



HAL
open science

Les bibliothèques et le web de données. D'un système de recherche informationnel à un dispositif de médiation

Kaouther Azouz

► To cite this version:

Kaouther Azouz. Les bibliothèques et le web de données. D'un système de recherche informationnel à un dispositif de médiation. Journées doctorales de la SFSIC 2015, May 2015, Villeneuve d'Ascq, France. pp.341 - 251. hal-01393082

HAL Id: hal-01393082

<https://hal.science/hal-01393082>

Submitted on 6 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Les bibliothèques et le web de données

D'un système de recherche informationnel à un dispositif de médiation

Libraries and linked data

From an evolution information based research system to a mediation device

Kaouther AZOUZ

Laboratoire GERiCO (Groupe d'Études et de Recherche
Interdisciplinaire en Information et Communication – EA 4073)

Mots clés : Web de données, médiation, bibliothèques, contenu numérique, usagers

Key words: Linked data, mediation, digital content, web users, libraries, digitization

Résumé

Le web de données bouleverse les pratiques des bibliothèques qui ne limitent plus leur rôle à la numérisation et à la mise en ligne du patrimoine. La ré-exploitation et l'appropriation de leurs collections numériques sont leurs principales visées sur le web. Pour y parvenir, elles mettent en place des stratégies ciblant la revalorisation et l'enrichissement des contenus en fonction des besoins de l'utilisateur. Cette situation relègue aux bibliothèques le statut de lieux de savoir et de médiateurs numériques du web.

Abstract

Linked data has urged libraries to change their practices which are no longer confined to digitization and content uploading. Enabling reuse of their digital content by web users is increasingly becoming their major concern. In order to achieve such target, libraries should develop several online strategies of enrichment and upgrading of their content that will meet users' needs. Hence libraries will not only be perceived as temples of knowledge but also as digital mediators on the web.

[Si les bibliothèques seront considérées comme uniquement des groupes de données elles disparaissent, en revanche si elles seront considérées comme des lieux d'accès aux ressources et des lieux de médiation, à ce moment-là elles peuvent plutôt perdurer]

B.Bachimont

Introduction

Cette recherche vise l'analyse d'un nouveau paradigme des bibliothèques sur le web. Ne se limitant plus à l'accès et à la recherche de l'information, les usagers sont en quête de contenus intelligibles (Bermès et al, 2013) et d'une information précise qu'ils pourraient réutiliser. Cette pratique incite les bibliothèques à réviser leurs stratégies de valorisation des contenus et à comprendre les mutations d'usages que génèrent l'utilisation des moteurs de recherche (Chaudiron, Ihadjadene, 2008). L'évolution du web d'un dispositif d'accès, de recherche, de consultation et de diffusion de l'information vers un dispositif de partage et d'échange impacte les pratiques des bibliothèques et des usagers.

Cadre théorique

Quel ancrage pour la médiation en SIC ?

La médiation est définie par le centre national des ressources textuelles et lexicales comme le fait de servir d'intermédiaire entre deux ou plusieurs choses. Elle se caractérise en sciences de l'information et de la communication par la densité des écrits, la variabilité des significations et la multiplicité des approches. La richesse de la notion nous incite à étudier les nouvelles formes de communication (Davallon, 2003) du patrimoine numérique des bibliothèques. Dans son rapport au Ministère de la Culture et de la Communication, Pierre Lescure met l'accent sur la médiation culturelle des bibliothèques et la nécessité de sa réadaptation aux mutations du numérique. Igor Babou et Joëlle Le Marec définissent cette médiation comme l'interaction entre la bibliothèque comme organisation matérielle et spatiale des connaissances et les moyens permettant la communication de son savoir auprès du public (Babou et Le Marec, 2003).

La médiation numérique du patrimoine est présentée par Bernadette Dufrène comme les « [...] prolongements des médiations traditionnelles » (Dufrène et al, 2013) et une nouvelle forme de représentation de l'institution, de son patrimoine et de son offre de service. Xavier Galaup, considère que la médiation ne peut se faire que par le biais d'un tiers comme les interfaces qui accompagnent

l'utilisateur et facilitent les usages (Galaup, 2012). Bien qu'ils partagent cette approche, Cécile Gardiès, Isabelle Fabre et Vincent Liquète considèrent que la médiation permet la mise en relation de deux « [...] choses jusque là non rassemblées pour établir une communication et un accès à l'information » (Liquète et al, 2010). Ce sont ces différentes notions de la médiation qui intéressent notre recherche puisqu'elles rappellent notre problématique et réitérent la médiation dans un nouveau dispositif : le web de données.

Les approches d'interdépendance entre médiation et dispositif

La médiation numérique nécessite des dispositifs techniques permettant l'utilisation des contenus et l'accompagnement de l'utilisateur du web (Gardiès et al, 2010). Elle se présente pour Silvère Mercier comme une démarche permettant la mise en place de dispositifs « passerelles », « de flux » et « ponctuels » permettant l'accès, l'appropriation ou la dissémination de contenus à des fins de diffusion de savoir et de savoir-faire. Cette définition évoque l'interaction de l'utilisateur avec les contenus