualifiable avec ce qu'elles décrivent.

Sont-elles :

- au-dessus : hiérarchisation des ressources
- à côté : observation de l'audience ou de la fréquentation
- dans les données : popularité, folksonomie etc.
- sous les données : analyse globale de données agglomérées à des fins prédictives

On voit dorénavant se dessiner des points de convergence avec le modèle proposé par Dominique Cardon. On peut déjà observer que les bibliothèques travaillent au moins au-dessus des données et à côté, voire dans les données selon les contextes et les démarches. En fonction du contenu du silo considéré, la bibliothèque aura une position différente :

- le suivi des prêts, des fréquentations met la bibliothèque « à côté des données », en observatrice des pratiques et des usages, de l'audience de ses collections et de ses espaces, ou de ses services en ligne.
- le catalogue met la bibliothèque « au-dessus des données », puisqu'elle organise, hiérarchise ses ressources, choisit celles qu'il faut conserver, mettre en avant, celles qui obsolètes peuvent être pilonnées ou mises en magasin.

⁵¹ CARDON, Dominique, 2015. A quoi rêvent les algorithmes: nos vies à l'heure des Big Data. Paris : Seuil. La République des idées.

- les exemples de métadonnées sociales sont plus rares mais existent. Le recours de certains catalogues à l'indexation collaborative, au crowdsourcing peut permettre la collecte de métadonnées sociales et selon les cas d'associer l'intérêt pour des ouvrages, des collections à une communauté précise. L'indexation collaborative du catalogue Archives et manuscrits de la BnF peut illustrer cette orientation. Au niveau international, on compte de nombreux exemples⁵². Les pratiques collaboratives se sont également beaucoup développées, dans le sillage des Humanités Numériques, pour des projets nationaux (<http://www.bibliindex.info>) ou internationaux (<http://monasterium.net>), d'indexation, de catalogage, ou de transcription collaboratives de collections. Ces projets associent régulièrement des établissements patrimoniaux pour les collections qu'ils abritent, mais il est plus rare de trouver des bibliothèques universitaires associées à ces projets.
- La bibliothèque « sous les données », par l'analyse des traces est encore très marginale, y compris au niveau international. Cependant, on peut considérer quelques pratiques traditionnelles en bibliothèques relevant de cette démarche : les journées test pratiquées dans de nombreux établissements, consistant à relever tous les ouvrages consultés par les lecteurs pour connaître leurs usages relèvent clairement de l'analyse des traces. De même, la relève des compteurs en temps réel relève également de cette démarche. Le développement de l'internet des objets (IoT) et le lancement de bâtiments connectés (l'Illiad, dans une certaine mesure les bibliothèques de Jussieu PMC, et prochainement le Learning Center de Paris-Saclay) ouvrent la voie à des collectes de traces plus nombreuses par des portiques ou des détecteurs de mouvements : utilisation des salles de travail, utilisation des espaces en temps réel. Enfin, il faut noter le projet de recherche lancé par la BnF avec Mines-Telecom visant à analyser les traces de navigation des internautes sur Gallica à des fins de meilleure connaissance sociologique des Gallicanautes.

Les bibliothèques s'inscrivent donc bien dans cette taxonomie avec des pratiques plus ou moins répandues. Or, l'ambition de Dominique Cardon est également d'étudier les tensions entre ces modèles et entre les entreprises les ayant développées. Ces tensions relèvent d'une chronologie dans la mise en place des technologies, des stratégies et de la performance des modèles économiques afférents. Et si les tensions qu'il observe induite par la position des plateformes par rapport aux données étaient aussi les tensions qui animent les bibliothèques ? Elles prennent d'autres formes puisque l'ampleur et le contexte économique sont plus modestes ou placés au second plan.

L'intérêt de cette approche consiste à identifier un des ressorts des tensions observées dans le web, et à le transposer dans les bibliothèques à l'organisation des ressources documentaires et la connaissance de ses usagers.

Ces tensions peuvent également s'étudier à l'aune de l'analyse qu'avait faite Roger T Pédaque des 3 dimensions du document, explosé par le numérique

⁵² MOIREZ, Pauline, 2013. BIBLIOTHÈQUES, CROWDSOURCING, MÉTADONNÉES SOCIALES. [en ligne]. 1 janvier 2013. [Consulté le 18 décembre 2016]. Disponible à l'adresse : <http://bbf.enssib.fr/consulter/bbf-2013-05-0032-007>

comme l'a longuement étudié Jean-Michel Salaün dans « Le lu, le vu et le su »⁵³. Il pourrait en effet y avoir une corrélation entre une dimension du document et la position qu'on adopte pour interroger les données liées à cette dimension. En effet, quand on décrit « le vu », un ouvrage dans sa matérialité et son intégrité on est au dessus de l'information, on est dans la hiérarchisation de l'information selon le modèle de Dominique Cardon.

Quand on est dans « le lu », dans le texte on est dans le TextMining donc dans la démarche calculatoire et prédictive du modèle de Dominique Cardon.

Quand on est dans « le su », la popularité, la réputation d'un ouvrage, on est dans la médiation et donc dans les données sociales selon le modèle de Dominique Cardon.

Cette mise en perspective avec le système de Dominique Cardon peut être utile pour appréhender aussi la question des représentations professionnelles. Certaines dimensions du document constituent moins le cœur historique du métier des bibliothèques que d'autres qui peuvent favoriser des tensions selon le positionnement culturel des bibliothèques.

Cependant, une approche exclusivement structurale, passant par des grilles d'analyse peut donner certaines clés de compréhension mais ne peut suffire à appréhender seule la situation des bibliothèques vis à vis de l'explosion des données. C'est l'objet du travail mené à partir des données de l'enquête que de sortir d'un cadre théorique et de tenter de donner corps à ce modèle. L'enquête quantitative et qualitative menée auprès de différents établissements a permis de recueillir des données objectives et subjectives sur la situation des bibliothèques universitaires. Les dix entretiens oraux réalisés auprès de SCD d'université, et les 28 questionnaires renseignés par différents établissements de l'enseignement supérieur (bibliothèques de grandes institutions et SCD) doivent donner un échantillon significatif qu'il s'agit désormais d'étudier.

53 SALAÛN, Jean-Michel, 2012. Vu, lu, su. Paris : La Découverte

ORGANISATIONS ET DONNEES

Emmanuelle Bermès souligne⁵⁴ que si le Big Data a commencé à faire évoluer les processus de travail en bibliothèque, les marges de progrès sont encore très importantes. Elle y décrit la convergence entre les technologies du Big Data et celles du Linked Data, notamment par l'automatisation d'alignements d'entités nommées de data.bnf.fr avec les notices correspondantes du catalogue général de la BnF. 100 000 liens entre manifestations et œuvres ont été injectés dans le catalogue sans aucune intervention humaine. Ceci témoigne d'une double « révolution copernicienne » pour les bibliothèques :

- Ce ne sont plus les documents qui sont au centre du catalogue mais les entités, les concepts. Ces nouveaux catalogues ne s'intéressent plus aux documents mais aux liens qu'il faut construire entre les manifestations et les entités. La chaîne de traitement du document structurant le travail, les organisations mais aussi les cadres d'emplois pourraient ainsi se voir profondément chamboulés si ce modèle de catalogue s'imposait. Cependant, il faut souligner que si le modèle FRBR n'est plus un modèle théorique, et façonne largement les stratégies et les outils des bibliothèques, il s'établit en parallèle des modes traditionnels de catalogage. Ainsi, le « modèle sémantique » associé aux technologies automatisées du Big Data, coexiste-t-il avec le « modèle document » associé au SIGB et au traitement du document dans la chaîne traditionnelle.
- Le second volet de cette « révolution copernicienne » concerne les modes d'organisation. Cette coexistence a entraîné la dissociation de la production des données d'une part de l'organisation des données d'autre part. On a ainsi un catalogage traditionnel décrivant des documents, d'une part et des équipes assurant l'organisation des accès par les entités nommées, et les technologies du Big Data d'autre part. On retrouve avec Isidore la même ligne de séparation : d'une part une production traditionnelle de signalement de collections par les institutions disposant des collections et d'autre part le traitement algorithmique des données pour une exposition en RDF dans un silo unique et qui permet d'être indexé sur le web par les moteurs. Cette coexistence de modèles implique la coexistence de différentes modalités de production dans une ou plusieurs structures.

Pour être efficiente cette coexistence doit être harmonieuse et cohérente. Sur les questions métier c'est l'objet de la transition bibliographique que d'assurer cette cohérence. Sur le plan organisationnel de nouveaux outils managériaux doivent être mis en œuvre pour assurer la cohésion de ces ensembles marchant visiblement à des rythmes différents et pouvant être vécus comme allant dans des directions différentes⁵⁵.

⁵⁴ BERMÈS, Emmanuelle (éd.), 2017. L'avenir des catalogues. Le cercle de la librairie, à paraître. Les éléments ont été communiqués avec la complicité bienveillante de l'auteure. Qu'elle en soit remerciée.

⁵⁵ ILLIEN Gildas « Une BnF agile, quand le développement logiciel fait bouger l'organisation du travail » in PÉRALES, Christophe, 2015. Conduire le changement en bibliothèque : vers des organisations apprenantes. Villeurbanne : Enssib.

Les bibliothèques ne sont donc pas au bout de cette révolution mais à la croisée des chemins où coexistent ici ou là, non seulement des modèles différents de gestion des données, de procédure de production, mais aussi d'encadrement, de management des équipes.

On voit ici l'intérêt de la notion de tension entre les modèles dans la taxonomie de D. Cardon. Dans le cas d'une organisation du travail autour du document, où la bibliothèque hiérarchise et prend de la hauteur sur les données, le travail est inscrit dans une chaîne documentaire qui structure l'organigramme, les cadres d'emploi et les fiches de poste. La chaîne de décision et d'organisation hiérarchique, en râteau, y a fait ses preuves depuis longtemps. Dans le cadre d'un travail basé sur l'automatisation des traitements sur les données, dans un niveau plus fin de granularité où convergent les compétences métiers et les compétences informatiques, la souplesse d'organisation et l'agilité de décision sont nécessaires. Cela s'articule autour des différents silos de données et unités organisationnelles de l'institution. Comment ces différents modèles d'organisation hiérarchique dans un cas et transversale dans l'autre peuvent-ils coexister sans tensions dans la même organisation ?

Il s'agit maintenant par l'analyse des questionnaires et entretiens d'identifier ce qui émerge, ce qui s'installe dans les pratiques ou les organisations.

2.1. LA PLACE DES DONNEES DANS L'ORGANISATION

2.1.1. Outils, données et maturation des processus de travail

Parmi les 19 établissements qui collectent des données automatiquement ou prévoient de le faire, 16 ont précisé les types de collecte. Tous travaillent d'une manière ou d'une autre sur la mesure de l'audience de leurs services :

- au moins 13 collectent les données depuis leur proxy d'accès aux ressources électroniques. La plupart précise qu'ils s'appuient sur les outils mis en place par Couperin ezPAARSE et EZproxy pour mieux connaître les usages de cette documentation.
- 4 établissements travaillent également sur la fréquentation de leurs services en ligne en s'appuyant sur les logs de connexion (ce qui nécessite ensuite un traitement expert par des outils dédiés) ou sur des services en ligne dédiés comme Google Analytics.
- 4 récupèrent également les données de fréquentation de leurs établissements soit en passant par les portiques de sécurité qui alimentent une base, soit par les services d'Affluences.

On voit ici se dégager un besoin unique, mesurer les usages des différents services proposés par la bibliothèque. Ils s'appuient sur des outils et des silos de données distincts selon les procédés de collecte, la nature des données d'origine, ou les services qui les exploitent. L'ensemble le plus large reste constitué par les

données concernant la documentation électronique et le service de Couperin, ezPAARSE, service mutualisé servant avantageusement la communauté desservie.

D'un autre coté, de nombreux établissements ont affirmé leur volonté de se doter d'une archive institutionnelle propre, malgré l'existence de HAL. Il faut souligner que tant dans les entretiens que dans les questionnaires, il apparaît que l'initiative ne vient pas forcément du SCD mais de l'université de tutelle elle-même.

Il y a comme un paradoxe à constater qu'une plateforme mutualisée telle que HAL ne retienne pas les initiatives locales de création de système propre à chaque université. Mais ce paradoxe n'est peut être qu'apparent.

En effet, les entretiens soulignent régulièrement les insuffisances et les limites de HAL en regard des besoins exprimés par les tutelles des SCD. Les outils que proposent HAL apparaissent insuffisants dans la gestion du signalement des ressources, mais surtout des auteurs, des laboratoires et des institutions dont la gestion s'avèrent longue et laborieuse. Par l'action des participants dans les laboratoires, les doublons de descripteurs se multiplient les rendant d'autant moins pertinents. Leur reprise ne corrige que momentanément le problème jusqu'à la prochaine création intempestive de doublons. Du point de vue des universités, valoriser leurs publications, quand elles sont si difficiles à identifier faute de normalisation, peut relever de la gageure. Face à la montée en puissance des critères d'évaluation des universités sur des grilles inspirées du classement de Shanghai, ce handicap pèse lourd. Si les universités sont tentées de se doter de leur propre système c'est que HAL laissent les modalités de signalement sous la responsabilité des déclarants et ne prévoient pas de les modifier. Mais malgré tout HAL a permis de démontrer l'utilité et la viabilité d'un système d'archive dans l'évaluation et la valorisation de la recherche. Les efforts nécessaires pour compter les publications dans HAL impliquent une organisation dans l'université concernée. Face à l'urgence des questions d'évaluation de la recherche, la création d'une archive institutionnelle est privilégiée car cela apparaît plus simple que de s'organiser au sein d'une archive mutualisée, au détriment de la qualité et de la gestion des données.

HAL semble avoir permis de fédérer une communauté d'usage, qui a pu partager et mûrir autour de cet outil des compétences, des organisations et des processus de travail. Cette maturité avançant, et HAL demeurant en deçà des besoins ou des capacités des déclarants de s'organiser, le recours à une solution locale s'affirme dans de nombreux établissements. Elle s'affirme d'autant plus que les systèmes de référencement croisé sont possibles, via le protocole OAI-PMH entre HAL et l'entrepôt institutionnel. Ainsi le projet de la BU de Paris-Sud est bien de commencer dans HAL un référencement des publications de l'université puis dans un second temps de lancer son propre entrepôt qui moissonnera les notices OAI HAL de Paris-Sud.

Un mouvement semblable autour des données de la recherche se dessine en creux dans l'articulation de celles-ci avec HAL ou l'archive institutionnelle. Plusieurs entretiens convergent sur ce point :

- le regret de ne pas voir émerger une plateforme nationale ou un cadre national de travail pour les données de la recherche
- le besoin de mutualiser la solution technique ou l'infrastructure de stockage afin de partager les coûts de développement, de croiser les pratiques, pour mieux appréhender l'ampleur de ce vaste ensemble que sont les données de la recherche. Dans un contexte universitaire de fusion et de fédération d'universités, la recherche de partenariats entre institutions se fait cependant le plus souvent hors des cadres de la COMUE. C'est que le système d'archive est déjà pensé à l'échelle de la COMUE.
- HAL est cité comme une référence soit parce que les données de la recherche sont à associer à des publications électroniques, soit parce que HAL est un modèle centralisé national d'appui à des services universitaires permettant l'émergence d'une communauté d'usage d'autant plus mature qu'elle est installée dans de nombreuses institutions, et d'autant plus nécessaire que les acteurs sont peu nombreux au sein de chaque institution. Les entretiens témoignent d'un besoin de partage sur ce modèle pour travailler sur la question des données de la recherche.

D'une certaine manière, on peut penser que ces communautés d'usage naissent autant du besoin fonctionnel pour lequel elles ont été créées (Couperin) que du besoin de sortir de la structure pour se décontextualiser et faire avancer hors structure des sujets sensibles. Une fois mûris et légitimés dans ces espaces de coopération ils peuvent être repris dans les structures concernées.

Il ne faut pas considérer le recours à des solutions locales d'entrepôts institutionnels comme un échec de HAL, mais comme une réussite du modèle, une maturité des pratiques, malgré un manque d'agilité et de souplesse du modèle centralisé. Une articulation entre les entrepôts institutionnels et un référencement dans HAL pourrait devenir à terme un modèle majoritaire. En effet, 11 établissements sur 28 interrogés disposent déjà d'une archive institutionnelle et 5 prévoient de s'en doter.

L'exemple d'ezPAARSE va dans le même sens. Les chiffres de l'échantillon témoignent d'une large adhésion au projet. Dix établissements le mentionnent dans leur mode de collecte automatique de données. EzPAARSE répond à des besoins et à des problématiques très répandus dans les SCD : manque d'outillage adapté, besoin de travailler en communauté pour apprendre les uns des autres sur des sujets à la fois technique et métier. Porté par le consortium Couperin, le cadre communautaire existait déjà et a permis l'émergence du projet et sa réalisation puis désormais sa dynamique. L'organisation de Couperin permet la prise en charge par la communauté d'utilisateurs des décisions et des priorités de développement, mais les choix se sont portés sur des outils libres dont les différentes communautés au-delà de Couperin peuvent se saisir.

Si ezPAARSE s'appuie sur une communauté ancienne, les besoins n'ont pas l'antériorité de HAL. Ce dernier lancé en 2001 bénéficie de 15 années d'expérience et de maturation. ezPAARSE n'a encore que peu de recul. Ses

premiers travaux commencent en décembre 2012 et la version 1 a été lancée en septembre 2013. Des besoins moins matures peuvent plus facilement se mutualiser et converger.

Les autres pratiques de récupération de données automatisées ne sont pas associées à des outils spécifiques (compteurs d'entrée par exemple) ou fonctionnent avec des outils du marché (Affluences ou Google Analytics pour l'audience des sites). Dans ce cadre on ne trouve pas de communauté d'usage en bibliothèques, mais une relation client/prestataire. On peut poser l'hypothèse que si Affluences continue à être utilisé par un large panel de bibliothèques en France, une coordination sera assurée pour faire émerger les besoins communs. Deux modèles de coopération sont possibles, l'un piloté par Affluences et l'autre organisé par la profession elle-même. Un tel schéma n'est pas envisageable pour Google Analytics.

A dessein, l'enquête a séparé :

- la récupération des données automatiques
- la récupération d'informations externes à la bibliothèque auprès de partenaires
 - la question de la visualisation des données
 - la question de l'informatique décisionnelle

Si ces différents sujets peuvent être liés, ils ne le sont pas forcément ou ne sont pas forcément perçus comme tels. Les résultats montrent que ces sujets sont rarement envisagés sous l'ensemble de ces aspects.

Par ailleurs, l'enquête fait une distinction entre les données collectées à des fins statistiques, de visualisation de données ou d'informatique décisionnelle. Or, il n'est pas possible de travailler sur des solutions d'informatique décisionnelle sans collecte de données en amont. C'est pourtant ce que donnent comme information deux des établissements du panel. On peut en déduire qu'il y a une collecte de données mais dont la nature ne peut pas être connue.

De même, on peut envisager que certains établissements déclarant récolter des données, travaillent en fait sur des solutions d'informatique décisionnelle ou de visualisation de données sans savoir nommer ces outils. D'ailleurs certains outils sont à la frontière entre l'informatique décisionnelle et la data visualisation. Ainsi le logiciel Tableau Public, utilisé à la bibliothèque de l'université de Yale, doit permettre d'accompagner les acquéreurs en leur fournissant des données visuelles sur leurs champs disciplinaires (emprunts, budget etc.). Il s'agit donc bien dans ce contexte d'un outil d'aide à la décision appuyé par une représentation graphique des données.

Ces domaines sont moins matures notamment du fait de leur grande évolution ces dernières années. Le vocabulaire n'est pas fixé et contribue à un certain flou qui explique pour une part certaines incohérences des réponses du questionnaire.

A ce stade on peut constater quelques écarts de maîtrise et d'appréhension des données selon les sujets, et leur maturité dans les organisations. On peut s'interroger sur la compréhension de ces sujets émergents par les professionnels.

L'enquête permet de dessiner un paysage contrasté :

- Des champs très structurés et investis par la profession montrent une maîtrise des enjeux et des processus, sinon dans les SCD au moins dans les organisations professionnelles. (HAL, données de la recherche, mesures des audiences)
- Un champ difficilement mesurable par l'enquête quantitative, car peu structuré contenant des données variées, composites et qui se traduit par des initiatives diverses au vocabulaire imprécis. Cela concerne des activités liées aux données collectées par le SCD dans un contexte local qu'une enquête globale ne peut saisir. Ces champs seront davantage explorés au travers des entretiens individuels.

2.1.2. Maîtrise des données, maîtrise des organisations

L'émergence de nouveaux types de données, de nouveau silos de données, de nouveaux formats de données pose au-delà des compétences métiers à développer, des questions liées à l'intégration dans l'organisation de l'établissement. Ces questions sont d'autant plus tangibles que le sujet concerné est méconnu. Différents mécanismes ont pu être observés qui œuvrent à l'intégration de la nouveauté en gérant les risques organisationnels mais aussi juridiques etc.

A propos des données de la recherche, plusieurs entretiens soulignent les difficultés rencontrées pour faire comprendre les enjeux et les questions liées à la création d'un nouvel entrepôt de données. Les usages, les fonctions, le cadre juridique de ces entrepôts sont méconnus, incertains et donc forcément mouvants. Cette incertitude encouragerait les responsables à différer la décision pour en attendre la maturation. Or, l'urgence de certains dossiers comme l'est tout particulièrement celui des données de la recherche n'autorise pas forcément le luxe d'attendre. Selon les établissements on a observé des choix différents sur ce sujet qui disent aussi la difficulté à communiquer sur ce type d'obstacles.

« Ca fait 3 ans au moins que je dis partout à l'université : attention, il faut penser nos données, on va devoir les gérer proprement dans pas longtemps ... on a du mal à penser la gestion de la donnée dans l'institution parce qu'on ne la voit pas, on ne sait pas ce que c'est, d'une part ; et par ailleurs, très souvent, c'est le pré carré des DSI qui en fait sont très *old school*, ne comprennent pas les problèmes data ».

Dans ce cas précis, la nouveauté a été identifiée et n'a pas été traitée. Elle semble agiter trop de questions organisationnelles, et n'offrir que peu de garantie dans sa mise en œuvre du fait de questions juridiques qui n'avaient pas de réponses - du moins avant la loi sur la République Numérique - et de questions techniques.

Pour les établissements qui ont entrepris de se lancer sur de tels sujets, les entretiens remontent des difficultés, soit dans les relations avec les équipes informatiques qui se pensent responsables du sujet, soit avec les équipes métiers (hiérarchiques ou non) qui pensent que ça n'est pas de la responsabilité de la bibliothèque. Le questionnaire offre des résultats également assez similaires sur ce sujet : la plupart des établissements concernés par un projet sur les données de la recherche en sont soit les porteurs et commanditaires (ils précisent alors qu'ils travaillent à convaincre l'université et les chercheurs du bienfondé du projet) soit

sont consultés par des équipes de recherche pour des besoins ponctuels et spécifiques. Le questionnaire n'a pas remonté de cas où l'université porte le projet et fait appel aux compétences de la bibliothèque pour l'accompagner. Le modèle anglo-saxon n'a trouvé que de rares transpositions en France⁵⁶ (Sciences-Po et Lille1).

Au cœur de ces difficultés on trouve des questions d'appréhension, de responsabilité et de légitimité. Tout le monde n'a pas

« les compétences sur les formats et les métadonnées nécessaires pour prendre de la distance vis à vis de l'outil et voir que la fonction ne change pas. » « On a donc un truc invisible, géré par des spécialistes (informaticiens) qui ne comprennent pas les enjeux... »

Certains des établissements qui ont pu avancer sur ces sujets ont une approche par les processus. Cela a un double avantage. Les processus permettent de mieux lire la dimension métier sous-jacente. La maîtrise des processus garantit la qualité des données. Ainsi les réticences opposées sur l'appréhension ou la responsabilité peuvent être contournées par cette approche processus, et celle de la légitimité par la maîtrise qualité qui en découle.

Si la démarche qualité s'est développée dans les bibliothèques ces dernières années, c'est le plus souvent au travers de la qualité de service, de la charte Marianne. La certification ISO 9001 attribuée à certains établissements concerne exclusivement des missions dans la relation au public.

Dans un des établissements rencontrés, il existe auprès de la direction une mission qualité et une mission indicateur. La mission qualité travaille sur la qualité des services aux publics, la mission indicateur alimente en indicateurs l'ensemble des services, y compris la mission qualité qui s'appuie dessus pour améliorer les services rendus. La qualité des données relève de la culture d'établissement en lien avec la collecte de données pour fournir des indicateurs :

« Il était de plus en plus difficile pour la direction de maîtriser toutes les données qui permettent de diriger une bibliothèque. (...) c'est un outil dont on a besoin tous les jours aujourd'hui. »

Mais cette maîtrise va dans les deux sens : il s'agit de savoir ce qu'il se passe dans l'établissement mais aussi de maîtriser ce qui sort comme information vers les services de direction de l'université.

« Je veux garder la maîtrise de ce qui sort comme information... Je ne veux pas que l'université s'empare des données brutes que la bibliothèque produit. En revanche je mets un point d'honneur à ne donner que des données fiables même quand elles nous sont défavorables».

Pour cela, les silos de données quels qu'ils soient doivent être fortement intégrés au système et à l'organisation de l'établissement. L'informatique décisionnelle est ici un outil de pilotage, un levier par lequel la maîtrise de l'organisation se mesure et s'opère.

⁵⁶ GAILLARD, Rémi, 2014. De l'Open data à l'Open research data : quelle(s) politique(s) pour les données de la recherche. 2014. Enssib.

Pour autant on a constaté dans différents contextes l'utilisation des outils qualité pour maîtriser les données, les appréhender, les consolider. Cette démarche a été observée en plusieurs occasions lors des entretiens.

Des processus définis qui permettent de mieux appréhender la dimension métier mais aussi de garantir que les tâches sont effectuées de la même façon. Ils permettent donc de mesurer des résultats comparables. Ces derniers sont nécessaires à l'évaluation de la qualité en fin de processus. Ils concernent le traitement des données de la recherche sur des corpus ciblés à Sciences Po mais aussi la collecte des données de pilotage de la bibliothèque universitaire de Grenoble pour le système d'informatique décisionnelle. Dans le premier cas la comparaison permet la vérification de la qualité, dans le second cela permet la cohérence des indicateurs et donc une lecture sur la durée et leur évolution. Cependant dans le premier cas, seule une partie de l'activité de la bibliothèque est concernée, le traitement de données collectées par un laboratoire de Sciences Po. A Grenoble l'ensemble des processus métier est concerné. Les impacts organisationnels ne sont pas les mêmes.

« Le but du jeux était que ça ne rajoute pas de travail aux équipes, au contraire ça va vous faciliter le travail ».

Ici les processus de collecte sont totalement automatisés mais les processus métiers en amont, pour garantir la qualité des données et leur cohérence ont fait l'objet d'une formalisation avancée. Ce sont donc concrètement l'ensemble des services qui ont été impliqués par le travail sur les processus nécessaires à la collecte automatique et cohérente des données. Une telle démarche ne peut se faire sans l'appui constant de la direction de l'établissement.

« Il faut une volonté forte de la direction, c'est un outil de direction, c'est un outil de pilotage. Il faut que la direction s'implique. En revanche j'ai la conviction que toutes les bibliothèques ont l'usage de cet outil. Je suis même étonné que toutes les bibliothèques ne soient pas équipées de tels outils. »

Ainsi, si de nouveaux silos de données apparaissent pour des besoins locaux sur les processus, un travail isolé peut se réaliser, parfois d'envergure mais sans conséquences sur les autres services. L'implication de l'établissement est secondaire. Les données et leur usage restent limités au contexte initial et ne sont pas considérés par les autres services puisque le plus souvent ignorés. Le risque de perte de compétence est ici très important en cas de départ.

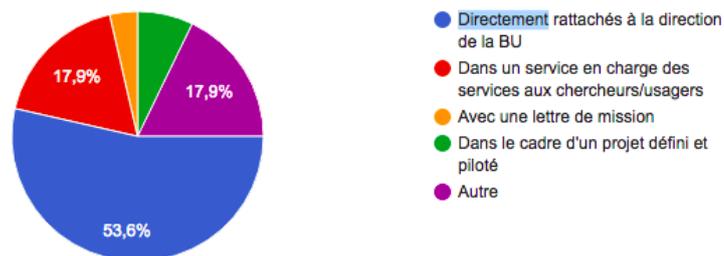
A l'inverse quand il est souhaité par la direction que les données soient croisées, consolidées et issues de processus métiers qui en garantissent la qualité et la cohérence, l'ensemble des services est concerné. Cela implique une plus grande prise de conscience par l'ensemble des acteurs, hiérarchiques et opérationnels. Cette gestion des données et de la qualité par les processus est sinon un mode de management, au moins un outil au cœur du management. Elle ne peut se mettre en place que par une décision managériale et donc hiérarchique.

« La qualité des données est une culture d'établissement, nous en parlons à chaque formation. C'est un investissement constant. Nous formons inlassablement à la qualité des données produites. »

Sur 28 établissements d'enseignement supérieur, 5 déclarent utiliser un outil d'informatique décisionnelle pour produire les indicateurs, soit un peu moins de 20% de l'échantillon. Les éléments recueillis notamment lors des entretiens témoignent de différents facteurs quant au développement de ces outils. Il s'agit pour l'essentiel de répondre au besoin du directeur du SCD d'un tableau de bord régulier. L'automatisation que permettent les outils d'informatique décisionnelle constitue un atout.

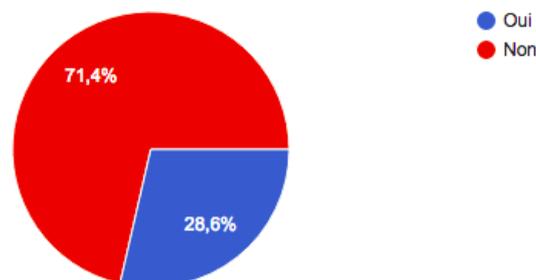
On peut observer plus largement, dans le questionnaire au delà de la seule question de l'informatique décisionnelle la volonté de maîtrise par la direction. On constate que dans les différentes structures consultées les établissements traitent dans un premier temps les nouveaux silos de données à côté des autres services de la bibliothèque. Ils sont dans la majorité des cas rattachés directement à la direction (53,6% soit 15 des 28 établissements consultés). Dans quelques cas ils sont positionnés dans un service ayant en charge l'appui à la recherche (5 établissements travaillant sur les données de la recherche ou des archives institutionnelles) ou dans celui de l'informatique documentaire (5 établissements). Deux d'entre eux ont fait le choix d'un mode projet et un celui d'une lettre de mission. Enfin, 2 établissements n'ont pas de personnels dédiés à ces missions.

Dans quels cadres ces personnels travaillent-ils ? (28 réponses)



Est-il envisagé au sein de votre établissement de modifier l'organigramme afin d'améliorer la gestion des données ?

(28 réponses)



L'émergence de ces sujets se passe donc le plus souvent dans le giron de la direction du SCD. On mesure par cette proportion l'importance de l'appui de la direction. Les projets concernent tous les types de données évoqués dans le questionnaire, même quand ils sont attachés à des sujets légitimes et traditionnels

comme le FRBR. C'est le caractère nouveau, innovant, qui explique le plus souvent ce rattachement à la direction qui peut suivre et mesurer l'avancement du projet. Il serait naturel que lorsque le projet est assez mature, il se détache de la direction et rejoigne les services. Cependant on constate que la modification de l'organigramme n'est envisagée que par 8 établissements sur la totalité de l'échantillon. Parmi ces 8 établissements, pour 3 d'entre eux les fonctions concernées sont rattachées à la direction, pour 4 d'entre eux à un service d'appui à la recherche, pour le dernier cas à un service documentaire. 12 établissements sur 15 dont les projets sont directement rattachés à la direction n'envisagent donc pas de modification de l'organigramme. On peut attribuer cela au fait que les projets sont encore peu matures et nécessitent d'être plus avancés. C'est par exemple le cas des établissements travaillant sur les données de la recherche où la phase d'étude est encore en cours. Pour les 5 établissements travaillant sur l'informatique décisionnelle, on peut comprendre que l'usage de ces outils est à destination de la direction et que donc c'est sa place naturelle. Pour d'autres travaux, les raisons ne semblent pas évidentes et sont sans doute liées à des situations locales.

Les organigrammes consultés⁵⁷ ne montrent pas de façon claire où sont les responsabilités sur les données. Les services d'appui à la recherche peuvent avoir à travailler sur HAL ou les archives institutionnelles, mais dans d'autres établissements ce sera le service d'informatique documentaire (Lyon 1)⁵⁸ qui en a la responsabilité. La question dans un cas est envisagée comme un service donné à un public spécifique, dans le second cas, le sujet est envisagé au travers de l'accès aux ressources. Ailleurs, il est envisagé au travers des processus de travail et du risque de multiplication de ces processus.

« Autour du catalogue où on faisait du travail de dentelière pour le SUDOC, convergent d'autres acteurs, des collègues du département de l'informatique documentaire, des collègues qui travaillent sur la documentation électronique... »

L'intégration, la maîtrise des processus qui ailleurs se multiplient passent ici par le catalogue et la chaîne de traitement du document.

« Il faut qu'on ait les bons intervenants et qu'on évite les chaînes parallèles. On n'aborde pas le problème de la même façon. On cherche à éviter que les collègues fassent deux métiers différents selon qu'ils traitent de la « docélec » ou des thèses. »

Dans un autre établissement rencontré en entretien, on observe que l'intégration de nouveaux silos de données liés à de nouvelles activités, s'effectue par un équilibre entre la stabilité des organisations et la distribution des moyens humains selon les priorités et les besoins.

« L'esprit de l'organisation ne change pas, c'est la concrétisation de l'organigramme qui change » ...

... pour intégrer une GED (Gestion électronique des documents) sur le dépôt des mémoires de M2 et de thèses, pour travailler sur SGBM etc. L'organigramme n'est cependant pas figé. Pour des besoins plus pérennes on peut le modifier par

⁵⁷ Voir quelques exemples en annexes, ainsi que MARCEROU-RAMEL, Nathalie, « Nouvelles compétences, nouvelles organisations » in CAVALIER, François et POULAIN, Martine, 2015. Bibliothèques universitaires: nouveaux horizons. Paris : Éditions du Cercle de la librairie. Bibliothèques

⁵⁸ <http://portaildoc.univ-lyon1.fr/medias/fichier/departement-sid-septembre-20161472807230841-pdf> (consulté le 1 décembre 2016)

« petites touches »; c'est ainsi que fut créée la mission d'appui à la recherche. L'ensemble est en fait géré par une démarche qualité qui a diffusé une culture de l'adaptation des moyens humains, des priorités aux besoins utilisateurs. Deux pôles de stabilité demeurent, un organigramme à « haute échelle » qui ne change pas et une chaîne du traitement du document qui structure encore largement la direction des collections de cet établissement et tout le service public. Cette organisation autour de la chaîne du document est maintenue car

« elle rend toujours des services attendus sur les documents, des prêts de livre papier... ».

On mobilise des moyens humains dans les différents services existants pour les placer sur les projets en dehors de la chaîne de traitement du document. L'allocation de moyens humains est donc décidée sur des projets dont la direction est assurée par des chefs de département ou leurs adjoints, mais pas forcément sur leurs effectifs. Tout cela est décidé

« après échange avec la hiérarchie fonctionnelle et hiérarchique en comité de direction ».

Le mode projet prend bien des moyens aux activités de production mais pour des objectifs et des priorités vérifiés par la démarche qualité pour une période limitée. Le projet terminé, les personnes sont remises à 100% sur leur fonction d'origine ou sur un autre projet.

Il a pu être établi qu'une certaine surqualité⁵⁹ était observée dans l'activité de catalogage. La qualité des notices créées n'est pas pleinement utilisée par les publics. Renoncer à cette sur-qualité permet de dégager des moyens humains sur les nouvelles activités permanentes. Ainsi, malgré le sentiment d'une certaine stabilité, l'activité de la bibliothèque change mais sans rupture, par « petites touches », moins consommatrice d'énergie qu'un grand changement d'organigramme. Cependant le changement est réel et doit être accompagné. Un plan d'accompagnement au changement était ainsi prévu pour l'ensemble du personnel avec différents parcours selon les situations individuelles, nous rapporte l'entretien. Ainsi, la démarche qualité peut faire l'économie de refontes lourdes des organigrammes mais pas de la prise en charge des changements qui irriguent la vie des services.

Ce type d'organisation a donc pris appui sur la démarche qualité et la culture métier qu'elle permet, pour faire évoluer les activités peu à peu. L'organisation offre une certaine souplesse, une agilité, encadrée et appuyée par la hiérarchie. Pour autant, elle permet de sortir momentanément des organisations hiérarchiques les personnels affectés à de nouveaux besoins. Dans ce cas précis, cela ne semble pas poser de difficultés particulières de territoire, de concurrence entre les directions.

On a ainsi pu observer plusieurs approches pour gérer la multiplication des silos de données et processus de travail associés. Une approche passant :

- par la démarche qualité,

⁵⁹ La sur-qualité est une qualité au-dessus du niveau attendu entraînant un surcoût https://fr.wikipedia.org/wiki/Management_de_la_qualit%C3%A9 (consulté le 12/12/2016)

- par le pilotage via des indicateurs,
- par l'ouverture du catalogue à différents types de signalements accompagnée d'une réflexion puis d'une refonte de l'organigramme du SCD.

Ces approches ne sont pas exclusives les unes des autres et pourraient s'observer dans le même établissement.

- Par la démarche qualité on observe la maîtrise des processus de travail et de la qualité des données produites.
- Par l'informatique décisionnelle on observe une recherche de la maîtrise du travail et de son efficacité.
- Par la réflexion sur l'organigramme il s'agit de repenser les moyens et les métiers par rapport aux priorités et aux missions.

Chaque démarche commence depuis un point de départ différent mais recherche dans chaque cas la maîtrise des processus métiers et des organisations. La production d'indicateurs est le passage obligé. Ces indicateurs procèdent d'une maîtrise des données,

- par la maîtrise des nouveaux processus (démarche qualité),
- par la mesure des processus de travail (informatique décisionnelle).
- par l'intégration aux processus établis (intégration au catalogue)

Quand un directeur de SCD insiste en entretien sur le passage du modèle des bibliothèques d'un « just in case » au « just in time », on peut sans doute considérer que les modèles ici observés convergent vers ce « just in time » par des moyens différents. Il s'agit de fournir le service attendu ni plus ni moins. Tous cherchent par la maîtrise des données et des processus à garantir le bon niveau de qualité, et dans certains domaines comme le catalogage à éviter la sur-qualité. Il s'agit d'optimiser les moyens humains alloués en fonction des priorités pour faire face à l'élargissement des missions, à la multiplication des silos de données avec des moyens contraints. Pour cela on s'appuie sur les outils dont on dispose :

- La démarche qualité n'empêche pas un accompagnement régulier des équipes car une telle logique impose le changement dans les dynamiques de travail.
- Les indicateurs permettent de mesurer les marges de manœuvre dans l'organisation et de dialoguer avec l'université pour les ajustements possibles.
- L'intégration au processus de travail traditionnel dans la chaîne de traitement documentaire.

Quand les établissements se saisissent pleinement de l'élargissement des processus de travail et des silos de données, en s'appuyant selon les contextes sur une culture métier comme levier pour le changement, le défi semble totalement relevé. Il passe par une maîtrise de l'encadrement des nouveaux processus et une maîtrise de l'intégration à l'organisation de travail. L'intégration, la traduction dans l'organigramme n'apparaissent pas toujours et pas dès le lancement du projet. L'agilité de ces dossiers ne trouve pas aisément de traduction dans une organisation figée. Le défi relevé par ces établissements est finalement de trouver des leviers pour conserver une organisation cohérente et dans laquelle chacun connaît sa place. Ces leviers permettent l'adaptation et le déplacement des compétences et des priorités.

Ces modalités peuvent être perçues comme peu compatibles avec un cadre administratif et une chaîne de responsabilité et de décision hérités de Napoléon et donc d'un modèle militaire très vertical. Là, la fiabilité repose largement sur la stabilité et où le changement est considéré comme un perturbateur de l'organisation.

2.1.3. La place du SCD et des données

La question des données dans les organisations agite largement la littérature informatique et professionnelle⁶⁰, mais aussi la sphère politique⁶¹.



Une partie de la question s'est structurée pendant les débats et la préparation de la Loi pour une République numérique qui consacre largement l'ouverture des données produites par les administrations et donc par les bibliothèques.

Cette volonté vient de l'idée que des données publiques gratuites et réutilisables sont une preuve de la transparence des politiques publiques et un vecteur de leur évaluation, d'une part, et que d'autre part des acteurs privés sur un modèle marchand ou non marchand peuvent en utilisant ces données, produire des services sortant des missions du service public et éventuellement créer de la valeur.

Comme le souligne Gautier Poupeau⁶², l'ouverture des données des bibliothèques a montré que l'intérêt du monde extérieur s'est porté sur des données non bibliographiques. La culture professionnelle concentrée sur les données bibliographiques n'encourage pas l'ouverture du spectre des données à exploiter. Ainsi les administrations ne sont pas forcément les mieux placées pour identifier les bénéfiques et l'intérêt des données qu'elles produisent.

Pour les bibliothèques, ouvrir et mettre à disposition leurs données, oblige à repenser largement les modèles afin de garantir l'anonymat et la confidentialité de ces données. La maintenance des données (formats, actualisation etc.) doit être

⁶⁰ La gazette des communes, n°46/2344, dossier data, p 28... et 44...

⁶¹ <https://twitter.com/mandonthierry/status/78912653055338754/photo/1> (consulté le 12/12/2016)

⁶² POUPEAU, Gautier « La donnée nouvelle perspective pour les bibliothèques » in BERMÈS, Emmanuelle (éd.), 2017. L'avenir des catalogues. Le cercle de la librairie.

pensée à la base, pour en permettre l'exploitation. C'est là une culture différente à mettre en place et à maintenir. Elle doit être nécessairement portée par la direction et soutenue par la formation.

Les données changent de place, les données internes ont vocation à devenir pour l'essentiel des données externes. C'est un changement de paradigme : cela nécessite de changer de logique dans l'administration de ces données. La maintenance prend souvent

« insuffisamment en compte l'impact des changements sur les réutilisateurs potentiels, l'hétérogénéité éventuelle tant des formats d'exposition que de la structure oblige à les traiter avant toute exploitation⁶³. »

Pour les administrations elles-mêmes cette ouverture peut être l'occasion de reverser des données dans leurs propres SI après correspondance de référentiels. C'est par l'intégration progressive de référentiels et de ressources qu'Isidore s'est développé⁶⁴.

Les questions qui sont posées sont des questions techniques de maintenance, d'adaptation et de risque d'évolution des formats ou de la structure des référentiels que l'on récupère.

Si le monde extérieur aux bibliothèques a montré son intérêt pour certaines données peu valorisées, les bibliothèques restent encore assez timides quant à l'exploitation, à la valorisation de données qu'elles sont en mesure de produire. Ainsi, les technologies des objets connectés sont très peu exploitées en bibliothèque alors même qu'elles furent parmi les premières à adopter la RFID. Les semaines test sont toujours de mise pour mesurer la consultation des collections, avec les biais statistiques que cela implique, sans compter la mobilisation des personnels. Pourtant la RFID peut donner des solutions de comptage automatisé. Le comptage peut également être affiné en disposant des détecteurs de mouvements dont les coûts sont aujourd'hui assez minimes.

Si certains établissements témoignent d'un intérêt pour ce genre de solution, les réticences sont importantes. La crainte d'outrepasser ce qu'il est possible de collecter sur les usages des lecteurs est légitime, et l'emporte le plus souvent. Cette prudence risque de laisser à d'autres le soin de collecter ces données. On peut citer l'exemple d'Affluences.

Ce service fourni par une jeune entreprise française s'appuie sur deux leviers. Des capteurs qui collectent les compteurs d'entrée et de sortie dans les établissements. Ces capteurs sont dans certains cas ceux de la bibliothèque et Affluences en récupère seul les données en temps réel. Dans d'autres cas il s'agit de capteurs posés par l'entreprise elle-même. Avec ces éléments et les jauges de chaque établissement, Affluences peut produire un service au travers d'une application sur téléphone, donnant pour les bibliothèques environnantes le nombre de places libres et les horaires. D'autres usages des données sont envisagés par Affluences, dans un cadre qui ne sert pas la bibliothèque cliente, quand bien même les données proviennent des dispositifs de la bibliothèque. Mais seul Affluences

⁶³ POUPEAU, Gautier, « la donnée : nouvelles perspectives pour les bibliothèques » in BERMÈS, Emmanuelle (éd.), 2017. L'avenir des catalogues. Le cercle de la librairie

⁶⁴ <http://rechercheisidore.fr/referentiels> (consulté le 12/12/2016)

les récupère. Dans une logique Big Data, toutes les données ont leur importance. Ici on devrait appliquer le « just in case ».

Par ailleurs, cela pose des questions de responsabilité sur les données. Pour les cas où les outils de collecte appartiennent à Affluences, on peut considérer que les données sont leur propriété, mais pour les cas où ce sont les outils de la bibliothèque qui produisent les données, il ne s'agit que d'une mise à disposition pour un service précis. Or cette mise à disposition ne permet pas de garantir et de maîtriser les réutilisations qui sont faites des données fournies par la bibliothèque.

La maîtrise des données n'est pas toujours garantie quand on fait appel à un prestataire et si le contrat ne précise pas clairement le cadre d'utilisation des données et les garanties à fournir pour s'en assurer. Il y a dans ce type de contrat un risque de voir sortir des bibliothèques des données qui sont leur cœur de métier, et dont le cadre d'usage par les sous-traitants n'a pas les mêmes contraintes juridiques que les administrations. Par cette fuite de données, il peut y avoir une perte de maîtrise de l'activité, non désirée, non maîtrisée et un risque supplémentaire de platformisation.

« Partout dans les administrations, le numérique a été traité à partir de l'informatique, service par service, l'ensemble composant le système d'information, et problème par problème : l'insatisfaction des usagers, l'inefficience de l'information descendante, la concurrence imprévue de services Web grand public (...) Des outils s'ajoutent aux outils, des interfaces conçues localement améliorent la fluidité, mais compliquent la maintenance et les mises à jour. Ce temps est passé. Sous l'alerte, la compréhension doit devenir systémique (...) Dans beaucoup de grandes organisations, ce rôle de facilitation et d'acculturation est dévolu au Chief Digital Officer, CDO (...) En réalité, l'enjeu est bien plus profond pour l'ESR. Le risque n'est pas seulement une baisse de la performance de l'organisation et de ses processus. C'est celui de la fuite des usagers vers des outils privés plus intuitifs, de sorte que les contenus de valeur se trouvent externalisés dans une grande indifférence, une perte grave quand les universités pourraient constituer un grand réseau social et faire de la recherche sur les données de l'apprentissage⁶⁵. »

Thierry Mandon, secrétaire d'état au numérique appuyait très récemment, devant les vice-présidents pour le numérique des universités le rôle attendu de ces CDO⁶⁶.

Il faut voir ici la volonté forte appuyée par les CDO de modifier les logiques institutionnelles au profit de logique d'innovation, de disruption. Les bibliothèques, leurs données et l'expertise acquise en matière d'échange, de transformation et de réutilisation des données leur donnent une place privilégiée dans cette dynamique nouvelle. Il est trop tôt pour dire si ce sera largement le cas. Cependant quelques initiatives peuvent être lues comme allant dans ce mouvement.

⁶⁵ Conseil national du numérique, 2016. Université numérique du temps des explorateurs à celui de la transformation Avis du Conseil national du numérique [en ligne]. [Consulté le 20 décembre 2016]. Disponible à l'adresse : https://cnumerique.fr/wp-content/uploads/2016/05/CNNum_Avis2016-1_ESR-4.pdf

⁶⁶ *Ibid* (consulté le 12/12/2016)

Ainsi le projet PREVU⁶⁷ porté par l'université de Paris 8 illustre la valorisation de données d'usage des collections de la bibliothèque, par une coopération entre des chercheurs de l'université et le SCD. A partir des données d'emprunts, il est possible de visualiser aisément les éditions les plus consultées de Bourdieu ou Deleuze, selon le sexe, ou le champ disciplinaire des usagers. Ce projet de recherche orienté vers la visualisation de données constitue un outil de connaissance sociologique pour l'université et pour le SCD. L'outil développé dans le cadre d'un projet ANR ne semble pas prévoir d'autres contextes d'utilisation. Pour autant, un tel outil transformé en plateforme comme ezPAARSE couvrirait des besoins importants.

Même si les initiatives de dataviz, sont encore rares - 2 établissements parmi les 28 du questionnaire, dont le projet PREVU - elles témoignent avec l'informatique décisionnelle de la volonté de sortir les données de leurs silos pour les faire parler. Dans leurs silos, prises dans les applications métiers, les données sont plus difficiles à manipuler, plus difficiles à faire parler. Les sortir du silos n'est pas qu'un enjeu technique ou organisationnel. Cela répond aussi au besoin de mieux communiquer la situation, l'activité, les enjeux d'un SCD dans ses différents aspects vis à vis de leur hiérarchie. Depuis la loi LRU, les équipes de direction d'université désormais seule tutelle des SCD doivent apprendre à connaître les bibliothèques. Ces mises en forme des données hors de leur silo sont aussi un effort de transparence et de pédagogie, d'abord vers les hiérarchies mais aussi vers les bibliothécaires eux-mêmes.

Ainsi à Yale, le logiciel Tableaux Public, logiciel de data visualisation capable de traiter un grand nombre de données et de les rapprocher quand elles proviennent de différents silos, est utilisé pour aider les acquéreurs dans leurs choix. Quelques éléments sont présentés en ligne⁶⁸.

Certains établissements font également ce travail sur les données bibliométriques de leur université. Il s'agit le plus souvent d'évaluer les données concernant l'université de rattachement à partir des archives (HAL ou archives institutionnelles). A ce titre, comme fonction support, relativement indépendante des missions de recherche de l'établissement, et disposant d'une expertise dans les questions de signalement des publications, les SCD jouent le rôle de tiers de confiance pour évaluer l'activité éditoriale de l'université. C'est un prolongement naturel des fonctions documentaires telles que pensées par Paul Otlet et repris par J.M. Salaün.

« L'article scientifique est autant le lien de l'exposé de la preuve que celui du partage des connaissances. Il deviendra même plus tard l'outil principal de régulation de la communauté scientifique dans une version de la bibliométrie que Paul Otlet n'avait pas envisagée⁶⁹. »

Les archives institutionnelles servent à la bibliométrie donc à l'évaluation mais de façon imparfaite comme déjà évoqué.

On retrouve également la valeur probante dans la communauté scientifique des publications d'articles et plus encore en lui adjoignant les données de recherche associées. Les incitations et les obligations de l'UE (Union Européenne) dans

67 <http://prevu.fr> (consulté le 12/12/2016)

68 Certains de ces outils sont consultables ici <http://public.tableau.com/profile/sarah.tudesco#!/> (consulté le 12/12/2016)

69 SALAÜN, Jean-Michel, 2012. Vu, lu, su. Paris : La Découverte.

H2020 devraient favoriser un effet de levier intéressant pour remettre les bibliothèques au centre de l'économie et de la régulation scientifique par la gestion des données de signalement des publications et des données de la recherche. Garanties de la cohérence des données et donc de leur exploitabilité, à des fins probantes et à des fins d'évaluation, les bibliothèques universitaires sont en bonne position pour accompagner ce mouvement, à condition qu'elles sachent porter les solutions et les mettre en œuvre.

Les outils des éditeurs privés, comme les réseaux sociaux de la recherche qui concurrencent les entrepôts de publications, épargnent encore les données de la recherche. Mais pour combien de temps ? Une prise de position rapide sur ce dossier devrait permettre de remettre les bibliothèques dans leur rôle de tiers de confiance dans l'économie de la connaissance.

Dans la perspective du modèle de Dominique Cardon, les tensions sont ici entre deux économies et deux modèles :

- Le modèle des réseaux sociaux de la recherche qui se placent dans les données, et dans une économie à court terme de rentabilité économique prenant appui sur une position de conglomérat industriel. La plateforme de réseau social est là pour inciter les chercheurs à rester dans un écosystème unique, celui de l'éditeur.

- Le modèle des entrepôts de publications est celui de la mesure, de l'évaluation. La bibliothèque est donc au dessus des données qu'elle collecte et exploite. Elle permet de mesurer « l'autorité des méritants », dans la stricte application des principes de la bibliométrie que Google a repris pour inventer le PageRank.

Avec les données de la recherche, la bibliothèque sert la communauté scientifique, et met à disposition ses capacités et ses compétences dans la gestion des données, pour appuyer les publications des données faisant foi de scientificité. Elle n'est là que pour classer et collecter sans aller dans les données qu'elle collecte. C'est un tiers de confiance pour la communauté scientifique. Sa position devrait paraître plus légitime et l'est déjà d'après les études et données réunies par mon enquête. Ce pourrait permettre un fort effet de levier pour enrichir et développer les entrepôts de publications.

On le voit, faire parler les données implique de plus en plus de les faire sortir de leurs silos d'origine. Cela nécessite une expertise particulière en terme de maîtrise des formats et d'analyse de l'information, expertise reconnue des métiers des bibliothèques et de la documentation au sein des universités⁷⁰. A ce titre, le traitement, le format, les modes de collecte et d'exploitation des données, bref leur gestion, sont un enjeu stratégique pour les établissements universitaires, car ils sont des outils indispensables de la politique des universités.

2.2. MISES EN ŒUVRE DE LA GESTION DES DONNEES

Le questionnaire venu compléter les entretiens a proposé huit catégories de données ou de projets dont on verra qu'ils ne sont pas exclusifs et peuvent même

⁷⁰ FAUSSURIER, Bérengère, 2016. Le SCD et son université de tutelle. [en ligne]. 1 janvier 2016. [Consulté le 18 décembre 2016]. Disponible à l'adresse : http://bbf.enssib.fr/matieres-a-penser/le-scd-et-son-universite-de-tutelle_66698

se recouper : données de la recherche, archives institutionnelles, données collectées automatiquement par RFID ou autres objets connectés, données de services partenaires, data visualisation, FRBR, Informatique décisionnelle. Sur 28 établissements, seuls deux ont répondu négativement à l'ensemble de ces huit catégories, cinq ont répondu positivement à une seule de ces catégories. En d'autres termes, la question des données en dehors de la forme traditionnelle des catalogues mobilise une bonne part des SCD. 93% de l'échantillon est impliqué dans un projet de cette nature, et 75% dans plusieurs projets de cette nature.

La multiplication des silos est une réalité largement partagée en BU, mais recouvre des réalités différentes selon les contextes et les priorités.

Cette réalité influe d'ailleurs sur les orientations des structures « support ». L'INIST confirme son rôle d'appui pour la gestion des données de la recherche⁷¹, après avoir traduit les recommandations⁷² en la matière, et porté des formations dans le domaine⁷³. Il propose désormais un outil, OPIDOR⁷⁴, pour une Optimisation du Partage et de l'Interopérabilité des Données de la Recherche. Une collaboration a également été mise en place pour une enquête de besoin auprès des chercheurs⁷⁵ entre l'INIST et l'URFIST de Nice. Ce projet est porté par BSN 9⁷⁶ et BSN 10, les volets « formations » et « données de la recherche » du principal dispositif national d'action publique en matière d'information scientifique et technique porté par le MENESR (Ministère de l'Éducation Nationale, et de l'Enseignement Supérieur et de la Recherche) depuis 2009. Il s'agit d'étudier les besoins des chercheurs en matière de données de la recherche aussi bien en terme de compétences à mutualiser (que l'INIST peut porter) qu'en terme de formation que les URFIST ont vocation à donner. Ce projet lancé en 2015 souligne les dispositifs en cours de mise en place sur le sujet. Il ne s'agit pas pour l'heure de mettre en place des infrastructures lourdes et mutualisées sur le modèle de HAL, mais d'établir un cadre de travail, porté par l'INIST qui pourrait favoriser la mutualisation des compétences, l'émergence d'une communauté de pratiques, formée par les URFIST. On observe donc une stratégie en deux volets, l'un de coordination et d'expertise nationale en matière de recommandations, de bonnes pratiques, avec un appui local pour appuyer la coordination par des formations et un relai régulier auprès de la communauté de chercheurs. Pour cela, le ministère a doté les URFIST d'une personnalité juridique collective qui leur permet désormais de travailler officiellement ensemble, un GIS (Groupement d'Intérêt Scientifique) dont l'établissement porteur, récemment désigné est l'université de Côte d'Azur. Les URFIST jusque-là souffraient d'un certain isolement avec, environ 20 personnes réparties sur 7 unités régionales soit moins de 3 personnes par unité et d'un manque d'identification par les personnels de l'Enseignement Supérieur, en dehors d'un cercle d'initiés assez restreint. Cet éclatement administratif et territorial limitait les capacités de mobilisation communes soit pour des raisons de coordination, soit pour des raisons administratives et comptables. Afin de palier ces difficultés les URFIST s'étaient déjà organisées en association en 2014 et et

⁷¹ <http://www.inist.fr/?Donnees-de-la-recherche> <http://www.inist.fr/?Pour-faciliter-la-redaction-de&lang=fr&utmcontent=buffer63798&utmmedium=social&utmsource=twitter.com&utmcampaign=buffer> (consultés le 12/12/2016)

⁷² <http://openaccess.inist.fr/?Traduction-francaise-des-guides-795> (consulté le 12/12/2016)

⁷³ <http://www.inist.fr/?Donnees-de-la-recherche> (consulté le 12/12/2016)

⁷⁴ <https://opidor-preprod.inist.fr/> (consulté le 20/12/2016)

⁷⁵ <http://www.doranum.fr/2016/03/04/le-projet-doranum/> (consulté le 12/12/2016)

⁷⁶ Bibliothèque Scientifique Numérique : <http://www.bibliothequescientifiquenumerique.fr/> (consulté le 20/12/2016)

doté d'un système commun d'inscription aux formations dispensées. L'activité des URFIST dispose donc désormais d'une visibilité nationale en terme de formation (portail mutualisé) et de recherche (GIS). Le projet Doranum porté par l'URFIST de Nice a vocation à être appuyé et relayé par les autres URFIST via le GIS.

Unités mixtes d'enseignants-chercheurs et de conservateurs de bibliothèque, elles relèvent directement de la présidence des universités et à ce titre disposent d'un positionnement tout à fait intéressant. Ces différents éléments témoignent de la volonté de l'INIST de devenir avec le relais des URFIST un interlocuteur de premier plan dans le domaine des données de la recherche, sur le plan métier et sur le plan technique, en mêlant la force d'action d'un gros établissement, comme l'INIST et la souplesse et la proximité de petites structures que sont les URFIST. L'équilibre entre le monde universitaire où évolue les URFIST et celui de la seule recherche où l'INIST évolue, par son lien au CNRS, reste sans doute encore à trouver. Ainsi on voit émerger des services extérieurs encourageant la mutualisation, le partage de bonnes pratiques au sein des universités et des structures de recherche. La coopération entre deux structures support, INIST et URFIST, apparaît dans ce schéma, comme une nouveauté en cherchant à articuler une démarche nationale et une coordination locale.

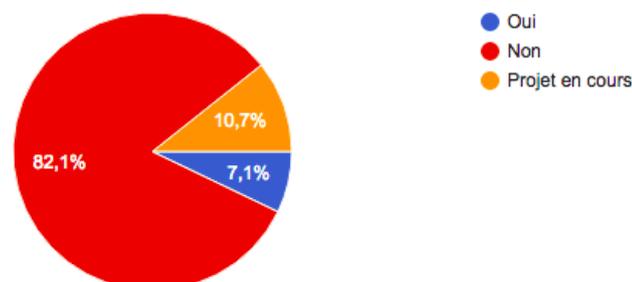
Une des personnes interviewée évoquait d'ailleurs la possibilité de se rapprocher de l'INIST qui

«se propose d'être un prestataire de service pour les données de la recherche ».

Cet appui semble en effet nécessaire à différents égards. Les DMP (data management plan, obligatoire pour H2020 et toute forme d'archivage de données) sont encore peu nombreux en France. Seuls 2 SCD signalent l'existence d'un DMP dans l'université, 3 signalent un travail en cours sur le sujet, soit moins 18% de l'échantillon. Un SCD signale l'existence d'un DMP mais n'y a pas été associé. Dans deux autres cas le DMP est à l'initiative du SCD. Dans un cas, le SCD est associé mais pas pilote. Enfin, dans le dernier cas le SCD est pilote de l'entrepôt et du DMP mais sans qu'on puisse savoir si l'initiative du projet lui revient. Au vu de ces éléments, la gestion des données relève d'une pratique minoritaire et est très souvent portée par le SCD lui-même.

Votre université a-t-elle un plan de gestion des données de la recherche ?

(28 réponses)



Cependant d'autres données éclairent cette réalité sous un autre angle. Les établissements ayant répondu à la question sur l'association du SCD à la

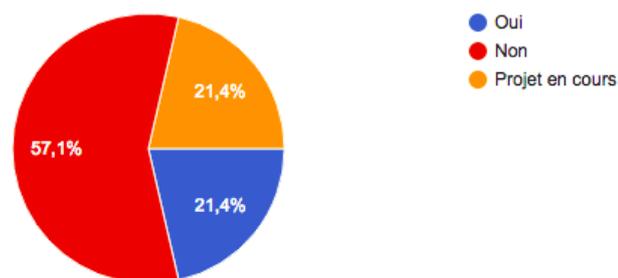
conception du DMP sont plus nombreux (12 SCD) que ceux qui ont dit avoir connaissance de l'existence d'un DMP ou d'un travail en cours sur un DMP (6 SCD). Cette apparente incohérence laisse penser que 43% des SCD ayant répondu travaillent sur un DMP. Ils y réfléchissent sans qu'il y ait un réel projet autour des données de la recherche, mais un DMP est nécessaire pour des besoins ponctuels en lien avec des projets de recherche, sur le modèle H2020 probablement.

« Ponctuellement, nous sommes sollicités par un chercheur ou un labo pour gérer ses données. Il n'existe pas de démarche globale au niveau de l'établissement »

précise l'un d'entre eux dans le questionnaire. Plusieurs entretiens confirment ce type de collaboration, soit que la bibliothèque va au devant des chercheurs soit que ceux-ci sollicitent la bibliothèque pour archiver et conserver des données liées à un projet spécifique.

La bibliothèque est-elle associée au travail sur la gestion des données de la recherche ?

(28 réponses)



Ainsi, si un plan global n'est pas encore envisagé dans la majorité des SCD, près de la moitié de ceux ayant répondu à l'enquête travaillent de façon effective mais plus localisée, plus opérationnelle sur ces questions. Cela témoigne d'une tactique d'avancée par petits pas, de rapprochements progressifs. Cela permet de construire des habitudes de collaboration et des liens de confiance pouvant aider à terme à travailler de façon plus globale. Une des approches de cette politique des petits pas est la bibliométrie, que l'on trouve mentionnée à deux reprises dans le rôle attendu du SCD en matière de gestion des données de la recherche. Cela signifie que dans ce cas précis, le lien entre publications dans une archive ouverte et données de la recherche passe pour le SCD par la bibliométrie, un rôle d'évaluation et de mesure. Le rôle de signalement et de gestion est donc dévolu à d'autres acteurs de l'université :

« On a du mal à penser la gestion de la donnée dans l'institution parce qu'on ne la voit pas, on ne sait pas ce que c'est, d'une part ; et par ailleurs, très souvent, c'est le pré carré des DSI qui en fait sont très *old school*, ne comprennent pas les problèmes data ».

Une fois les liens établis sur la gestion ponctuelle de données de la recherche, au delà de la seule bibliométrie, l'étape suivante serait celle décrite par un des établissements travaillant sur un DMP :

« La bibliothèque pilote le projet : nous travaillons actuellement avec quatre projets pilotes de mise en ligne de données de recherche en SHS. Nous évaluons les besoins des chercheurs, nous rédigeons les cahiers des charges, nous proposons des scénarios de réalisation, nous assurons la curation des données, nous nous chargeons des questions juridiques ou du lien avec les services de l'université compétents, et nous sommes maître d'ouvrage pour la réalisation, les services informatiques de l'université étant maître d'œuvre. En parallèle, nous mettons au point le plan de gestion des données avec le service des archives et les services informatiques. Nous créons également un entrepôt de données. »

Parmi les six établissements ayant répondu ne pas connaître l'existence d'un projet de DMP, trois sont clairement dans cette politique des petits pas. C'est également le cas de trois des établissements ayant participé à un entretien. Le TGIR Huma-Num, contacté pour un entretien insistait d'ailleurs sur sa politique d'implication des SCD dans les projets de recherche, même si les collaborations ne sont pas toujours simples dans certains établissements universitaires. Cependant les trois SCD de l'enquête quantitative parmi les six n'ayant pas de DMP connus s'inscrivent dans une démarche un peu différente. Ils disent travailler de façon prospective sur le sujet, et sans relation effective avec les chercheurs sur des projets concrets :

- « Le SCD avec l'enseignant délégué aux humanités numériques de l'Université »
- « Démarche prospective conjointe de l'URFIST et de la mission Recherche de la BU » ; « le SCD mène un travail prospectif sur les données, mais pour l'instant sans implication d'autres acteurs de l'université »
- « Insérée dans une petite structure de recherche à l'étranger, la bibliothèque est l'interlocuteur unique de la direction sur le sujet. Et elle espère en faire un objectif du prochain contrat quinquennal »

Ces SCD travaillent sur un terrain décisionnel plutôt qu'opérationnel. Ici la démarche est de se préparer à ces évolutions, en s'inscrivant dans l'écosystème universitaire, sous un angle stratégique ou politique, en préparant les décideurs, pas les chercheurs.

On peut cependant souligner qu'aucun établissement n'a mentionné dans l'enquête de relation avec l'INIST ; un l'a évoqué pendant l'entretien comme une possibilité. Il est vrai que la politique de l'INIST sur le terrain des données de la recherche semble s'affirmer depuis la rentrée 2016 et ne peut pas déjà être mesurée. Des travaux ultérieurs sur le sujet pourraient aider à mieux cerner le type de collaborations attendues et son efficacité.

On peut souligner également dans l'enquête la mention d'une collaboration avec l'URFIST, sans qu'il soit possible de préciser s'il s'agit de celui de Nice ou celui de la région du SCD ayant répondu à l'enquête en ligne.

On observe à la fois une mobilisation significative des bibliothèques sur ces nouveaux terrains (données de la recherche, web sémantique etc.) et en même temps une permanence des organisations de travail. L'organisation change à la marge, « par petite touche », comme le montrent les organigrammes ou différents entretiens. Ainsi, après avoir évoqué longuement les travaux sur le web de

données, et sur les données de la recherche dans son établissement, un interlocuteur répondant à la question sur la place de la chaîne du traitement du document la voit comme un élément central et structurant du travail de la plupart des agents :

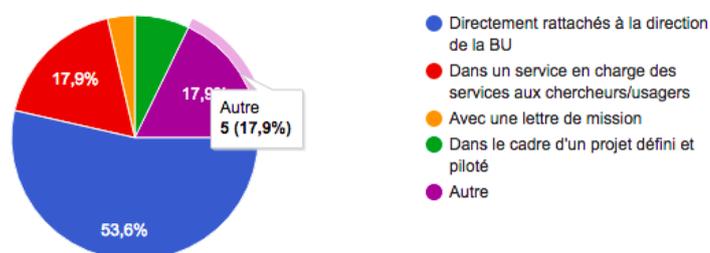
« c'est encore très présent, c'est le SIGB qui est le cœur de l'outil métier. Mais du côté de la GED⁷⁷ il y a une réorganisation pour tenir compte de la production. »

Dans un autre établissement, on l'a déjà évoqué, c'est le SIGB qui est utilisé pour intégrer les nouveaux circuits documentaires afin de ne pas créer de nouveaux silos. La mise en place de nouveaux types de circuits, de nouveaux services nécessitant des processus de travail distincts, ne se fait que marginalement à la place des circuits antérieurs. De même que les nouveaux formats de données et protocoles d'échanges (RDF, XML, Dublin Core, OAI-PMH etc.) coexistent avec leurs aînés (MARC, Z-39-50), les organisations de travail anciennes et nouvelles coexistent au sein du SCD.

Ici ou là des outils, des dispositifs managériaux sont utilisés pour piloter ces nouveaux circuits le plus harmonieusement possibles (démarche qualité, ou pilotage par les indicateurs). Parfois ils sont intégrés grâce aux outils traditionnels quittes à les tordre un peu, tout en travaillant sur l'organigramme et sur les missions de chaque agent. Comme évoqué plus haut, ces différents types de données supposent une stratégie différente, un positionnement différent. Quand la même structure doit intégrer ces différents types de données et donc se positionner différemment vis à vis de ces données, plusieurs organisations sont nécessaires vis à vis de chaque unité organisationnelle ce qui peut être facteur de tensions dans le SCD.

Les entretiens ne témoignent pas de telles tensions⁷⁸ de façon explicite. Cependant plusieurs éléments laissent penser qu'elles existent. Nous avons évoqué plus haut le fait que la plupart des personnels travaillant sur ces nouveaux silos de données sont soit rattachés à la direction (53% des cas), soit rattachés à un service en charge des services aux chercheurs/usagers (17%) soit dans un service d'informatique documentaire (17%).

Dans quels cadres ces personnels travaillent-ils ? (28 réponses)



La forte proportion de personnels rattachés à la direction ne peut s'expliquer par le seul fait de la nouveauté ou du caractère stratégique du dossier. Comme

⁷⁷ Utilisée pour intégrer les mémoires, thèses de l'université concernée

⁷⁸ Le terme employé dans tout ce développement n'est pas à prendre dans le sens de frictions dans des relations inter-personnelles mais au dans celui du résultat de forces pouvant s'exercer sur des structures mécaniques mues par des forces contraires

beaucoup de sujets nouveaux, ils peuvent perturber le reste de l'organisation et il s'agit de les protéger tout en protégeant l'organisation de leurs éventuels effets.

Par ailleurs, différents entretiens évoquent des outils de régulation interne. Ils sont décrits pour gérer des difficultés, des obstacles liés à des arbitrages, à l'allocation de moyens humains. Il est difficile de rentrer plus dans les détails sauf à trahir la confidentialité des entretiens sur un sujet délicat. Il ressort de cela l'existence de tensions qui s'expriment par des retards de décision de la direction (SCD ou université selon les cas) sur des sujets stratégiques, par des méthodes de régulation interne dans la définition de priorités, et par des solutions d'accompagnement collectif ou individuel (accompagnement au changement, formation principalement).

Le recours à des structures tierces pour mutualiser un projet, pour partager des outils peut aussi être considéré comme une forme de régulation des tensions internes. Il peut s'agir d'une communauté d'usage qui rassemble des individus dans des situations similaires dans leurs organisations où certaines difficultés sont difficilement compréhensibles. Leur isolement au sein de l'établissement car le sujet est nouveau, n'aide pas au partage et à la bonne compréhension de ces difficultés. Se rassemblant pour des raisons techniques, financières ils peuvent également partager ces difficultés et les stratégies mises en œuvre pour les gérer, les contourner. Ces formes d'organisation en dehors de la structure d'origine sont de ce point de vue bénéfique à tous. Elles sortent les questions et les tensions éventuelles de la structure d'origine et aident à revenir avec des réponses qui peuvent relever de l'organisation de la structure d'origine mais aussi de l'outil mutualisé. A condition que l'organisation mutualisée laisse la place au cas particuliers, aux situations individuelles, et soit en mesure de proposer des solutions.

Cependant certaines tensions ne peuvent trouver de résolution dans ce cadre tant elles sont consubstantielles à la différence de nature des différents silos de données et de leur contexte d'usage. Ces questions animent non seulement les établissements mais la profession et l'interprofession. Elles concernent pour une large part les conditions de réutilisation des données produites par les bibliothèques, gratuité, mention d'origine, licence de réutilisation. Si les données bibliographiques ne font que peu débat sinon sur le choix de la licence CC ou Etalab, les données produites par la numérisation des fonds font l'objet de nombreuses polémiques. Elles traduisent la tension entre les modèles et posent la question de la valeur au sens marchand. Réside-t-elle dans le document numérisé ou dans les données d'usage que l'on peut collecter ? Le premier choix repose sur le modèle historique des bibliothèques où la valeur est dans la sélection et la mise à disposition, et se place au dessus des données. Le deuxième choix se situe à coté des données pour en connaître l'usage et le mesurer.

Cependant, une part d'un débat qu'il ne s'agit ici que d'évoquer pour illustrer les tensions entre les modèles sans le prolonger, est réglé avec la loi sur la République numérique. Le caractère public, ouvert que doivent prendre les données des administrations afin de favoriser les réutilisations pose différentes questions aux BU. Quelle part de nos missions peut-on consentir à laisser réinvestir par la collectivité en libérant les données, afin de rendre de nouveaux services ou des services repensés ? Et quelle part doit être investie en priorité par

la communauté professionnelle dans le double objectif de concentrer nos compétences là où elles peuvent pleinement s'exprimer et de ne pas favoriser la création de plateformes de services qui se glissent entre les bibliothèques et leur cœur de métier⁷⁹?

Quand le choix est fait de libérer les données, la dissémination, les alignements multiples, structurant accès et visibilité des ressources posent de façon plus accrue la question de la qualité des données. Leur masse questionne les méthodes traditionnelles de gestion de la qualité et donc l'organisation du travail pour la garantir. La question de la gestion de la qualité dans les organisations de travail si elle est présente n'est pas aussi répandue que le travail massif sur les données. Il faut espérer et s'attendre à un développement significatif des démarches qualité dans les établissements impliqués dans des processus en lien avec le web de données, du fait de la forte réutilisation des données produites et de la nécessaire maîtrise de la fiabilité par les processus.

Le développement des humanités numériques est également un facteur qui peut modifier le rôle, la place et donc l'organisation des SCD dans les universités concernées. La mise à disposition d'outils tels que « Voyant Tools » par Humanum, ainsi que l'explosion des projets de recherche dans ces domaines implique un appui métier sur les données ce qui n'est pas sans poser question notamment en termes d'organisation.

Le TGIR Huma-Num ne propose pas de services aux bibliothèques, c'est un service à destination des chercheurs. Il propose des solutions de traitement, de stockage, d'organisation et d'enrichissement des données, de signalement et d'accès. Si ces services ne s'adressent pas directement aux bibliothèques, Huma-Num souligne l'importance d'associer aux projets de recherche des professionnels de la documentation, documentaliste ou bibliothécaire :

« C'est ce qu'on fait quand ils viennent nous voir en amont. Ils nous disent souvent, j'ai un corpus constitué et pour le traiter je vais prendre un doctorant. Nous on dit attention, le job d'un doctorant c'est de faire une thèse, pas de concevoir un modèle de données. Plutôt que de prendre un doctorant, prenez un documentaliste ou un bibliothécaire dont le boulot c'est de concevoir des modèles de données. On insiste beaucoup là-dessus. »

Ainsi, le développement des humanités numériques construit et rend nécessaire le rapprochement entre le monde de la documentation et des bibliothèques d'une part et celui de la recherche d'autre part. La nécessité ne tient pas seulement à des questions de compétence à mettre autour d'une table pour mener à bien un projet. C'est aussi rendu nécessaire pour la conservation des données, la gestion du cycle de vie des données du projet de recherche. Les chercheurs, d'après cet entretien ne sont pas du tout conscients des enjeux que recouvre la gestion des données.

« C'est absent, c'est peu perçu car les chercheurs ont transposé ce qui se passait dans le monde analogique dans le monde numérique. Le chercheur

79 PÉRALES, Christophe, 2016. Future is now: « ubérisation » de la science et économie de plateforme | BibliOpen. [en ligne]. 2 octobre 2016. [Consulté le 18 décembre 2016]. Disponible à l'adresse : <http://blog.educpros.fr/christophe-perales/2016/10/02/future-is-now-uberisation-de-la-science-et-economie-de-plateforme/>

n'a pas de responsabilité dans la conservation⁸⁰ de ses œuvres ou des ses archives de travail. (...) Alors on essaie d'impliquer les chercheurs sur cette question. On fait de la gestion partagée. On implique les chercheurs sur le cycle de vie des données, est-ce que ces données vont être obsolètes ou est-ce qu'il faudra leur donner une continuité. Vous <les bibliothèques> avez la responsabilité scientifique là-dessus. »

Cependant cette question reste peu traitée par les projets de recherche. Quand la bibliothèque est sollicitée directement sur ces sujets, c'est pour le stockage ou des questions de modèle de données pour les traiter. Jamais pour les conditions d'accès, leur conservation, ou les questions juridiques qui peuvent les entourer (captation complète de site web, données personnelles etc.) Pour autant les bibliothèques peuvent apporter une véritable expertise sur ces questions et fournir aux chercheurs un appui qui sécurise les données, leur viabilité et ce faisant celles des recherches elles-mêmes. Si les recommandations et maintenant les obligations des projets H2020 vont dans ce sens, cela ne concerne que H2020, et donc une faible portion des projets. Les projets ANR ne sont pas encore soumis à de telles obligations. Il serait sans doute du devoir des bibliothèques de l'enseignement supérieur que de faire des propositions aux tutelles pour que les projets de recherche émettent des conditions sur la gestion des données. Les bibliothèques peuvent offrir un appui dans ce domaine. H2020 et l'exemple anglo-saxon sont les éléments de référence pour appuyer une telle démarche.

Cependant, apparaissent dans l'enquête des obstacles organisationnels à cette démarche de la part des bibliothèques. Evidemment la question des ressources humaines est majeure, tant en terme de personnels disponibles dans ce rôle d'appui, mais aussi en terme de formation et de compétence. Si ces points sont essentiels, ils ne sont pas les seuls.

Les projets de recherche ont besoin d'agilité, de prise de décision rapide, d'exécution rapide, évaluer, revenir, refaire pour optimiser, expérimenter.

« Je le dis pas assez souvent, sans doute, Isidore en terme de temps de décision c'était 10 minutes. Et un an pour la réalisation. Je m'astreignais à ne jamais prendre de décision quand il fallait plus de 7 mn de réflexion. Cela a beaucoup dynamisé le projet. » Huma-Num en tant qu'« UMS (unité mixte de service) relève plus de la start-up que d'autre chose. On laisse 20% du temps aux équipes pour du travail perso, trouver des collaborations etc. Mais il y a des organisations dont ça n'est pas la culture historique ni métier. (...) On fonctionne intégralement en méthode agile⁸¹. Ce qui est compliqué avec la méthode agile c'est (sic) des grades de gens très différents. Pour les techniciens (de numérisation) c'est difficile de sortir du workflow. Donc maintenant on embauche des gens du même grade et qui peuvent sortir du workflow. »

On mesure ici par la nature des missions historiques des bibliothèques, acquérir, signaler, communiquer des collections et par l'organisation qui en découle (organisation des corps et organisation des missions afférentes, organigramme etc.) l'écart qui les sépare d'autres services support à la recherche.

⁸⁰ La réalité est plus complexe puisque les chercheurs sont supposés déposer leurs archives et bibliothèques professionnelles à la bibliothèque de leur établissement de rattachement.

⁸¹ ILLIEN Gildas « Une BnF agile, quand le développement logiciel fait bouger l'organisation du travail » in PÉRALES, Christophe, 2015. Conduire le changement en bibliothèque : vers des organisations apprenantes [en ligne]. Villeurbanne : Enssib.

On peut supposer à ce stade que les bibliothèques qui sauront se positionner comme un appui aux missions de la recherche sont celles qui sauront trouver des modes d'organisation permettant de dégager des ressources humaines d'encadrement au profit des projets de la recherche, et qui sauront régulièrement adapter leur organisation interne aux changements de priorités. On voit ici la portée de la démarche qualité dans le management global d'un établissement qui sait se transformer :

« l'esprit de l'organisation ne change pas c'est la concrétisation de l'organigramme qui s'adapte aux projets et à leurs besoins ».

Une bibliothèque capable de dégager pendant la durée d'un LABEX⁸², par exemple, un temps partiel ou un ETP pour appuyer le LABEX sur les questions de données non seulement renforce de façon durable les liens entre la communauté de chercheurs de l'université, mais aussi crée une véritable valeur aux projets car elle pourra ensuite inscrire la conservation des résultats des projets à la fin du LABEX dans ses collections.

Différents outils managériaux existent qui peuvent apporter aux établissements la capacité de modeler l'organisation à ses priorités. Parmi eux on a déjà évoqué la méthode agile. Cependant les cadres d'emploi, les identités professionnelles, le fonctionnement hiérarchique des administrations peuvent constituer des obstacles considérables à une évolution pourtant considérée comme souhaitable par les politiques publiques depuis une dizaine d'année et que les bibliothèques tentent d'intégrer⁸³. Mais quand les recrutements, les cadres d'emploi décidés depuis bien plus de 10 ans sont encore en exercice ou en application, les capacités d'évolution des établissements ne peuvent s'envisager sans affecter profondément les personnels et les carrières. C'est ce dernier aspect que nous allons désormais traiter.

⁸² LABEX : Laboratoire d'Excellence. Structure provisoire chargée pendant 5 ans de porter des projets de recherche et financée par les investissements d'avenir.

⁸³ INSPECTION GÉNÉRALE DES BIBLIOTHÈQUES, 2013. Quels emplois pour les bibliothèques ? Etat des lieux et perspectives - Livre et Lecture - Ministère de la Culture et de la Communication [en ligne]. [Consulté le 11 juillet 2016]. Disponible à l'adresse : <http://www.culturecommunication.gouv.fr/Politiques-ministerielles/Livre-et-Lecture/Documentation/Publications/Rapports-de-l-IGB/Quels-emplois-pour-les-bibliotheques-Etat-des-lieux-et-perspectives>

QUELS CHANGEMENTS POUR LE TRAVAIL EN BIBLIOTHEQUE ?

La multiplication des silos de données, l'exposition des données par le web sémantique ou les entrepôts OAI-PMH, la recherche et la constitution de nouveaux indicateurs pour évaluer l'action des bibliothèques sont autant de facteurs, parmi quelques autres, qui déplacent l'activité de la bibliothèque centrée sur son catalogue et la chaîne de traitement du document pour favoriser le développement d'autres services et d'autres manières de travailler. Ce recentrage est accompagné par un mouvement déjà connu et amplement étudié celui des services aux usagers, placés au cœur de l'organisation. Ce mouvement ne se fait pas sans remodeler les identités professionnelles laissant parfois certaines personnes sur le côté. Cela n'empêche pas un mouvement de fond qui s'appuie sur des compétences, des organisations de travail et les projets d'établissement pour incarner une logique de service permise par de nouvelles façons d'organiser les données.

Ce sont ces aspects successifs des changements dans le travail (organisation des données, organisation du travail et logique de service) qui s'opèrent qu'il s'agit donc d'étudier ici.

3.1. L'USAGER AU CŒUR ... OU LA STRATEGIE DU CHANGEMENT

Héritée de la vision documentaire de Paul Otlet, l'organisation du travail en bibliothèque est établie par un lien structurant entre la division des tâches et l'amélioration des services documentaires. Elle s'inscrit dans une vision tayloriste du travail documentaire, déjà évoquée, dont la chaîne de traitement du document est le pivot et dont le catalogue est la finalité.

Les catégories d'emploi en bibliothèque, les SIGB, les organigrammes bref, bon nombre des éléments structurants le travail sont aujourd'hui encore largement pensés par rapport à cette vision de l'organisation des fonctions documentaires de la bibliothèque.

Pour autant, plusieurs mouvements éclairés par des visions alternatives des missions du catalogue émergent ou s'affirment. La vision de Paul Otlet, issue de la pensée de son temps, structure les tâches pour rendre un service, une mission. La normalisation des tâches aboutit à la normalisation du service, équivalent pour tous, à caractère standard, qui colle assez peu aux évolutions des besoins et des usages en matière de service. Les bibliothèques universitaires ont entamé depuis une quinzaine d'années une évolution importante pour adapter leur fonctionnement et rendre des services plus adaptés aux usages.

Cette évolution coïncide avec la massification de l'informatique et le développement d'internet. Elle s'est donc largement appuyée sur cette révolution technologique pour adapter ses services aux nouveaux usages (catalogue en ligne et services distants autour du prêt, accès aux ressources électroniques, formations usagers en lien avec ces technologies). Mais l'organisation elle-même des processus de travail a été largement modifiée par cette révolution technologique

(échange de données entre bibliothèques, puis entre services documentaires de la chaîne du livre etc.) Ainsi non seulement le « front office » des bibliothèques s'est modifié, mais aussi le « back office », dans un mouvement parallèle sinon conjoint.

Si l'informatisation du catalogue et le développement d'internet ont pu dans les années 80-90 être un choc professionnel, nécessitant des formations aux nouveaux outils de travail, ils n'ont finalement que peu modifié la structure et le fonctionnement des bibliothèques universitaires. La chaîne tayloriste de traitement du document a été informatisée, mais ne s'est pas modifiée. On retrouve les mêmes étapes, et les mêmes acteurs. Seuls les outils ont changé. Ces outils ont cependant permis de libérer du temps pour commencer l'autre évolution des BU celle visant à replacer l'utilisateur au cœur de ses préoccupations, intégrant et s'appuyant sur la révolution internet pour ce faire.

Apparaissent ainsi les services d'informatique documentaire, puis les missions et services d'appui aux chercheurs et les missions qualité⁸⁴. La satisfaction des usagers se mesure, se pilote et pilote même les organisations de certaines structures, quand le pilotage des établissements ne se fait pas par les indicateurs eux-mêmes. Si ces évolutions ne s'observent pas partout avec la même ampleur aujourd'hui encore, le mouvement vers les services et les usagers est clairement amorcé.

Se développent également les plans d'accompagnement au changement, que l'informatisation des catalogues n'avait, en son temps, pas ou peu nécessité. Cela témoigne de la profondeur des changements en cours. Entre le numérique et une activité centrée sur les usages et moins sur les processus de travail, les bibliothèques partagent la même révolution que bon nombre de structures de service (publiques ou privées) : un déplacement de l'organisation du travail par les processus métier vers la satisfaction client. Selon les cas, cette satisfaction passe par une organisation produit ou usage. Cette évolution s'inscrit dans une transition numérique. Ces évolutions des organisations du travail ont fait l'objet de nombreuses études en sociologie du travail, sociologie des organisations notamment. Quelques éléments de certains de ces travaux semblent illustrer des situations rencontrées au cours de cette enquête.

L'ouvrage de François Dupuy, *Lost in management*⁸⁵ rassemble au travers de différentes expériences d'organisation du travail en Europe dans les secteurs privés et publics, les témoignages des évolutions en cours. Parmi ces organisations de travail, on trouve l'orientation client :

« Cette entreprise « globale » a entrepris de redéfinir son approche commerciale en passant d'une logique initiale très orientée sur les produits qu'elle fabrique et distribue à une logique davantage orientée vers les clients et les solutions qu'ils attendent. (...) Nous avons vu que cette démarche fait intrinsèquement partie de l'évolution qui touche tous les secteurs de l'économie marchande, à l'exception notable de ceux qui sont (encore) en situation de monopole. »

84 MARCEROU-RAMEL, Nathalie, « Nouvelles compétences, nouvelles organisations » in CAVALIER, François et POULAIN, Martine, 2015. Bibliothèques universitaires: nouveaux horizons. Paris : Éditions du Cercle de la librairie. Bibliothèques...

85 DUPUY, François, 2013. *Lost in management: la vie quotidienne des entreprises au XXIe siècle*. Paris : Éd. du Seuil. Points,

De ce point de vue, la bibliothèque qui n'appartient pas au monde marchand est on le sait bien en concurrence non seulement avec le web et l'accès direct à l'information, mais aussi avec l'ensemble des tiers lieu⁸⁶ qui offrent un espace intermédiaire entre le bureau et le domicile. L'orientation « client » est donc un réflexe de survie dans ce climat de concurrence. Pour autant d'après François Dupuy

« le changement ne se produit pas quand il est nécessaire mais quand il est possible ».

Ce sont pour partie, les marges de manœuvre dégagées par l'informatisation des catalogues et l'automatisation d'une partie des opérations de traitement du document qui ont rendu cette transformation possible.

Par ailleurs le recours aux outils numériques, l'analyse des indicateurs, des traces de navigation, sur le site ou sur les ressources électroniques sont guidés par cette préoccupation des usages, mais d'usages évalués, pas potentiels. On est ici dans des usages contrôlés, définis par le principe déjà évoqué du « just in case » pour la récupération des données afin de servir « just in time » le client. Ce principe était dans une certaine mesure à l'œuvre dans plusieurs des structures rencontrées au travers d'outils différents. Il se mesure, se construit dans chaque structure par la création d'outils documentés par des silos de données, à l'image des GAFA⁸⁷ et de leurs stratégies au service des clients.

Or, les travaux de sociologie du travail en général et ceux de sociologie du numérique⁸⁸ plus encore témoignent que ces stratégies s'appuient sur des organisations du travail très incertaines pour les agents qui y opèrent. A l'inverse, les organisations en râteau peuvent s'avérer très protectrices pour les agents, mais aussi pour les métiers et leurs identités.

« La routine peut abaisser mais aussi protéger. Elle peut décomposer le travail mais aussi composer une vie⁸⁹. »

Si le mode projet a été un outil privilégié pour désenclaver les unités organisationnelles, force est de constater que ni l'enquête ni les entretiens ne laissent beaucoup de doute : le mode projet⁹⁰ n'est pas en vogue dans les bibliothèques sur les questions étudiées.

Sur 28 établissements seulement 2 travaillent sur des sujets data dans le cadre d'un projet défini et piloté. De plus, un établissement travaille via une lettre de mission. Tous les autres travaillent dans leurs services respectifs ou directement auprès de la direction. Cependant on peut souligner que la formulation de la question dans l'enquête quantitative peut laisser une marge d'interprétation sous-estimant le recours au mode projet. Pour autant, les entretiens témoignent de la même réalité. Les nouveaux services construits autour des données le sont dans les services métiers concernés (informatique documentaire, appui à la recherche) ou sous la responsabilité de la direction quand les missions sont très transversales

86 SERVET, Mathilde, 2009. Bibliothèques troisième lieu (Les). Mémoire d'étude. Villeurbanne. [Consulté le 20 décembre 2016]. Mémoire DCB. Disponible à l'adresse : <http://www.enssib.fr/bibliotheque-numerique/notices/21206-les-bibliotheques-troisieme-lieu>

87 Acronyme utilisé pour désigner les géants des industries du numérique en général, GAFA désigne spécifiquement Google, Amazon, Facebook et Apple

88 SENNETT, Richard, 2000. Le Travail sans qualités: Les conséquences humaines de la flexibilité. Paris : Albin Michel.

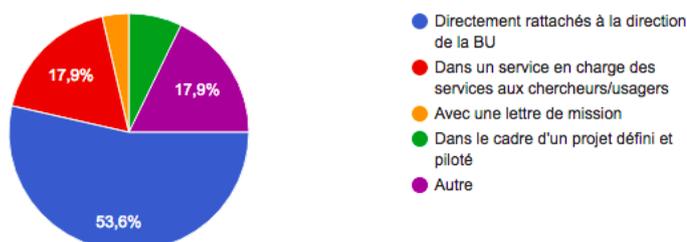
89 SENNETT, Richard, 2000. Le Travail sans qualités: Les conséquences humaines de la flexibilité. Paris : Albin Michel. p.56

90 DUPUY, François, 2013. Lost in management: la vie quotidienne des entreprises au XXIe siècle. Paris : Éd. du Seuil. Points,

Quels changements pour le travail en bibliothèque ?

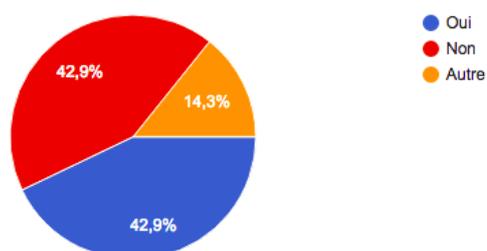
(pilotage et informatique décisionnelle, collecte automatique des données). La proportion importante de personnes travaillant sans que l'intitulé de leur poste soit particulièrement explicite peut laisser penser à une relative souplesse de l'organisation.

Dans quels cadres ces personnels travaillent-ils ? (28 réponses)



Les fonctions de ces personnels de la bibliothèque en lien avec ces projets sont-elles explicites dans l'organigramme de votre établissement ?

(28 réponses)



Un des entretiens explique ainsi le rapport à la gestion de projet dans sa structure :

« le projet avec groupes de travail, comme on l'apprend à l'Enssib, peut être très lourd à gérer et à mener. La norme ISO 9001 est moins contraignante. » On comprend ainsi que les établissements ne disposent pas des ressources nécessaires pour mener des projets dans leur forme canonique, mais qu'ils en mènent sous des formes allégées, plus souple, ici dans le cadre des principes de la démarche qualité. « On expérimente même si c'est pas une offre bien maîtrisée (...) La démarche qualité gère le risque et donc l'incertitude. Ce sont plus les habitudes de travail en bibliothèque avec les groupes projets qui constituent un obstacle. »

Ailleurs, c'est la refonte de l'organigramme qui fut le projet lui-même et les sous-groupes étudient de nouvelles missions etc. Dans les établissements, on semble globalement revenu de ce mode de travail.

«Fonctionner « en mode projet » (ah ! délice du langage managérial), c'est décider que pour composer un produit ou une «solution», il faut réunir de façon transversale des acteurs qui jusque-là travaillaient - en paix ! - dans leurs «silos», les « métiers » comme disent les Français. En somme, on a «

dé-verticalisé » un certain nombre et on les fait travailler de façon horizontale, sous la direction (le leadership) d'un «chef de projet». Et on compte effectivement sur les capacités managériales de ce dernier pour mener à bien ce projet. Cela traduit une paresse intellectuelle qui envahit aujourd'hui la vie des entreprises et qui est sans doute la contrepartie de la pénibilité croissante du travail. »⁹¹

On retrouve le mode projet dans les structures mutualisées ou support en dehors des BU. On sort non pas de son unité mais de son organisation. L'organisation est lourde mais les protections au sein de la structure demeurent, et sont même doublées d'une dé-contextualisation du projet qui s'opère hors sol. Cela favorise notamment des prises de position favorables au projet mais qui dans l'établissement n'auraient pu être adoptées aussi facilement du fait de tension et de facteurs de risques. Dégagées de cela, les personnes travaillant sur des projets hors de leur structure peuvent penser uniquement au bien du projet.

Par la permanence des unités organisationnelles qui protègent les personnes et les métiers, et le plus souvent par le maintien de la chaîne de traitement du document comme flux fondamental de la gestion du travail, les bibliothèques intègrent de nouveaux silos de données dans des organigrammes renouvelés, enrichis de dimensions « services aux usagers ». Selon les établissements, on observe cependant une ligne de partage entre une organisation légère, agile, laissant une large part à l'initiative et l'expérimentation encadrée par des outils permettant un certain contrôle (démarche qualité et méthode agile) d'une part, et d'autre part, des organisations plus verticales, plus pilotées, structurées en râteau et un organigramme très présent. Ces dernières semblent les plus nombreuses si on analyse le questionnaire où plus de 50% des personnes en charge des projets sont directement rattachées à la direction. Mais il ne permet que difficilement de mesurer la part d'initiative, de liberté d'expérimentation dont les agents disposent. Cette proportion peut aussi bien illustrer le souhait de préserver cette initiative du pouvoir des unités organisationnelles et positionner les agents pour leur donner plus de liberté d'action.

On peut cependant constater deux forces en présence dans les bibliothèques particulièrement actives sur ces questions de données :

- Celle encourageant la prise d'initiative peu maîtrisée ;
- Celle privilégiant la seule mise en place de processus particulièrement maîtrisés, auquel la chaîne de traitement du document sait globalement répondre.

« Ce mode de fonctionnement fut initialement la réponse à une volonté d'«organisation scientifique du travail» (Taylor, Fayol) qui permettait de définir des tâches claires, accomplies de façon planifiée par des individus interchangeables. Comme je l'ai suggéré dans le chapitre précédent, la dépendance vis-à-vis de l'aléa humain s'en trouve réduite au minimum et la déviance n'est ni acceptable ni acceptée, car elle exprime un désaccord avec la « science » relevant soit d'une perversion morale soit de troubles psychologiques. C'est ainsi que furent posées les bases de la production de masse⁹². »

⁹¹ DUPUY, François, 2013. Lost in management: la vie quotidienne des entreprises au XXIe siècle. Paris : Éd. du Seuil. Points

⁹² DUPUY, François, 2013. Lost in management: la vie quotidienne des entreprises au XXIe siècle. Paris : Éd. du Seuil. Points

Une nouvelle étape dans les catalogues est en train de se dessiner. Le catalogue n'y est plus le centre de gravité de l'activité et des services. Ce sont les données et c'est autour d'elles qu'on construit les services. Si le catalogue n'est plus le centre c'est aussi qu'il s'ouvre,

« s'ancre dans ce réseau global d'information qu'est le web de données. Cependant, à ce jour, la plupart de ces identifiants sont consolidés à partir de processus de production qui se poursuivent en parallèle. On est encore loin d'un système unifié et réparti qui distribuerait la responsabilité de la production à des institutions distinctes en fonction des types d'entités⁹³. »

La question qui se pose alors est de savoir si à chaque silo de données correspond un service, à chaque contrainte technique sur la structure des données, un autre silo pour la contourner, au risque d'atomiser l'organisation des données et l'organisation du travail. S'il est trop tôt pour répondre à cette question, certains chemins se dessinent. Le choix fait dans l'enseignement supérieur de travailler à la FRBérisation des catalogues en s'appuyant sur des groupes de travail répartis sur le territoire dont les membres sont issus d'établissements volontaires, afin de piloter collectivement une solution partagée et externe aux SIGB de chaque établissement montre un double schéma déjà observé :

- le besoin de mutualiser les compétences, les expériences, le besoin de sortir des schémas d'organisation de chaque établissement pour identifier les questions structurelles et les anomalies conjoncturelles, transcender l'organisation locale
- créer de nouveaux silos de données avec un modèle apportant tous les avantages du web sémantique sans pour autant déstabiliser les fonctionnements établis de la chaîne de traitement du document dans chaque établissement. En créant un silo mutualisé à coté des SIGB, on favorise les nouvelles pratiques sans casser les anciennes. Plus tard, quand les nouveaux outils seront mieux maîtrisés, les processus de travail mieux définis, on pourra alors étudier les modalités éventuelles d'intégration de ces outils dans les processus de travail de chaque établissement. Ainsi, on mutualise les coûts et les compétences nécessaires, on externalise hors de la BU les impacts organisationnels, dans un silo indépendant du catalogue, protégeant l'organisation de celle-ci.

Les bibliothèques universitaires ont-elles pour autant changé de paradigme et quitté le monde tayloriste de Paul Otlet ?

Interrogé sur l'importance de la chaîne de traitement du document, après un long exposé sur le travail sur les données de la recherche, et d'autres démarches totalement hors cadre par rapport à cette chaîne, mon interlocuteur la perçoit comme encore « structurante » pour le travail de tous. En creusant, la plupart des nouveaux services échappent à la chaîne de traitement du document sauf la Gestion électronique de Document -GED- (intégration de mémoires, publications, doctorats de l'université) qui rejoint la chaîne documentaire. Le service public en salle vient compléter l'argumentaire du caractère structurant de la chaîne documentaire.

93 BERMÈS, Emmanuelle (éd.), 2017. L'avenir des catalogues. Le cercle de la librairie. à paraître

« Si on ne la connaît pas on ne peut pas faire correctement de renseignement en salle. »

Dans les représentations professionnelles, la chaîne de traitement et le service public en salle constituent deux piliers essentiels, qui ne s'opposent pas forcément aux évolutions des autres missions. Au contraire, la recherche d'un certain équilibre entre les deux est ici observée. Il s'agit de maintenir un bon niveau de qualité dans le signalement pour garantir la qualité d'accès aux collections en salle, mais sans aller trop loin dans la qualité du catalogage permettant de libérer des ressources sur de nouvelles priorités.

Si le centre de gravité s'est déplacé, si le paradigme de l'établissement a changé, les témoignages ne l'attestent pas forcément. Une enquête qualitative à petite échelle ne suffit pas toujours à éclairer ces évolutions. Entre rémanence de la perception malgré des changements significatifs ou continuité des cultures professionnelles et des organisations, la nuance est mince. Ce sont des éléments externes qui peuvent l'attester. Ainsi, le même interlocuteur affirme

« moi je serais plus d'avis d'aller au devant des besoins, d'expérimenter, même si c'est une offre peu maîtrisée ».

Sortir d'un processus très maîtrisé comme la chaîne de traitement du document pour proposer, expérimenter, essayer témoigne en fait qu'on est ici bien sorti d'une organisation tayloriste du travail et que le souhait est une organisation apprenante et agile qui s'autorise à essayer, ajuster, rater, de recommencer, d'améliorer, à petite échelle mais en production, en situation réelle, avant de s'étendre progressivement.

Ailleurs, au contraire, la chaîne de traitement reste le centre de gravité de l'organisation du travail. Le choix a été fait d'intégrer au maximum l'ensemble des ressources documentaires dans le SIGB quitte à « tordre un peu son fonctionnement », à adapter la logique d'ensemble à chaque filière d'entrée.

Dans le premier cas, l'organigramme n'avait pas été modifié depuis 7 ou 8 ans. Dans le second cas, il vient d'être profondément modifié après un long processus de travail avec l'ensemble des équipes.

- Dans le premier, cas

« les processus ne sont pas liés à un seul service. Ca pose pas mal de question mais nos outils sont tout de suite partagés et donc c'est très rapidement transversal ».

Ici la logique d'intégration s'appuie sur la transversalité et l'adaptation aux besoins des ressources indépendamment de leur service de rattachement. La logique territoriale des unités métiers semble ici totalement dépassée.

- Les changements étaient dans le second cas structurels, dans une logique verticale d'intégration forte de l'ensemble des fonctions.

On mesure ici le poids des représentations et l'écart entre les situations vécues et la persistance de représentation traditionnelle d'une part et le poids que ces représentations peuvent avoir dans les choix et les orientations organisationnelles pour l'accompagnement des stratégies orientées usagers, d'autre part.

3.2. STATUTS, CULTURE METIER ET REPRESENTATIONS PROFESSIONNELLES

Le caractère nouveau de certains métiers liés aux données peut laisser penser que, le recours à des contractuels ou des corps extérieurs au monde des bibliothèques devrait être sinon majoritaire au moins fréquent. Les deux volets de cette étude se sont donc intéressés aux statuts, aux corps des personnels impliqués (pour les Humanités Numériques, pour les données de la recherche ou le FRBR ou l'informatique décisionnelle). Les entretiens ne m'ont pas permis de dégager de tendances particulières. C'est une des raisons qui m'ont incité à compléter les entretiens par une enquête quantitative, dont les résultats sont très différents de ceux qu'on aurait pu penser.

Les personnels de la bibliothèque travaillant sur les projets évoqués en section 1 sont-ils :
(28 réponses)

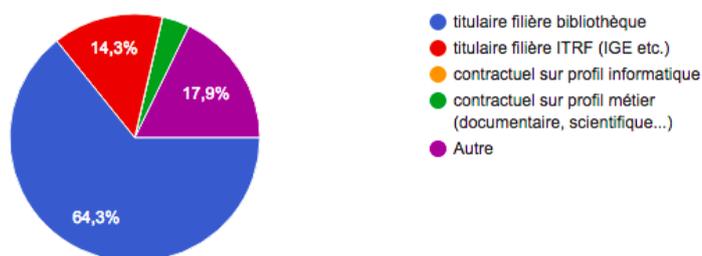


Figure 1

Près des deux tiers des personnels travaillant sur des projets liés aux données sont titulaires de la filière bibliothèque. Le second pôle (5 établissements) est constitué des différentes filières hors titulaires des bibliothèques (titulaire et contractuel sur profil métier). Aucun cas n'est remonté de recrutement sur profil informatique. Il en ressort une vraie conscience des enjeux profondément métier des questions de données. Les entretiens confirment appuient cette forte tendance. Les DSI des universités sont des appuis sur des questions d'infrastructure (bande passante, stockage, serveurs) ou d'acquisition de logiciels experts (informatique décisionnelle). Ils peuvent assurer sur certains dossiers une maîtrise d'ouvrage mais la maîtrise d'œuvre reste côté métier. En revanche, plusieurs établissements ont dans le SCD un ou deux informaticiens qui peuvent aider à certaines tâches (maintenance, petits développements pour l'essentiel) mais le pilotage reste au métier.

Ainsi le métier apparaît très présent sur les sujets liés aux données, mais comment est-il appréhendé ?

Dans le domaine des humanités numériques, les projets qui réussissent le mieux

« c'est quand on a toutes les compétences autour de la table et que chacun est capable de faire 50% du chemin vers l'autre »

c'est à dire savoir penser un peu comme un informaticien ou un peu comme un chercheur, tout en restant bibliothécaire et inversement. La réussite passe par l'hybridation des compétences de chacun.

« On repère les gens qui sauront dépasser leur compétence de base pour faire du développement informatique par exemple. (...) on est vigilant sur les projets où des acteurs ne sauront pas dépasser leur compétence d'origine (...) ça n'est pas une question de compétence mais une question de formation et une question institutionnelle. Dans les pays anglo-saxons les « curators » sont hybrides. En France, on a quelques formations hybrides comme l'Ecole des Chartes et l'Enssib mais ça ne suffit pas toujours. « J'étais surpris de voir récemment dans un projet ANR alors que des bibliothécaires étaient dans le projet, ils se réunissaient ensemble à côté du projet : on a des questions de bibliothécaires à régler. »

Ainsi, il faut non seulement une hybridation, être capable de penser un peu comme les autres interlocuteurs mais aussi laisser les autres venir à soi. Pour Huma-Num, c'est une question de formation, initiale et continue. Il y a trop peu de formations favorisant l'hybridation des compétences en France ainsi qu'une culture institutionnelle qui peut encourager des visions métiers segmentées. En effet, le constat selon lequel ce sont des bibliothécaires qui assurent les missions liées aux enjeux des données, y compris pour les données de la recherche, laisse penser que les bibliothèques de ce point de vue fonctionnent en circuit fermé sur un créneau où l'hybridation serait un bénéfice. Pourquoi ne pas placer un enseignant-chercheur dans la bibliothèque pour travailler sur le projet de gestion des données de la recherche ?

Malgré cela Huma-Num envisage de donner des formations avec une certaine prudence.

« On l'a fait pendant quelques années et au bout de deux trois ans on revoyait les mêmes personnes. Il n'y a pas encore assez de personnes impliquées dans les humanités numériques et on avait fait le tour. Alors on a arrêté mais il faudra sans doute qu'on reprenne dans quelques temps. »

Cette question reste « un enjeu important » mais à ce stade,

« on préfère s'appuyer sur les URFIST ».

Le choix des URFIST, espace hybride par nature est de ce point de vue très éloquent. Cependant, l'entretien insiste plus sur une culture de l'hybridation que sur des lieux d'hybridation, et encourage à mieux utiliser les URFIST.

En interne, la formation est en revanche un souci permanent :

« on laisse 20% de temps à chacun pour se former, du travail perso, ramener des projets. »

Cela implique une appétence pour cela, il faut accepter d'être en auto-formation permanente, de sortir de la routine. C'est un des facteurs qui a incité à recruter des personnels de même grade dans l'équipe, et de ne plus recruter de technicien.

Ce regard extérieur aux bibliothèques offre plusieurs perspectives :

- Une culture métier importante, nécessaire aux projets en humanités numériques, mais parfois trop fermée sur elle-même, pas toujours ouverte, pas assez dans le partage de ses questions et de ses réponses. La composition des

personnels des bibliothèques largement recrutés parmi les personnels du corps des bibliothèques notamment sur les données de la recherche en est une illustration.

- L'hybridation des compétences nécessaire aux domaines numériques s'appuie sur des organisations du travail et sur des recrutements adaptés, que les bibliothèques par leur mission et leur activité historique ne peuvent appliquer. Cet état de fait et la permanence d'une activité de production, alimentent sans doute la forte polarité métier des recrutements. Comment maintenir une activité essentielle et constitutive de la profession (acquérir, signaler, communiquer des collections) avec des effectifs déjà contraints si on recrute en plus des personnels en dehors de ces compétences métiers ?

La perception interne de ces évolutions a été plus délicate à saisir, dans la mesure où les personnes rencontrées sont impliquées dans la gestion des données. Il est possible cependant de souligner quelques points importants.

Nous évoquons plus haut qu'un établissement assez investi dans des sujets innovants sur les données, avait répondu que la chaîne de traitement documentaire restait fondamentale dans le travail des équipes. Cela renvoie aussi à la perception de cet interlocuteur qui bien que très impliqué sur de nouvelles questions perçoit la chaîne documentaire comme un élément structurant de l'organisation de son établissement. La représentation de la bibliothèque même pour les personnels en pointe, impliqués dans ces sujets, n'est pas fondamentalement modifiée. Si les outils changent, si les formats s'adaptent, les compétences s'enrichissent, les modalités de travail s'assouplissent, se mutualisent, à l'inverse la structure, elle, semble perdurer au moins dans ses représentations. Le fait que ces personnes soient intégrées à des services préexistants, ou rattachés directement à la direction et que l'évolution des organigrammes n'est pas envisagée à ce stade conforte certainement ce sentiment de permanence des structures métiers, malgré les évolutions technologiques. Pour autant, l'appréhension des fonctions ne semble pas simple, structurée et établie puisque les organigrammes en témoignent de façon peu explicite. Il est probable que ces changements importants ne soient pas perçus comme tels.

Est-il envisagé au sein de votre établissement de modifier l'organigramme afin d'améliorer la gestion des données ?

(28 réponses)

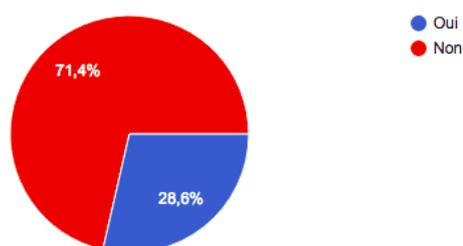


Figure 2

Les fonctions de ces personnels de la bibliothèque en lien avec ces projets sont-elles explicites dans l'organigramme de votre établissement ?
(28 réponses)

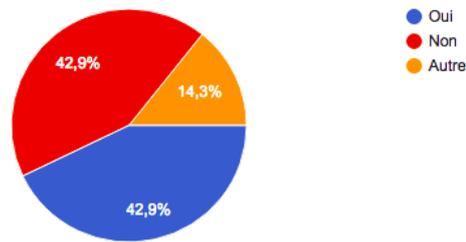


Figure 3

Pour autant, plusieurs établissements ont organisé ou organisent un accompagnement au changement. Ici dans le cadre d'une fusion d'université, ailleurs pour accompagner les équipes dans une démarche globale où le changement devient une constante de l'organisation. Il y a donc une certaine contradiction à ne pas percevoir de sentiment de changement, quand les organisations accompagnent de façon structurée les équipes dans le changement. Le changement serai-il ici un mode de management comme dans d'autres structures, et non pas une phase de transition ?

Ce paradoxe pourrait cependant s'expliquer. Si ces changements de modèles, de nature des données, de formats etc. n'étaient pas rattachés au métier, au quotidien du travail en bibliothèque ? En effet, certains travaux sont eux en production, mais par l'automatisation (informatique décisionnelle) ou parce que le silo de données concerné ne communique pas ou communique peu avec les étapes de la chaîne documentaire (ezPAARSE, PREVU, HAL ou les archives institutionnelles) alors

« c'est un travail que les bibliothécaires ont du mal à rattacher à leur métier. »

L'automatisation, la transparence permettent un changement invisible ou par « petites touches » qui ne modifie pas le reste de l'organisation. L'organisation n'a pas besoin d'être modifiée tant qu'un service peut accueillir une nouvelle activité, qui occupe une personne, parfois deux ou trois, rarement plus. Ainsi, le sentiment de permanence peut-il se construire et s'expliquer. Cependant, une enquête dénombrant sur 10 ou 15 ans la part des ressources affectées au traitement documentaire montrerait sans doute une réalité assez différente que celle perçue ici.

Ainsi, de nombreux établissements qui avaient au début des années 2000 confié à une personne la gestion des collections électroniques ont dû, quelques années plus tard, revenir sur ce choix, du fait de la disproportion de budget et de personnels entre les deux pôles (ressources papier et électroniques) : quand une personne doit gérer sous sa seule responsabilité plus de 50% du budget d'acquisition et 50% du personnel se partage moins de 50% du budget l'organisation semble ne pas avoir changé mais elle est pourtant profondément modifiée.

Enfin, plusieurs des projets de silos de données évoqués sont soit à l'état de projet et donc n'affectent pas l'activité en production de la bibliothèque, soit gérés dans

des structures interprofessionnelles comme pour ezPAARSE ou le SGBM et les groupes de travail sur le FRBR. Ainsi, le travail, décontextualisé est difficilement perçu comme impliquant directement les établissements. En travaillant ainsi, les questions d'organisation sont déportées et ne pèsent pas localement, mais minimisent dans les organisations les perceptions des changements en cours. Pour un projet comme la FRBÉRisation des catalogues, et SGBM, seules quelques personnes sont impliquées à l'échelle de la profession. Si, 46 établissements d'enseignement supérieur ont signé le groupement de commande SGBM, seulement 9 sites pilotes sont concernés en tant qu'organisation et dans chaque établissement, une ou deux personnes tout au plus sont impliquées. Pour un projet de cette nature, les changements ne se verront qu'une fois le système déployé. Si le modèle présenté est effectif, beaucoup de traitements se feront de façon transparente.

« L'intégration du SGBM aux systèmes d'information des universités se fera grâce à des API (Application Programming Interface) et des webservice⁹⁴. »

Ainsi le travail devrait en apparence être peu modifié, mais en profondeur radicalement différent, car mutualisé, sur des modèles de données très différents, ouvert sur le web etc. Le travail sera comparable, mais l'organisation encadrant ce travail devrait être modifiée en profondeur. (Notons ici, que le modèle prévu est une plateforme mutualisée, recourant fortement aux technologies des nuages pouvant poser les questions juridiques, et de désintermédiation évoquées plus haut.)

On peut comprendre alors ces différences de perception et le sentiment que tout reste inchangé quand pourtant tout a bougé, comme une variante au monologue du Comte Salina dans « Le Guépard⁹⁵ ».

94 GOURET Gaelenn, « Les BU engagées dans la mutualisation de leur système d'information », mercredi 22 juin 2016 pages 10 à 12 - Essentiel n° 71749, <http://education.newstank.fr/http://education.newstank.fr/fr/article/printable/71749/bu-engage> (consulté le 12/12/2016)

95 LAMPEDUSA, Giuseppe T. di, 2006. Le Guépard. Paris : Editions du Seuil.

CONCLUSION

Paul Otlet en pleine révolution tayloriste, établit un lien structurant entre la division des tâches et l'amélioration des services documentaires. Depuis lors, les bibliothèques comme les services de documentation ont systématisé, normalisé, « procéduré » leurs activités. Ces dernières sont découpées en étapes successives, reposant sur une division importante du travail et des compétences, mais ce processus vise à alimenter un silo unique d'information, le catalogue. L'informatisation des processus métiers en bibliothèque n'a rien changé à cette organisation, renforçant parfois même le caractère central du catalogue.

La façon dont on gère, inscrit, transforme ou traite l'information s'accompagne de nouvelles façons d'apprendre ou de travailler, sous l'effet des technologies de l'information. Ces nouvelles façons de faire ne sont pas à considérer comme l'indique Michel Serres⁹⁶ dans une vision optimiste des technologies, pour ce qu'elles ne font plus, mais pour ce qu'elles permettent de faire en plus. Michel Serres rappelle que les outils libèrent le cerveau de certaines contraintes. S'appuyant sur les travaux d'André Leroi-Gourhan⁹⁷, il souligne l'émergence de progrès sociaux et économiques, permis par des révolutions technologiques liées à la gestion de la connaissance.

Au-delà de la question du progrès positiviste que pourraient constituer ces technologies, émerge la question soulevée par J.-M. Salaün⁹⁸: quelles responsabilités accompagnent les acteurs du pentagone de l'information? A l'objectivité journalistique affichée jusque dans les années 90 succèdent la neutralité ou l'absence de mauvaises intentions des industriels de l'information. Google très tôt avec son « Don't be evil » mesure la responsabilité qui est la sienne comme pourvoyeur d'information sur le web. La question de la neutralité technologique ou algorithmique alimente toute une part des débats sur les technologies de l'information, depuis l'invention de PageRank, en passant par Wikipédia, jusqu'à la théorie de la bulle informationnelle des réseaux sociaux pour expliquer la surprise des résultats de l'élection américaine de 2016. « Le document, lui-même produit à la demande, n'est plus bien stable, mais un service labile » souligne J.-M. Salaün, « en phase avec la bibliothèque et avec l'évolution générale de l'économie vers le service ».

Or, au-delà de la question éthique posée, ces algorithmes sont à considérer ici comme les machines-outils de l'industrialisation des processus de travail. Ils achèvent ainsi le rêve de Paul Otlet d'intégration et d'automatisation d'un système documentaire par une division des tâches. Mais ce rêve n'est pas assuré dans un seul système documentaire, mais dans différents systèmes, concurrents, en tension les uns par rapport aux autres.

Ainsi, dans les bibliothèques, la division des tâches ne permet pas à elle seule de garantir la neutralité de l'accès à l'information. L'organisation des données, et l'organisation du travail créent des biais, des intentions, des modèles que

⁹⁶ Michel Serres, Philosophie magazine n°62, Entretien consultable : <http://www.philomag.com/les-videos/michel-serres-et-bernard-stiegler-moteurs-de-recherche-3244>, (consulté le 10/12/2016)

⁹⁷ LEROI-GOURHAN, André, 1964. Le Geste et la Parole, Albin Michel. Paris

⁹⁸ SALAÜN, Jean-Michel, 2012. Vu, lu, su. Paris, 2012, La Découverte.

Dominique Cardon nous a aidé à expliciter. L'architecture des données crée du sens. La bibliothèque induit un rapport biaisé entre les données et les usagers. Se positionner par rapport aux données est nécessaire, les biais sont donc irréductibles. Ces positions doivent faire l'objet de choix, répondre à des stratégies sur les services, aux besoins des usagers, aux souhaits de suivi et de pilotage des tutelles.

Si la permanence des systèmes mis en place dans les dernières décennies du XXe siècle témoigne de l'efficacité d'un modèle documentaire et organisationnel de la bibliothèque, on observe également le développement de nouveaux services rendus nécessaires par les nouveaux usages et par la numérisation. Ceux-ci s'articulent avec une organisation et une culture du travail très différentes, où la souplesse, l'expérimentation de la culture start-up, de l'organisation apprenante sont aussi indispensables que les capacités d'adaptation et de formation. On y cherche moins la maîtrise et l'efficacité d'un processus parfaitement rodé et préparé pendant des mois pour rendre un service attendu, que la capacité d'expérimenter, d'essayer, pour offrir des outils peut être moins maîtrisés mais plus en phase avec des besoins labiles.

Cette coexistence de fait est rarement observée dans les bibliothèques, où elle apparaît délicate à mettre en place du fait d'un cadre administratif et d'une structuration des cadres d'emplois hérités du modèle tayloriste du métier. Elle existe cependant, par la fréquentation et la coopération des établissements avec des structures plus légères: URFIST, Human-num ou groupes projet dans les organisations interprofessionnels. En mettant l'agilité et la souplesse dans des structures externes, les bibliothèques évitent d'affronter la délicate compatibilité entre les deux modes de management dans un cadre administratif, statutaire et budgétaire qui laisse peu de marge de manœuvre. Cette externalisation est une forme de régulation douce qui permet de laisser à certains des espaces d'innovation dans un cadre transversal et faiblement hiérarchisé, et les espaces protecteurs d'un travail en silo à la plupart des agents. C'est une stratégie de petits pas qui s'observe, petits pas vers les chercheurs pour gérer leurs données, petits pas vers d'autres modes d'organisations, petits pas vers les Humanités Numériques mais presque uniquement via les collections et non pas via les missions d'appui. Mais la vitesse d'évolution qui caractérise le numérique laisse peu de temps. Aussi cette stratégie rassurante pour l'organisation favorise-t-elle des risques importants de plateformes de certains services, par le recours volontaire ou non à des intermédiaires qui captent les données et une part des responsabilités qui sont initialement celles des bibliothèques.

Le choix qui s'offre aux bibliothèques est donc délicat : comment maintenir une cohérence d'ensemble quand il faut choisir entre l'efficacité des silos pour assurer un service et proposer une organisation stable et rassurante du travail d'une part et des modalités de management, d'autre part favorisant l'innovation dans un cadre moins contraignant mais plus flottant ? Pour autant, les règles administratives, les cadres d'emploi, la culture métier restent globalement inchangés ; les effectifs globalement contraints laissent peu de marge de manœuvre et offrent de moins en moins de souplesse à l'organisation existante pour évoluer. L'hybridation souhaitable des compétences, est d'autant plus délicate qu'en maintenant les services traditionnels le temps passé à se former et à travailler sur des sujets

innovants est fortement restreint. Et pourtant sauf à trouver les ressorts permettant d'évoluer, il est fort probable que les innovations évoquées dans ce travail s'ancrent dans d'autres structures que les BU si celles-ci n'adaptent pas leurs organisations pour les accueillir. C'est sans doute en actionnant les leviers des projets de recherche (LABEX, BSN etc.) et en s'associant plus systématiquement à ces dispositifs et aux chercheurs que les BU, en ouvrant les recrutements aux autres emplois d'appui à la recherche (IGE, ITRF) trouveront des marges de manœuvre et les leviers de l'hybridation.

BIBLIOGRAPHIE

Publications officielles et réglementaires

COMMISSION EUROPÉENNE et DIRECTION GÉNÉRALE DE LA RECHERCHE ET DE L'INNOVATION, 2015. Lignes directrices pour la gestion des données dans Horizon 2020 [en ligne]. 2015. Disponible à l'adresse : http://openaccess.inist.fr/IMG/pdf/14086_lignes_directrices_la_horizon_2020_tr_fr_version-oct2014.pdf

Dossiers législatifs - LOI n° 2016-1321 du 7 octobre 2016 pour une République numérique | Legifrance, 2016. [en ligne]. LOI n° 2016-1321. [Consulté le 18 décembre 2016]. Disponible à l'adresse : https://www.legifrance.gouv.fr/affichLoiPubliee.do;jsessionid=836E747AF734FC6DC229303D48E2ADF2.tpdila20v_1?idDocument=JORFDOLE000031589829&type=contenu&id=2&typeLoi=&legislature=14

LOI n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, 2016. [en ligne]. Loi n° 78-17. [Consulté le 18 décembre 2016]. Disponible à l'adresse : <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000886460>

LOI n° 2016-925 du 7 juillet 2016 relative à la liberté de la création, à l'architecture et au patrimoine, 2016. [Consulté le 18 décembre 2016]. Disponible à l'adresse : <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000032854341&categorieLien=id>

LOI n° 2016-1321 du 7 octobre 2016 pour une République numérique, 2016. [Consulté le 18 décembre 2016]. Disponible à l'adresse : <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000033202746&categorieLien=id>.

Note d'information du 5 avril 2016 relative à l'informatique dans les nuages, 2016. [Consulté le 18 décembre 2016]. Disponible à l'adresse : <http://www.archivesdefrance.culture.gouv.fr/static/9151>.

Publications par des organisations publiques ou professionnelles

ADBU, 2012. Introduction à l'enquête 2012 «Organisation fonctionnelle des équipes». ADBU. [en ligne]. 2012. Disponible à l'adresse : <http://adbu.fr/wp-content/uploads/2013/11/ADBU-Metiers.-Enquete-2012.-Introduction.pdf>

ADBU, 2016a. *Text and data mining in higher education and public research An analysis of case studies from the United Kingdom and France* [en ligne]. [Consulté le 18 décembre 2016]. Disponible à l'adresse : <http://adbu.fr/competplug/uploads/2016/12/TDM-in-Public-Research-Revised-15-Dec-16.pdf>

ADBU, 2016b. *La fouille de textes et de données dans l'enseignement supérieur et la recherche publique Une analyse d'études de cas menées au Royaume-Uni et en France* [en ligne]. [Consulté le 18 décembre 2016]. Disponible à l'adresse :

http://adbu.fr/competplug/uploads/2016/12/v9-Designed-Exec-Summary-ADBU-8pp_fra_final.pdf

BNF, 2016. *Les publics - Rapport d'activité 2015 de la BnF* [en ligne]. Paris. [Consulté le 18 décembre 2016]. Disponible à l'adresse : <http://webapp.bnf.fr/rapport/publics.html>

INIST, 2016a. Traduction française des guides pour Horizon 2020 - Libre accès à l'information scientifique et technique. [en ligne]. 2016. [Consulté le 18 décembre 2016]. Disponible à l'adresse : <http://openaccess.inist.fr/?Traduction-francaise-des-guides-795>

INIST, 2016b. *Préserver les données de la recherche à l'ère du Big Data. CNRS Le journal* [en ligne]. 2016. [Consulté le 14 septembre 2016]. Disponible à l'adresse : <https://lejournal.cnrs.fr/articles/preserver-les-donnees-de-la-recherche-a-lere-du-big-data>

INSPECTION GÉNÉRALE DES BIBLIOTHÈQUES, 2013. *Quels emplois pour les bibliothèques? Etat des lieux et perspectives - Livre et Lecture - Ministère de la Culture et de la Communication* [en ligne]. [Consulté le 11 juillet 2016]. Disponible à l'adresse : <http://www.culturecommunication.gouv.fr/Politiques-ministerielles/Livre-et->

ISMER, 2016. *Politique de gestion des données océanographiques de l'Institut des sciences de la mer de Rimouski*. 2016. [Consulté le 11 décembre 2016]. Disponible à l'adresse : <http://www.ismer.ca/Politique-de-gestion-des-donnees>

LERU, 2013. *LERU Roadmap for Research Data* [en ligne]. [Consulté le 19 décembre 2016]. Disponible à l'adresse : http://www.leru.org/files/publications/AP14_LERU_Roadmap_for_Research_data_final.pdf

UNIVERSITY COLLEGE CORK, 2014. *Data Management Policy*. 2014. [Consulté le 19 décembre 2016]. Disponible à l'adresse : <http://www.ucc.ie/en/media/support/itpolicies/policies/DataManagementPolicy.pdf>

Les enjeux du Big Data et du numérique

BABINET, Gilles et ORSENNA, Erik, 2015. *Big Data, penser l'homme et le monde autrement*. Paris : le Passeur éditeur. ISBN 978-2-36890-260-8.

BELOT, Laure, 2015. *La déconnexion des élites: comment Internet dérange l'ordre établi*. Paris : Les Arènes. Document. ISBN 978-2-35204-364-5.

BRASSEUR, Christophe, 2013. *Enjeux et usages du Big Data*. HERMÈS / LAVOISIER. ISBN 978-2-7462-4520-4.

BRUCHEZ, Rudi, 2015. *Les bases de données NoSQL et le Big Data*. [en ligne]. 2015. [Consulté le 19 décembre 2016]. Disponible à l'adresse : [/Informatique/Livre/les-bases-de-donnees-nosql-et-le-big-data-9782212141559](http://www.informatique.fr/Livre/les-bases-de-donnees-nosql-et-le-big-data-9782212141559)

CARDON, Dominique, 2015. *A quoi rêvent les algorithmes: nos vies à l'heure des big data*. Paris : Seuil. La République des idées. ISBN 978-2-02-127996-2.

CASEAU, Yves, 2015. *Big data: un changement de paradigme peut en cacher un autre opportunités et menaces liées à l'émergence de nouveaux écosystèmes rapport de l'Académie des technologies de la Commission TIC*. Les Ulis : EDP sciences. Avis et rapports de l'Académie. ISBN 978-2-7598-1780-1.

Choisissez mieux vos dataviz avec The Data Visualisation Catalogue | Outils Froids, 2016. [en ligne]. [Consulté le 14 septembre 2016]. Disponible à l'adresse : <http://www.outilsfroids.net/2016/04/choisissez-mieux-vos-datavisualisations-avec-the-data-visualisation-catalogue/>

COINTOT, Jean-Charles et EYCHENNE, Yves, 2014. *La révolution big data: les données au cœur de la transformation de l'entreprise*. Paris : Dunod. Stratégies et management. ISBN 978-2-10-071142-0.

DELORT PIERRE, 2015. *Le big data*. Paris : Presse universitaire de France.

GREENAN, Nathalie, GUILLEMOT, Danièle et KOCOGLU, Yusuf, 2010. *Informatisation et changements organisationnels dans les entreprises*. Paris : La Découverte. ISBN 978-2-7071-6452-0.

FEYT, Grégoire, NOUCHER, Matthieu, 2014. *La gouvernance informationnelle, outil et enjeu stratégiques des recompositions territoriales: vers l'émergence de nouveaux référentiels géographiques? 2^eeme colloque du GIS CIST*. 2014.

IoT, robots, big data...: sur quoi travaillent les développeurs?, 2016. [en ligne]. [Consulté le 14 septembre 2016]. Disponible à l'adresse : <https://rslnmag.fr/jeunesse/iot-big-data-developpeurs-web/>

JOST, Clémence, 2016b. *Lancement de l'Observatoire des usages du big data*. *Archimag*. 24 juin 2016.

LUCCHESI, Laure, A Etalab, nous assurons une mission de service public en mode start-up!, 2016. [en ligne]. [Consulté le 14 septembre 2016]. Disponible à l'adresse : <https://rslnmag.fr/cite/laure-lucchesi-etalab-service-public-startup/>

LEJEUNE, Yannick, 2014. *Big, fast, open data: décrire, décrypter et prédire le monde l'avènement des données*. Mercuès : Éd. Fyp. Innovation. ISBN 978-2-36405-121-8.

MAYER-SCHÖNBERGER VIKTOR, CUKIER Kenneth, 2014. *Big data: la révolution des données est en marche*. Paris : Robert Laffont.

MONINO et SEDKAOUI, 2016. *Big Data, Open Data et valorisation des données*. London. Collection innovation, entrepreneuriat et gestion. ISBN 978-1-7840. Bibliogr. p. [145]-154. Notes bibliogr. Index

PUYBAREAU, Florence, 2015. *La révolution big data*. Lignes de Repères. ISBN 978-2-36609-019-2.

RIFKIN, Jeremy et SAINT-UPÉRY, Marc, 2005. *L'âge de l'accès*. Paris : La Découverte. ISBN 978-2-

SADIN, Éric, 2015. *La vie algorithmique: critique de la raison numérique*. Paris : Éditions L'Échappée. Pour en finir avec. ISBN 978-2-915830-94-1.

VERDIER, Publié par Henri, 2016. Sciences et big data: La fin de la théorie? [en ligne]. 28 mars 2016. [Consulté le 14 septembre 2016]. Disponible à l'adresse : <http://www.henriverdier.com/2016/03/sciences-et-big-data-la-fin-de-la.html>

Les enjeux du Big Data et du numérique dans les métiers documentaires

BATTISTI, Michèle, 2013. Donnée, quand tu nous tiens! *Documentaliste-Sciences de l'Information*. 9 octobre 2013. Vol. 50, n° 3, pp. 1-1.

BERMÈS, Emmanuelle, 2015. Big Data et bibliothèques,. [en ligne]. 2015. Disponible à l'adresse : <https://figoblog.org/2015/01/13/big-data-et-bibliotheques/> consulté le 12/12/2016

BERMÈS, Emmanuelle (éd.), 2017. *L'avenir des catalogues*. Le cercle de la librairie.

Bibliothèques: quand les SIGB passent à la mutualisation, 2016. Archimag [en ligne]. [Consulté le 18 décembre 2016]. Disponible à l'adresse : <http://www.archimag.com/bibliotheque-edition/2016/01/21/bibliotheque-sigb-mutualisation>

CALDERAN, Lisette, LAURENT, LOWINGER et MILLET, 2012. Le document numérique à l'heure du web de données. [Paris]. Sciences et techniques de l'information. ISBN 978-2-8436.

CALDERAN, Lisette, LAURENT, Pascale, LOWINGER, Hélène et MILLET, Jacques, 2015. Big data: nouvelles partitions de l'information actes du séminaire IST Inria, octobre 2014. Louvain-la-Neuve [Paris] : De Boeck ADBS. ISBN 978-2-8041-8915-0.

CALDERAN LISETTE, LAURENT Pascale, 2015. Big data: nouvelles partitions de l'information /. In : Actes du séminaire IST Inria. de Boeck. 2015.

CLOT NATHALIE, 2016. Loi Lemaire: Open Access et TDM. ADBU – Association des directeurs et des personnels de direction des bibliothèques universitaires [en ligne]. 2 octobre 2016. [Consulté le 18 décembre 2016]. Disponible à l'adresse : <http://adbu.fr/loi-lemaire-open-access-et-tdm/>

Data mining: la loi ne doit pas enterrer la recherche française, 2016. Les échos [en ligne]. [Consulté le 11 juillet 2016]. Disponible à l'adresse : <http://www.lesechos.fr/idees-debats/editos-analyses/021875211332-fouille-de-donnees-la-loi-ne-doit-pas-enterrer-la-recherche-francaise-1217258.php?FXSgjV5pp7wlJx1i.99>

DELHAYE, Marlène, 2013. gestion des données de recherche | Marlène's corner. [en ligne]. 2013. [Consulté le 11 juillet 2016]. Disponible à l'adresse : <https://marlenescorner.net/tag/gestion-des-donnees-de-recherche/>

GAILLARD, Rémi, 2014. De l'Open data à l'Open research data: quelle(s) politique(s) pour les données de la recherche. 2014. Mémoire DCB, Enssib.

GILLIUM, Johann et JOLY, Monique, 2016. Big data et bibliothèques. 2016. Mémoire DCB, Enssib
Bibliogr. p. 63-65.

GRANDJEAN, Martin, 2016. L'archive et la massification des données. Martin Grandjean [en ligne]. 31 mars 2016. [Consulté le 14 septembre 2016]. Disponible à l'adresse : <http://www.martingrandjean.ch/archive-donnee-bruno-bachimont/>

GROUIN et FOREST, 2012. Expérimentations et évaluations en fouille de textes. Paris. Collection systèmes d'information et organisations documentaires. ISBN 978-2-7462.

JEAN-PHILIPPE MAGUÉ, 2016. Humanités numériques, Architecture de l'information: des réponses à un changement de paradigme documentaire. *dlis.hypotheses.org*. juin 2016.

JOSSE, Isabelle, 2013. La BnF engagée dans un projet de crowdsourcing pour la conception de la plateforme Correct (Correction et enrichissement collaboratifs de textes). [en ligne]. 1 janvier 2013. [Consulté le 19 décembre 2016]. Disponible à l'adresse : <http://bbf.enssib.fr/consulter/bbf-2013-05-0037-008>

JOST, Clémence, 2016a. Open data: quels enjeux pour les bibliothèques, les archives et les musées? *Archimag*. 5 septembre 2016.

JOST, Clémence, 2016c. Alerte: des robots bibliothécaires envahissent les bibliothèques de Singapour. *Archimag*. 6 juillet 2016.

LAPÔTRE, Raphaële et VELCIN, Julien, 2015. *Faire parler les données des bibliothèques* [en ligne]. Mémoire d'étude. Mémoire DCB. Disponible à l'adresse : <http://www.enssib.fr/bibliotheque-numerique/notices/65117-faire-parler-les-donnees-des-bibliotheques-du-big-data-a-la-visualisation-de-donnees>
Bibliogr. p. 87-96

MOIREZ, Pauline, 2013. Bibliothèques, Crowdsourcing, métadonnées sociales. [en ligne]. 1 janvier 2013. [Consulté le 18 décembre 2016]. Disponible à l'adresse : <http://bbf.enssib.fr/consulter/bbf-2013-05-0032-007>

POUPEAU, Gautier, 2016. De la responsabilité humaine des algorithmes | Les petites cases. [en ligne]. 16 octobre 2016. [Consulté le 18 décembre 2016]. Disponible à l'adresse : <http://www.lespetitescases.net/de-la-responsabilite-humaine-des-algorithmes>

RECHERCHE & INNOVATION, La BnF allie sociologie et big data pour cerner ses gallicanautes, 2016. RECHERCHE & INNOVATION [en ligne].

[Consulté le 18 décembre 2016]. Disponible à l'adresse :
<https://blogrecherche.wp.mines-telecom.fr/2016/11/03/bnf-sociologie-big-data-gallica/>

SALAÜN, Jean-Michel, 2012. Vu, lu, su. Paris : La Découverte. ISBN 978-2-7071-7135-1.

VENTRESQUE, Vincent, BANAT-BERGET, Françoise et BERMÈS, Emmanuelle, 2013. Les mutations des collections numériques à l'heure du web de données [en ligne]. Mémoire d'étude. [S.l.]. Mémoire DCB. Disponible à l'adresse :
<http://www.enssib.fr/bibliotheque-numerique/notices/64112-les-mutations-des-collections-numeriques-a-l-heure-du-web-de-donnees>

SPITZ, Brad, CROUZET, Juliette, BERNIER, Claire et CULLAFFROZ-JOVER, Sandrine, 2013. Droit de l'information. *Documentaliste-Sciences de l'Information*. 9 octobre 2013. Vol. 50, n° 3, pp. 16-21.

TILLI, Nicolas, 2013. La protection des données à caractère personnel. *Documentaliste-Sciences de l'Information*. 9 octobre 2013. Vol. 50, n° 3, pp. 62-69.

Les questions métiers aujourd'hui

BRED, Olivier Klein, 2016. Révolution digitale, mutation managériale. lesechos.fr [en ligne]. 28 juillet 2016. [Consulté le 2 août 2016]. Disponible à l'adresse :
<http://www.lesechos.fr/idees-debats/editos-analyses/0211155766756-revolution-digitale-mutation-managériale-2017300.php#Xtor=AD-6000>

CAVALIER, François et POULAIN, Martine, 2015. Bibliothèques universitaires: nouveaux horizons. Paris : Éditions du Cercle de la librairie. Bibliothèques. ISBN 978-2-7654-1469-8.

COTTE, Dominique, LAHARY, Dominique, GENOVA, Françoise, GANDON, Fabien, DUFOUR, Jean, WILLAERT, Christophe, SCIANDRA, Diane, CHAILLOUX, Alexis, AUBRY, Christophe et HUOT, Charles, 2013. Les métiers, entre traditions et modernité. *Documentaliste-Sciences de l'Information*. 9 octobre 2013. Vol. 50, n° 3, pp. 42-59.

COTTE, Dominique, 2013a. Sources et ressources. *Documentaliste-Sciences de l'Information*. 9 octobre 2013. Vol. 50, n° 3, pp. 60-61.

COTTE, Dominique, 2013b. Qui (re)travaillera la donnée? *Documentaliste-Sciences de l'Information*. 9 octobre 2013. Vol. 50, n° 3, pp. 24-25.

DUPUY, François, 2013. Lost in management: la vie quotidienne des entreprises au XXIe siècle. Paris : Éd. du Seuil. Points, 719. ISBN 978-2-7578-3662-0.

Efficacité collective, numérique et nouveaux rapports sociaux, 2016. RECHERCHE & INNOVATION [en ligne]. [Consulté le 19 décembre 2016]. Disponible à l'adresse :
<https://blogrecherche.wp.mines-telecom.fr/2016/04/06/efficacite-collective-numerique-rapports/>

ENLART BELLIER, Sandra et CHARBONNIER, Olivier, 2013. À quoi ressemblera le travail demain? technologies numériques, nouvelles organisations et relations au travail. Paris : Dunod. ISBN 978-2-10-

ERTZSCHEID, Olivier, LACHAL, Jérémy et GAUCHER, Mady, 2013. Métiers et compétences. *Documentaliste-Sciences de l'Information*. 9 octobre 2013. Vol. 50, n° 3, pp. 4-9.

FAUSSURIER, Bérengère, 2016. Le SCD et son université de tutelle. [en ligne]. 1 janvier 2016. [Consulté le 18 décembre 2016]. Disponible à l'adresse : http://bbf.enssib.fr/matieres-a-penser/le-scd-et-son-universite-de-tutelle_66698

HOLOGNE, Odile, BERRIAU, Nathalie, 2013. «Politique, gouvernance et organisation». *Documentaliste-Sciences de l'Information*. 2013. Vol. 2/2013 (Vol. 50), pp. p. 38-45. DOI URL: www.cairn.info/revue-documentaliste-sciences-de-l-information-2013-2-page-38.htm. DOI: 10.3917/docs.502.0038.

HOLOGNE, Odile, 2013a. Du réseau à l'organisation apprenante: l'animation de la fonction IST. *Revue Documentaliste Sciences de l'Information*,. 2013. Vol. 50, n° (2), pp. 38-39. DOI 10.3917/docs.502.0038.

HOLOGNE, Odile, 2013b. Données de la recherche: rôle des professionnels IST à l'Inra. In : *43. congrès de l'ADBU; Le Havre (France) - (2013-09-19) / Congrès* [en ligne]. ADBU. 2013. [Consulté le 11 juillet 2016]. Disponible à l'adresse : <http://prodinra.inra.fr/?locale=fr#!ConsultNotice:207958>

HOLOGNE, Odile, 2014. Mise en perspective de l'évolution des métiers de l'IST dans un organisme de recherche. [en ligne]. 25 octobre 2014. [Consulté le 11 juillet 2016]. Disponible à l'adresse : <http://prodinra.inra.fr/?locale=fr#!ConsultNotice:210664>

HOLOGNE, Odile, 2015a. Quels services IST dans un contexte d'Open Science. [en ligne]. 27 novembre 2015. [Consulté le 11 juillet 2016]. Disponible à l'adresse : <http://prodinra.inra.fr/?locale=fr#!ConsultNotice:334446>

HOLOGNE, Odile, 2015b. Partage des données de la recherche : Politiques, Données, Publications, Compétences ... Inra. [en ligne]. 24 février 2015. [Consulté le 11 juillet 2016]. Disponible à l'adresse : <http://prodinra.inra.fr/?locale=fr#!ConsultNotice:192704>

ILLIEN, Gildas, HOLOGNE, Odile, POUYLLAU, Stéphane, ALFONSI, Gilles, TROEIRA, Jean-Pierre, DELAHOUSSE, Jean, DALBIN, Sylvie, REKOWSKI, Ursula von, AUBRY, Christophe et HUOT, Charles, 2013. Enjeux professionnels. *Documentaliste-Sciences de l'Information*. 9 octobre 2013. Vol. 50, n° 3, pp. 26-41.

BONNET, François, La modélisation des processus pour accélérer la transformation numérique François Bonnet, W4, 2016. *InformatiqueNews.fr* [en ligne]. [Consulté le 14 septembre 2016]. Disponible à l'adresse : <http://www.informatiquenews.fr/la-modelisation-des-processus-pour-accelerer-la-transformation-numerique-francois-bonnet-w4-43880>

MAXIME GRIMBERT, 2016a. La performance des entreprises serait liée à la maîtrise de leurs données. Archimag. 31 mai 2016.

MAXIME GRIMBERT, 2016b. Big Data : L'Europe plus intrépide et radicale que les États-Unis. Archimag. 6 février 2016.

MERCIER-LAURENT, Eunika, FRITZ, Olivier, COURRIER, Serge et BOULLIER, Dominique, 2013. Méthodes techniques et outils. Documentaliste-Sciences de l'Information. 9 octobre 2013. Vol. 50, n° 3, pp. 10-15.

MIRIBEL, Marielle de, ÉVANO, Brigitte, GRELET, Christophe, HAON, Sandrine, LIZÉE, Benoît, MOUCHARD, Martin et ROCHE, Julien, 2016. Diriger une bibliothèque: un nouveau leadership. Paris : Éditions du Cercle de la librairie. Collection Bibliothèques. ISBN 978-2-7654-1497-1.

« La multiplication des chefs de projet est une catastrophe managériale majeure », affirme le sociologue François Dupuy - Ressources humaines. usinenouvelle.com/ [en ligne]. 16 janvier 2015. in L'Usine Nouvelle [Consulté le 14 septembre 2016]. Disponible à l'adresse : <http://www.usinenouvelle.com/article/la-multiplication-des-chefs-de-projet-est-une-catastrophe-manageriale-majeure-affirme-le-sociologue-francois-dupuy.N307730>

PÉRALES, Christophe, 2015. Conduire le changement en bibliothèque : vers des organisations apprenantes. Villeurbanne : Enssib.

PÉRALES, Christophe, 2016. Future is now: «ubérisation» de la science et économie de plateforme | BibliOpen. [en ligne]. 2 octobre 2016. [Consulté le 18 décembre 2016]. Disponible à l'adresse : <http://blog.educpros.fr/christophe-perales/2016/10/02/future-is-now-uberisation-de-la-science-et-economie-de-plateforme/>

SENNETT, Richard, 2000. *Le Travail sans qualités: Les conséquences humaines de la flexibilité*. Paris : Albin Michel. ISBN 978-2-226-11501-0.

ANNEXES

Table des annexes

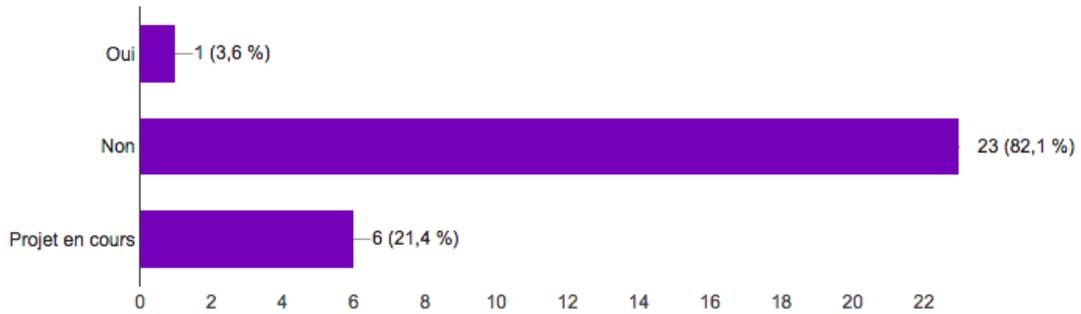
| | |
|--|----|
| ANNEXE 1 – LE QUESTIONNAIRE EN LIGNE | 83 |
| ANNEXE 2 – TABLEAU DE RESULTATS DE L'ENQUETE QUANTITATIVE | 90 |
| ANNEXE 3 – LE GUIDE D'ENTRETIEN EN PRESENTIEL | 93 |
| ANNEXE 4 – LES ORGANIGRAMMES | 97 |

ANNEXE 1 – LE QUESTIONNAIRE EN LIGNE

1 Votre université dispose-t-elle d'un entrepôt sur les données de la recherche ?

Votre université dispose-t-elle d'un entrepôt sur les données de la recherche ?

(28 réponses)

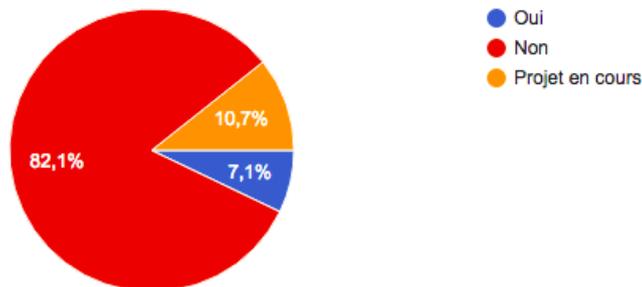


2 Si "oui" ou "projet en cours", comment la bibliothèque est associée à ce travail ?

3 Votre université a-t-elle un plan de gestion des données de la recherche ?

Votre université a-t-elle un plan de gestion des données de la recherche ?

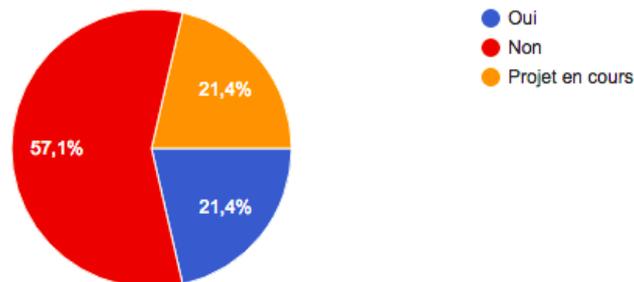
(28 réponses)



4 La bibliothèque est-elle associée au travail sur la gestion des données de la recherche ?

La bibliothèque est-elle associée au travail sur la gestion des données de la recherche ?

(28 réponses)

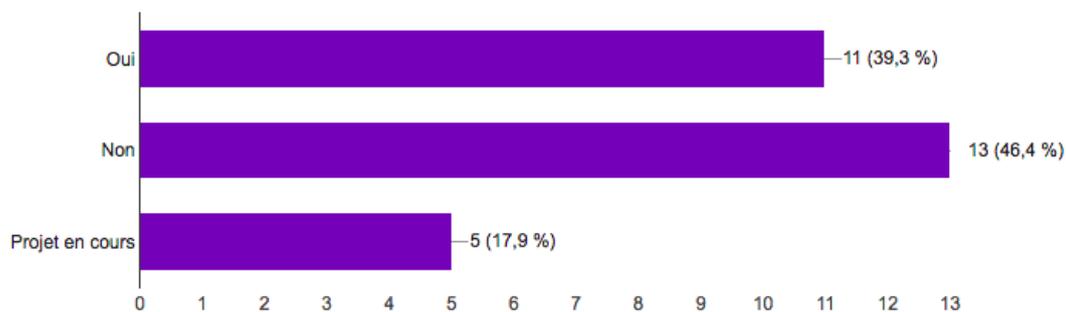


5 Si "oui" ou "projet en cours" précisez le rôle attendu de la bibliothèque et quel service de l'université pilote ce projet

6 Votre université dispose-t-elle d'une archive institutionnelle propre ? (voir note de bas de page 18)

Votre université dispose-t-elle d'une archive institutionnelle propre ?

(28 réponses)

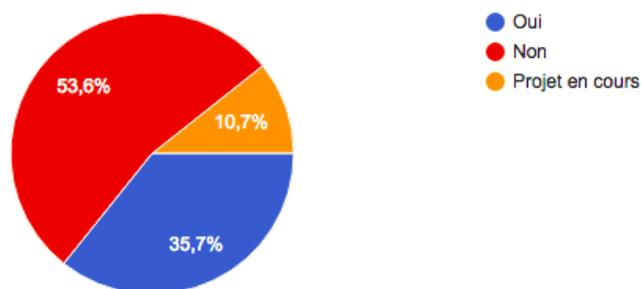


7 Si "oui" ou "projet en cours" comment la bibliothèque est associée au travail ?

8 Votre bibliothèque a-t-elle une ou plusieurs personnes dédiées au travail sur les données (hors données bibliographiques) ?

Votre bibliothèque a-t-elle une ou plusieurs personnes dédiées au travail sur les données (hors données bibliographiques) ?

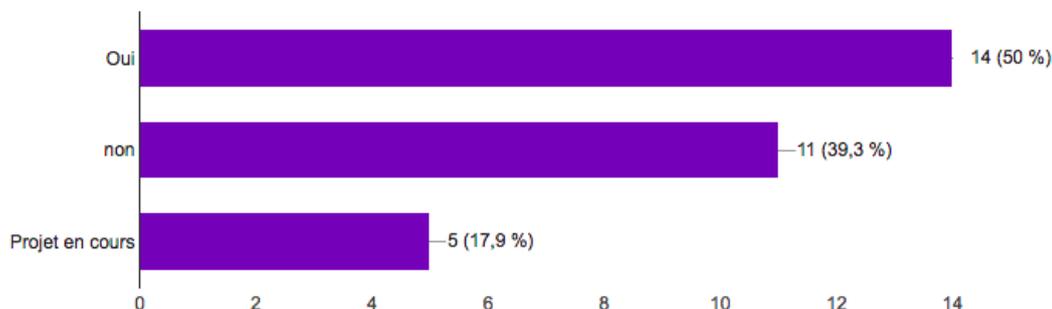
(28 réponses)



9 Votre bibliothèque collecte-t-elle automatiquement des données d'usage (domotique, objets connectés, RFID, logs des sites web, applications mobiles ...) ?

Votre bibliothèque collecte-t-elle automatiquement des données d'usage (domotique, objets connectés, RFID, logs des sites web, applications mobiles ...)

(28 réponses)

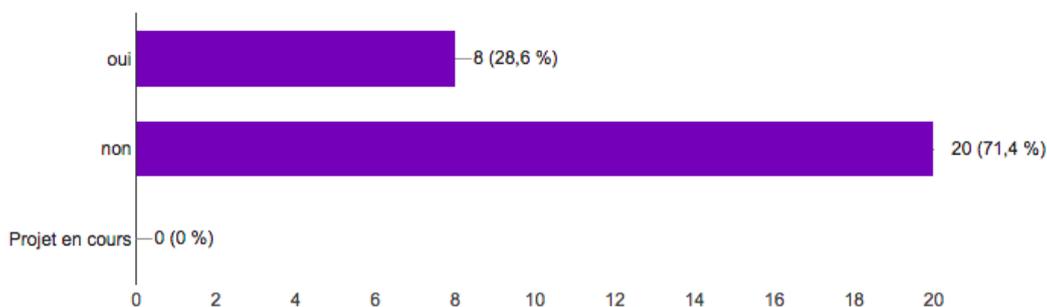


10 Si "oui" ou "projet en cours" précisez le ou les types de données collectées (nature, forme, mode de collecte, finalité de la collecte ...)

11 Votre bibliothèque collecte-t-elle des données auprès d'autres services de l'université à des fins statistiques ou d'améliorations des services ?

Votre bibliothèque collecte-t-elle des données auprès d'autres services de l'université à des fins statistiques ou d'améliorations des services ?

(28 réponses)

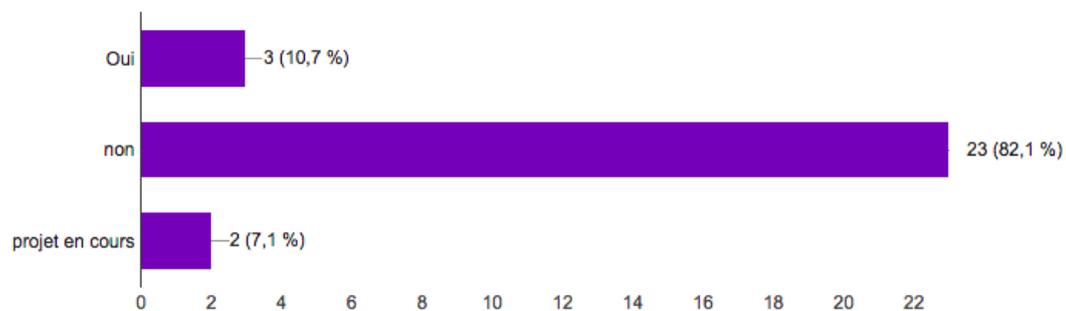


12 Si "oui" ou "projet en cours", précisez les données collectées (nature, forme, finalité de la collecte)

13 Votre bibliothèque travaille-t-elle sur des projets en dataviz ? (statistiques de la bib, en lien avec un ou des projets de recherche, d'informatique décisionnelle)

Votre bibliothèque travaille-t-elle sur des projets en dataviz ? (statistiques de la bib, en lien avec un ou des projets de recherche, d'informatique décisionnelle)

(28 réponses)

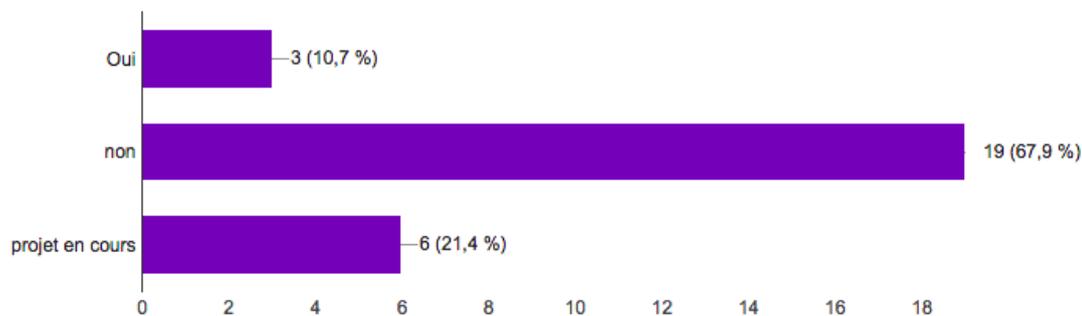


14 Si "oui" ou "projet en cours", précisez les données collectées (nature, forme, finalité de la démarche)

15 Votre bibliothèque travaille-t-elle sur un projet de catalogue en FRBR ou de FRbérisation de son catalogue ?

Votre bibliothèque travaille-t-elle sur un projet de catalogue en FRBR ou de FRbérisation de son catalogue ?

(28 réponses)

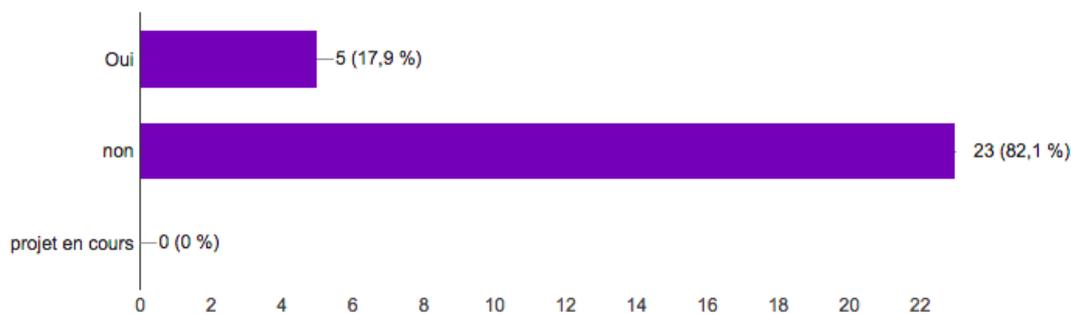


16 Si "oui" ou "projet en cours", précisez le cadre, la démarche et l'articulation avec les outils bibliographiques (professionnels et publics) existants

17 Votre bibliothèque travaille-t-elle sur des outils d'informatiques décisionnelles (business object, cognos...) ?

Votre bibliothèque travaille-t-elle sur des outils d'informatiques décisionnelles (business object, cognos...) ?

(28 réponses)

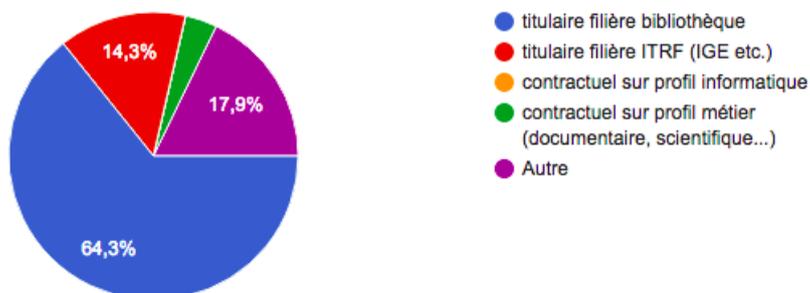


18 Si "oui" ou "projet en cours", précisez le type et l'origines des données collectées (nature, forme, finalité de la collecte)

19 Les personnels de la bibliothèque travaillant sur les projets évoqués en section 1 sont-ils :

Les personnels de la bibliothèque travaillant sur les projets évoqués en section 1 sont-ils :

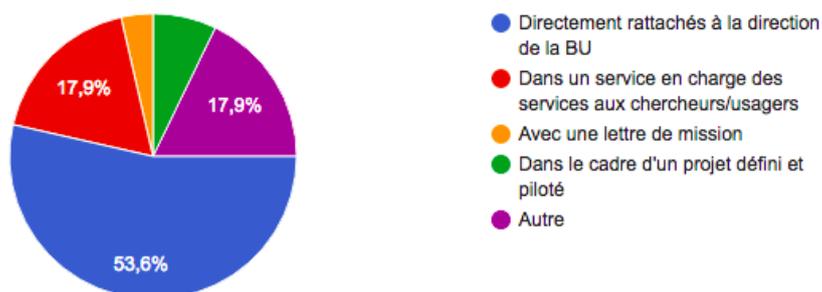
(28 réponses)



20 Si "autre", précisez :

21 Dans quels cadres ces personnels travaillent-ils ?

Dans quels cadres ces personnels travaillent-ils ? (28 réponses)

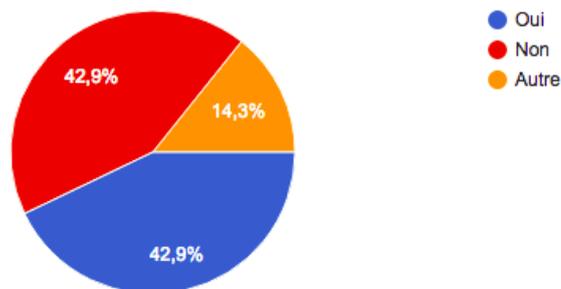


22 Si autre, précisez :

23 Les fonctions de ces personnels de la bibliothèque en lien avec ces projets sont-elles explicites dans l'organigramme de votre établissement ?

Les fonctions de ces personnels de la bibliothèque en lien avec ces projets sont-elles explicites dans l'organigramme de votre établissement ?

(28 réponses)

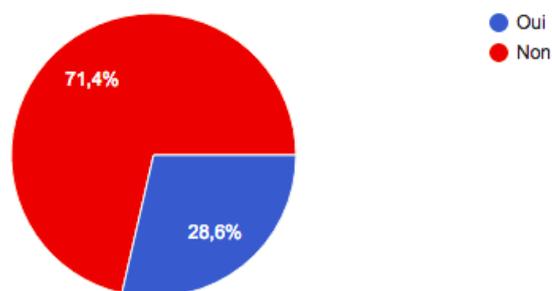


24 Si "autre", précisez :

25 Est-il envisagé au sein de votre établissement de modifier l'organigramme afin d'améliorer la gestion des données ?

Est-il envisagé au sein de votre établissement de modifier l'organigramme afin d'améliorer la gestion des données ?

(28 réponses)



26 Si "oui", précisez :

ANNEXE 2 – TABLEAU DE RESULTATS DE L'ENQUETE QUANTITATIVE

| Horodateur | les données de la recherche ? | plan de gestion des données de la recherche ? | archive institutionnelle propre ? | travail sur les données (hors données bibliographiques)? | automatiquement des données d'usage (domotique, objets connectés, RFID, logs des sites web, applications mobiles ...) | d'autres services de l'université à des fins statistiques | dataviz | FRBR ou de FRbérisation | BI | quels projets data signalés | nombre projets data signalés | Dans quels cadres ces personnels travaillent-ils ? |
|--------------------|-------------------------------|---|-----------------------------------|--|---|---|-----------------|-------------------------|-----|-------------------------------------|------------------------------|---|
| 10/3/2016 14:55:18 | Non | Non | Non | Non | non | non | non | non | non | Non | 0 | Autre |
| 10/3/2016 15:11:27 | Non | Non | Oui | Non | Oui | oui | Oui | non | non | Archive, auto, dataviz | 3 | Directement rattachés à la direction de la BU |
| 10/3/2016 15:16:32 | Non | Non | Projet en cours | Oui | Oui | oui | projet en cours | projet en cours | non | archive, données, auto, FRBR, | 4 | Dans un service en charge des services aux chercheurs/enseignants |
| 10/3/2016 15:39:57 | Non | Non | Non | Non | Oui, Projet en cours | non | non | non | non | auto, | 1 | Directement rattachés à la direction de la BU |
| 10/3/2016 16:02:10 | Non, Projet en cours | Projet en cours | Non | Projet en cours | non | non | non | Oui | non | DR, données, FRBR | 3 | Directement rattachés à la direction de la BU |
| 10/3/2016 16:41:19 | Non | Non | Non | Projet en cours | Oui | non | non | projet en cours | non | données, auto, FRBR | 3 | Dans un service en charge des services aux chercheurs/enseignants |
| 10/4/2016 4:58:17 | Non | Non | Non, Projet en cours | Non | non | non | non | projet en cours | non | archive, FRBR | 2 | Directement rattachés à la direction de la BU |
| 10/4/2016 8:49:13 | Non | Oui | Non | Non | non | oui | non | non | non | NON | 0 | Directement rattachés à la direction de la BU |
| 10/4/2016 10:26:59 | Projet en cours | Non | Oui | Oui | non | oui | non | non | non | DR, Archives, données, indicateurs, | 4 | Directement rattachés à la direction de la BU |
| 10/4/2016 17:16:22 | Projet en cours | Non | Non | Non | Oui | non | non | non | non | DR, auto, | 2 | Dans le cadre d'un projet défini et piloté |

Annexes

| Horodateur | les données de la recherche ? | plan de gestion des données de la recherche ? | archive institutionnelle propre ? | travail sur les données (hors données bibliographiques)? | automatiquement des données d'usage (domotique, objets connectés, RFID, logs des sites web, applications mobiles ...) | d'autres services de l'université à des fins statistiques | dataviz | FRBR ou de Frbérisation | BI | quels projets data signalés | nombre projets data signalés | Dans quels cadres ces personnels travaillent-ils ? |
|---------------------|-------------------------------|---|-----------------------------------|--|---|---|---------|-------------------------|-----|-----------------------------|------------------------------|---|
| 10/6/2016 15:23:20 | Non | Non | Oui | Oui | Oui, Projet en cours | oui | non | non | non | archive, données, auto, | 3 | Directement rattachés à la direction de la BU |
| 10/9/2016 17:42:02 | Non | Non | Non | Non | Oui | non | non | non | non | auto, | 1 | Directement rattachés à la direction de la BU |
| 10/10/2016 12:38:28 | Non | Non | Non | Projet en cours | non | non | non | non | non | données | 1 | Directement rattachés à la direction de la BU |
| 10/14/2016 16:17:20 | Non | Non | Oui | Non | non | non | non | non | non | archive, | 1 | Dans un service en charge des services aux chercheurs/enseignants |
| 10/14/2016 16:56:39 | Non | Non | Oui | Oui | non | non | non | projet en cours | non | archive, données, FRBR | 3 | Dans un service en charge des services aux chercheurs/enseignants |
| 10/21/2016 11:11:28 | Non | Non | Oui | Oui | Projet en cours | non | Oui | non | non | archive, données, frbr | 3 | Autre |
| 10/4/2016 11:29:44 | Non | Non | Oui | Non | non | non | non | non | Oui | archive, BI | 2 | Directement rattachés à la direction de la BU |
| 10/4/2016 15:33:16 | Non | Non | Non | Non | Projet en cours | non | non | projet en cours | Oui | auto, FRBR, BI | 3 | Autre |
| 10/5/2016 14:57:14 | Projet en cours | Projet en cours | Projet en cours | Oui | non | non | non | non | Oui | DR, Archives, données, | 3 | Dans le cadre d'un projet défini et piloté |
| 10/5/2016 18:47:37 | Non | Non | Non | Non | non | non | non | Oui | Oui | FRBR, BI, auto, univ, | 2 | Autre |
| 10/12/2016 15:43:20 | Non | Non | Non | Non | Oui | oui | non | Oui | Oui | FRBR, BI | 4 | Directement rattachés à la direction de la BU |
| 11/2/2016 21:50:10 | Non | Non | Oui | Non | Oui | non | non | non | non | archive, auto | 2 | Avec une lettre de mission |

Annexes

| Horodateur | les données de la recherche ? | plan de gestion des données de la recherche ? | archive institutionnelle propre ? | travail sur les données (hors données bibliographiques)? | automatiquement des données d'usage (domotique, objets connectés, RFID, logs des sites web, applications mobiles ...) | d'autres services de l'université à des fins statistiques | dataviz | FRBR ou de Fbrérisation | BI | quels projets data signalés | nombre projets data signalés | Dans quels cadres ces personnels travaillent-ils ? |
|---------------------|-------------------------------|---|-----------------------------------|--|---|---|-----------------|-------------------------|-----|-------------------------------------|------------------------------|---|
| 11/2/2016 22:10:10 | Non | Non | Projet en cours | Non | Oui | non | non | non | non | archive, auto | 2 | Dans un service en charge des services aux chercheurs/enseignants |
| 11/3/2016 15:53:09 | Non | Non | Oui | Oui | Oui | oui | non | non | non | archive, auto, données, université | 4 | Directement rattachés à la direction de la BU |
| 11/3/2016 17:59:12 | Projet en cours | Projet en cours | Oui | Oui | Projet en cours | non | projet en cours | non | non | DR, archive, données, auto, dataviz | 5 | Autre |
| 11/17/2016 11:02:54 | Non | Oui | Oui | Oui | Oui | non | Oui | non | non | archives, données, auto, daaviz | 4 | Directement rattachés à la direction de la BU |
| 11/17/2016 16:22:42 | Non | Non | Non | Non | Oui | non | non | non | non | auto | 1 | Directement rattachés à la direction de la BU |

A consulter [ici](#) :

<https://docs.google.com/spreadsheets/d/14pWX6c-0xqy57QYuFNXgaOX89ZsMFWvda8qwkmmeWUrU/edit?usp=sharing>

ANNEXE 3 – LE GUIDE D’ENTRETIEN EN PRESENTIEL

Les questions sont abordées sous trois angles, auxquels s’ajoute une dimension prospective :

1. Les données
2. L’organisation du travail et les ressources humaines
3. L’organisation en silos
4. Questions prospectives

| Thématique | Questions ouvertes | Questions fermées | traité |
|--|---|---|--------------------------|
| Les données : | Quels types de données ont-ils fait l’objet de projet spécifiquement data ? | Votre projet autour des données porte-t-il sur des données de la bibliothèque ou d’autres entités ? | <input type="checkbox"/> |
| | | S’agit-il de données bibliothéconomiques, d’usage, scolarité, web, recherche | <input type="checkbox"/> |
| | A quelles fins ces données ont-elles été collectées et/ou traitées | Pour produire des indicateurs d’usage ? autre ? | <input type="checkbox"/> |
| | | S’agit-il de constituer de nouveaux services aux usagers ? | <input type="checkbox"/> |
| | | S’agit-il de produire des indicateurs de qualité de service ? | <input type="checkbox"/> |
| | A quel niveau la bibliothèque est-elle intervenue dans le travail sur ces données ? | Travail amont ou aval ? | <input type="checkbox"/> |
| | | Conception complète ou en partenariat ? | <input type="checkbox"/> |
| | | Production ? archivage ? | |
| | Comment a été traitée la question de la qualité, propreté des données ? | Y a-t-il une démarche qualité en amont sur les processus qui collectent et produisent les données ? | <input type="checkbox"/> |
| L’organisation et les ressources humaines | Comment ont été organisés les projets data ? | Y a-t-il une gouvernance projet ? | <input type="checkbox"/> |
| | | Y a-t-il une gouvernance des données ? | <input type="checkbox"/> |

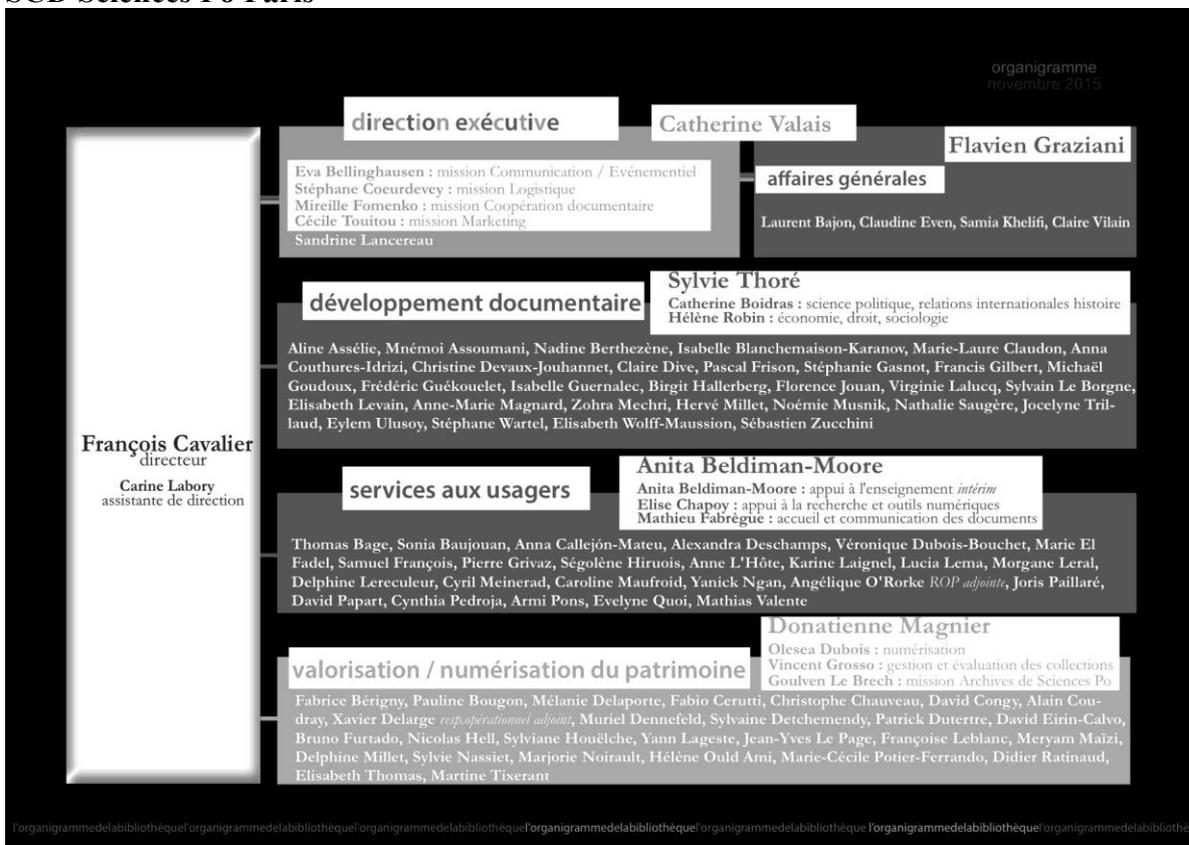
| | | |
|---|--|--|
| | Cette gouvernance est-elle technique (CRI) ou métier ? | <input type="checkbox"/> |
| Où se trouvent les compétences sur lesquelles le projet s'appuie ? | Les compétences sont-elles identifiées et cohérente avec la gouvernance ? | <input type="checkbox"/> |
| | Les compétences sont elles réparties parmi les différents partenaires du projet ? Au delà ? | <input type="checkbox"/> |
| | Ces besoins de compétences ont-ils guidé la gouvernance du projet ? | <input type="checkbox"/> |
| Pourquoi avoir sollicité des partenaires ou des sous-traitants ? | Y a-t-il une stratégie, une démarche sur l'identification, l'acquisition des compétences nécessaires ? | <input type="checkbox"/> |
| Comment se sont construites les compétences internes à la bibliothèque ? à l'université ? au projet ? | Y a-t-il eu des formations ? en interne ou via un organisme ? | <input type="checkbox"/> |
| | Y a-t-il eu des recrutements spécifiques ? de quelles natures ? pérennes ou non ? | <input type="checkbox"/> |
| | Les compétences préexistaient-elles au projet ? | |
| | Avez vous rencontré des difficultés pour la montée en compétence des équipes ? => difficulté de recrutement ? obstruction des équipes sur ces compétences. | <input type="checkbox"/> |
| | Avez eu le soutien de votre institution pour la formation, l'accompagnement des équipes ? | <input type="checkbox"/> |
| Quelles compétences ont été nécessaires dans la bibliothèque ? | Compétence projet ? démarche qualité ? informatique (technique bdd ou infrastructure) compétence numérique au sens large, compétence big data etc. | <input type="checkbox"/> |
| | Dans l'université ? | <input type="checkbox"/> |
| | Dans le groupe projet ? | <input type="checkbox"/> |
| Les silos et les représentations professionnelles | Quelles difficultés les équipes ont rencontré pour mener à bien le travail sur les données ? | Difficultés lors de la collecte ? (changement de processus de travail) |
| | | <input type="checkbox"/> |
| | Difficultés liées à la qualité des données elles-mêmes (architectures des données ou qualité des données elles-mêmes) | <input type="checkbox"/> |
| | Difficultés liées aux outils (absence ou manque d'outils, formation, ajout au travail existant etc.) | <input type="checkbox"/> |

| | | | |
|---|--|---|--------------------------|
| | Comment le projet data s'articule-t-il avec le fonctionnement courant de l'institution ? | Le projet est-il une verrou à côté des processus concernés, venant le perturber pour collecter les données, modifier les processus, sans associer les personnes travaillant sur ces processus ? | <input type="checkbox"/> |
| | | Les équipes concernées voient-elles le lien entre leurs activités traditionnelles et le travail sur les données ? | <input type="checkbox"/> |
| | | La hiérarchie du SCD ou de l'université voyaient-elles le lien entre le travail des équipes et celui sur les données ? | <input type="checkbox"/> |
| | Quelles difficultés organisationnelles ou humaines ont été rencontrées ? | Les équipes pensent-elles qu'il s'agit bien de la mission de la bib que de travailler sur ces sujets ? | <input type="checkbox"/> |
| | | Les équipes pensent-elles qu'il s'agit de leur travail dans leur service ? sinon de quoi s'agit-il selon elles ? | <input type="checkbox"/> |
| | Comment ont été gérés les obstacles humains et organisationnels ? | Le projet data a-t-il été autonomisé au maximum des activités de la bibliothèque ? | <input type="checkbox"/> |
| | | Les silos organisationnels ont-ils été conservés pour protéger le projet de ces obstacles ? | <input type="checkbox"/> |
| | | Un accompagnement des équipes impactées et pas forcément associées au projet a-t-il été envisagé, ou mis en œuvre ? | <input type="checkbox"/> |
| | Une autre organisation aurait-elle permis de lever ces difficultés ? ou en causer d'autres ? | Les difficultés sont-elles dues au projet et son organisation ou au fonctionnement de l'institution et de ses équipes ? | <input type="checkbox"/> |
| | | | <input type="checkbox"/> |
| La projection de projets modifiant la place et le rôle des données dans le travail | Quels types de projets sont identifiés impliquant la place ou le rôle des données ? | Le projet est-il en lien avec le catalogue, les données de la recherche, les indicateurs de pilotage, les données internes de l'université ? autre ? | <input type="checkbox"/> |
| | Quelles questions se posent pour la mise en place d'un tel projet ? | Les questions organisationnelles sont-elles identifiées ? | <input type="checkbox"/> |
| | | Les questions de pilotage sont-elles identifiées ? | <input type="checkbox"/> |
| | Quelles compétences sont identifiées pour mettre | | <input type="checkbox"/> |

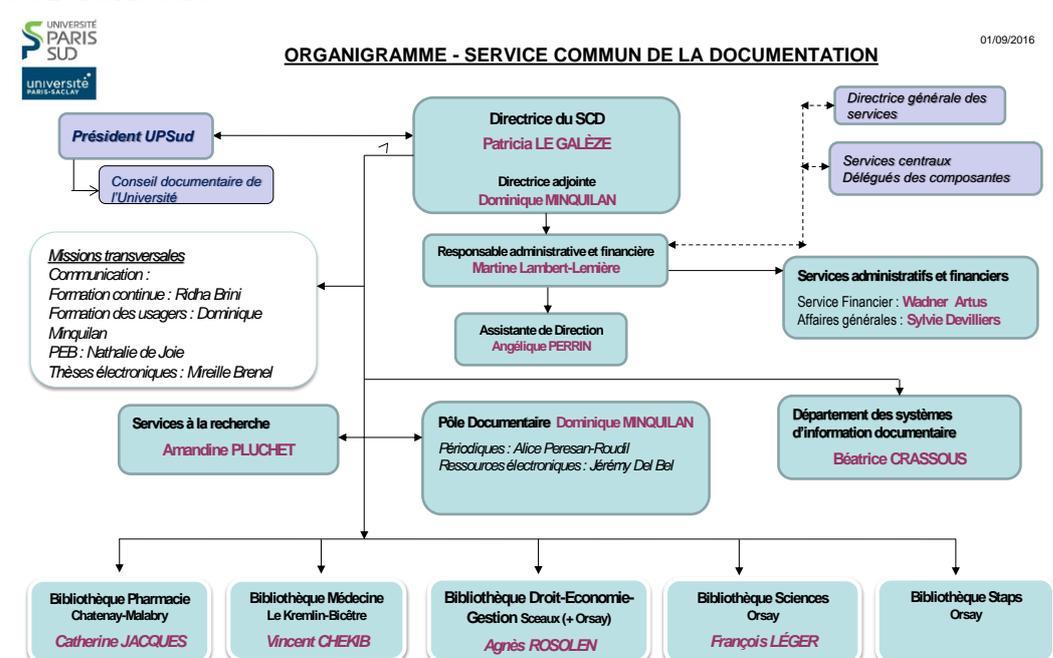
| | | |
|--|---|--------------------------|
| en œuvre ce projet ? | | <input type="checkbox"/> |
| Quelle organisation est envisagée entre les données et les services qu'elles permettront ? | | <input type="checkbox"/> |
| Quels facteurs de risque sont identifiés ? | Politique ? organisationnels ? humains ? compétences ? | <input type="checkbox"/> |
| Quelle structure aura le pilotage du projet ? | | <input type="checkbox"/> |
| Quelle part est envisagée pour une sous-traitance ? | | <input type="checkbox"/> |
| Envisagez-vous qu'il est nécessaire ou utile de mettre en place une gouvernance propre aux données | La question des droits sur les données, ainsi que sur leur qualité et le lien avec les processus implique-t-il une gouvernance plus clairement identifiée ? | <input type="checkbox"/> |

ANNEXE 4 – LES ORGANIGRAMMES

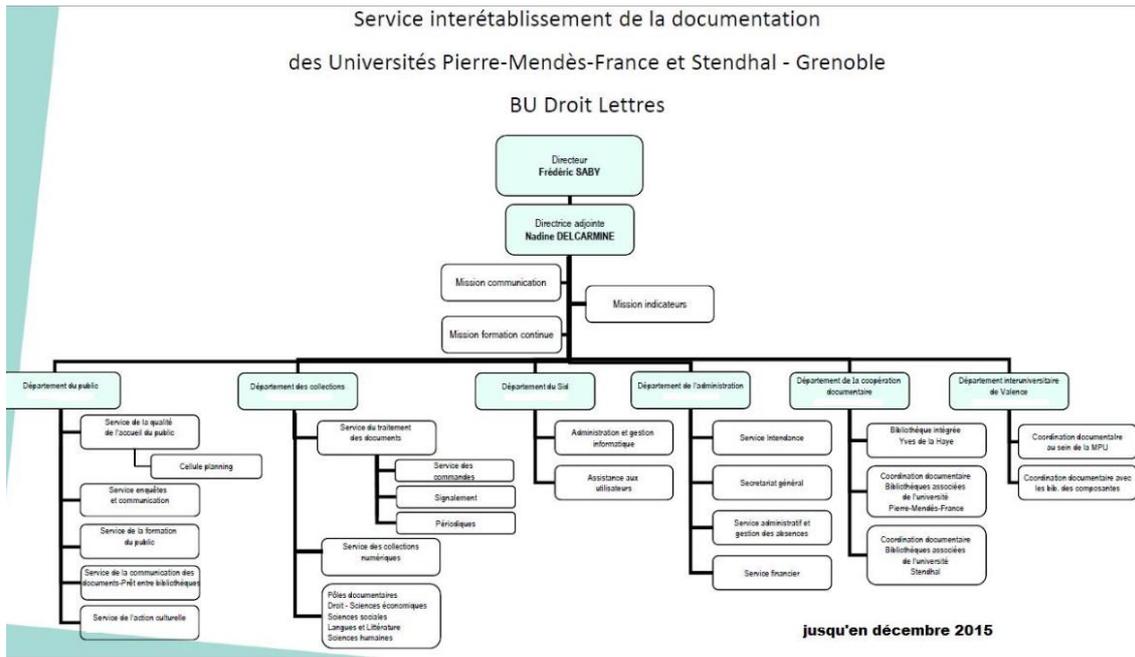
SCD Sciences Po Paris



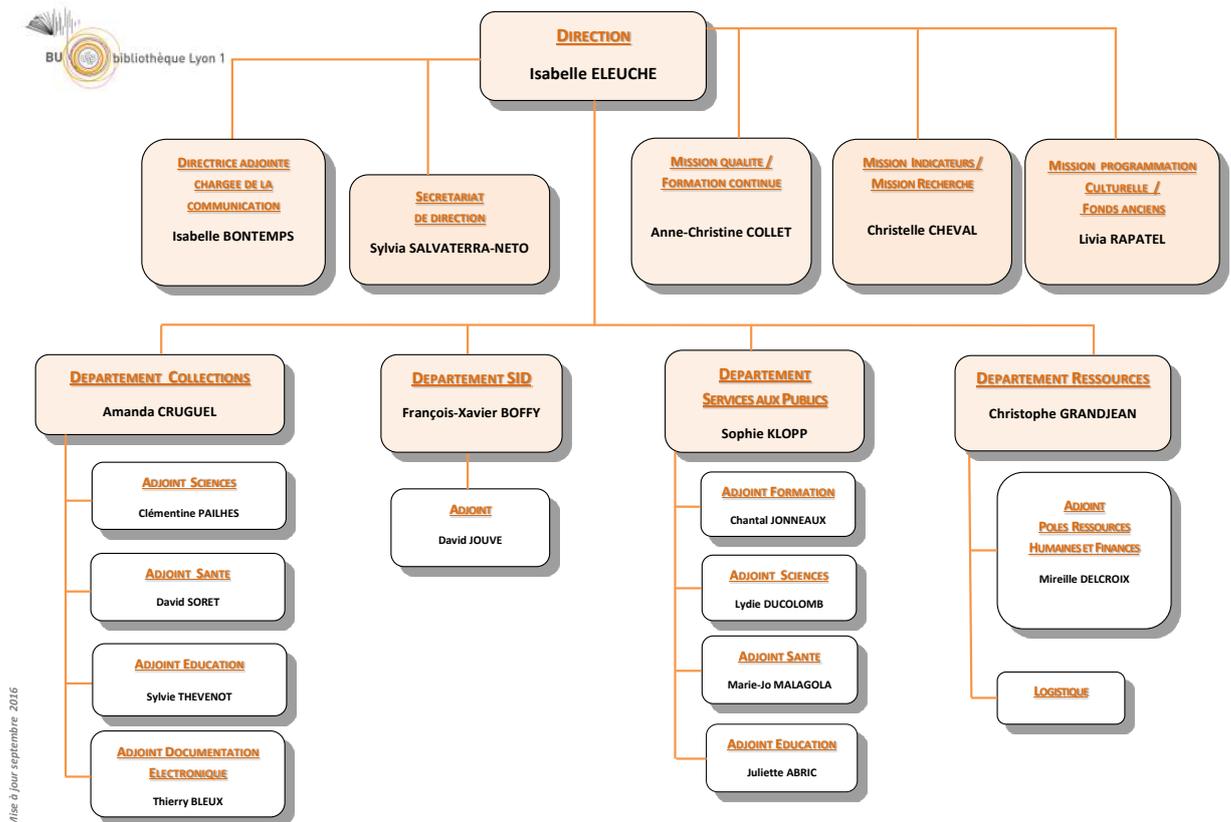
SCD Paris Sud



SCD Grenoble



SCD Lyon 1



INDEX

Affaires

| | |
|--|---|
| ADBU..... | 20 |
| adresse IP..... | 24 |
| AdWords..... | 23 |
| Affluences..... | 14, 20, 24, 26, 31, 34, 43, 44 |
| Amazon..... | 17, 19, 22 |
| ANR..... | 45, 54, 65 |
| AOI-PMH..... | 32, 57 |
| API..... | 68 |
| Application Programming Interface..... | 68 |
| arXive..... | 24 |
| BdD..... | 19 |
| Bibliothèque du Congrès..... | 15 |
| Big Data..... | 13, 15, 16, 17, 18, 19, 27, 30, 44 |
| BnF..... | 17, 20, 28, 30, 55 |
| British Library..... | 18 |
| BSN..... | 47 |
| BU..... | 9, 13, 17, 18, 25, 32, 47, 50, 53, 58, 61, 62, 68 |
| CAIRN..... | 25 |
| catalogue Archives et manuscrits..... | 28 |
| CDO..... | 44 |
| Chief Digital Officer..... | 44 |
| CINES..... | 25 |
| cloud..... | 22, 23, 24 |
| CNRS..... | 16 |
| COMUE..... | 33 |
| crowdsourcing..... | 28 |
| data management plan..... | 48 |
| data visualisation..... | 15, 34, 45 |
| data.bnf.fr..... | 30 |
| data.gouv.fr..... | 25 |
| dataviz..... | 45 |
| DBpedia..... | 17 |
| Design Thinking..... | 20 |
| Dilicom..... | 25 |
| DMP..... | 48, 49, 50 |
| Doranum..... | 48 |
| DRM..... | 25 |
| DSI..... | 35, 50, 64 |
| Dublin Core..... | 51 |
| eBooks..... | 25 |
| Enssib..... | 60, 65 |
| ETP..... | 55 |
| ezparse..... | 13 |

| | |
|---------------------------|--|
| eZproxy | 31 |
| Facebook..... | 17, 19, 21, 27 |
| FMSH | 16 |
| FOAF..... | 17 |
| folksonomie | 27 |
| FRBérisation | 18, 62 |
| FRBÉRisation..... | 15, 68 |
| FRBR..... | 30, 39, 47, 64, 68, 87 |
| Gallica | 28 |
| GED..... | 39, 51 |
| GIS | 47, 48 |
| Google | 17, 19, 23, 27, 31, 34, 46 |
| Google Analytics..... | 20, 23, 34 |
| H2020 | 13, 46, 48, 49, 54 |
| Hadoop | 19 |
| HAL..... | 24, 25, 26, 32, 33, 35, 39, 45, 47, 67 |
| humanités numériques..... | 65 |
| Humanités Numériques..... | 13, 15, 16, 64 |
| Huma-num | 24, 53 |
| Huma-Num | 16, 50, 53, 55, 65 |
| Ibooks | 25 |
| idref..... | 24 |
| IGB..... | 9 |
| Illiad | 28 |
| INIST | 47, 48, 51 |
| INSEE..... | 21, 25 |
| IoT..... | 28 |
| ISNI..... | 17 |
| ISO 9001 | 36, 60 |
| Jussieu PMC | 28 |
| Kindle | 25 |
| Kobo | 25 |
| LABEX..... | 55 |
| Learning Center..... | 28 |
| Linked Data | 17, 30 |
| LRU..... | 9, 13, 45 |
| MARC | 18, 51 |
| MCC..... | 25 |
| médiamétrie | 27 |
| MESR | 47 |
| Mines-Telecom | 28 |
| Nakalona..... | 16 |
| Netflix | 19 |
| OAI..... | 32 |
| OAI-PMH | 51 |
| OCLC | 17, 18 |
| Open Data | 17, 26 |
| Orcid..... | 24 |

| | |
|----------------------------------|--|
| PageRank..... | 27, 46 |
| Paris 8..... | 20 |
| Paris-Saclay | 28 |
| PNB..... | 25 |
| <i>PREVU</i> | 20, 45, 67 |
| RDF..... | 51 |
| Reed-Elsevier..... | 25 |
| RFID..... | 13, 14, 21, 43, 47, 84 |
| SCD..... | 10, 13, 29, 32, 33, 35, 38, 41, 42, 45, 47, 48, 49, 50, 51, 52, 53, 64 |
| SGBM..... | 18, 39, 68 |
| SI..... | 43 |
| SIGB..... | 18, 20, 21, 22, 24, 30, 51, 57, 62, 63 |
| SUDOC | 39 |
| Tableaux Public..... | 45 |
| Text and Data Mining (TDM) | 13 |
| TextMining | 29 |
| UE | 46 |
| URFIST..... | 47, 48, 50, 51, 65 |
| URI..... | 17 |
| URL..... | 24 |
| UX..... | 20 |
| VIAF..... | 17 |
| voyant Tools..... | 16 |
| Voyant Tools..... | 53 |
| webservices | 68 |
| Wikidata | 17 |
| Wikimedia..... | 17 |
| workflow..... | 21 |
| Xiti | 20, 23, 24 |
| XML..... | 51 |
| Z-39-50..... | 51 |

TABLE DES MATIERES

| | |
|--|------------|
| SIGLES ET ABBREVIATIONS | 7 |
| INTRODUCTION..... | 9 |
| LES ENJEUX D'UNE CONCEPTION ELARGIE DES DONNEES : OBJET ET METHODE | 13 |
| 1.1. Problématique générale | 13 |
| 1.2. Présentation de l'enquête et des données collectées | 14 |
| 1.3. Etat de l'art et questions identifiées | 15 |
| 1.4. Les révolutions de la data | 17 |
| 1.5. Extension du domaine des données en bibliothèque..... | 20 |
| <i>1.5.1. Les données internes à la bibliothèque.....</i> | <i>21</i> |
| <i>1.5.2. Les données externes à la bibliothèque</i> | <i>23</i> |
| <i>1.5.3. Disruption et position de la bibliothèque par rapport à ses données.....</i> | <i>27</i> |
| ORGANISATIONS ET DONNEES | 30 |
| 2.1. La place des données dans l'organisation..... | 31 |
| <i>2.1.1. Outils, données et maturation des processus de travail</i> | <i>31</i> |
| <i>2.1.2. Maîtrise des données, maîtrise des organisations.....</i> | <i>35</i> |
| <i>2.1.3. La place du SCD et des données</i> | <i>42</i> |
| 2.2. Mises en œuvre de la gestion des données..... | 46 |
| QUELS CHANGEMENTS POUR LE TRAVAIL EN BIBLIOTHEQUE ?.... | 57 |
| 3.1. l'utilisateur au cœur ... ou la stratégie du changement | 57 |
| 3.2. Statuts, culture métier et représentations professionnelles..... | 64 |
| CONCLUSION | 69 |
| BIBLIOGRAPHIE..... | 72 |
| ANNEXES..... | 81 |
| INDEX | 99 |
| TABLE DES MATIERES..... | 103 |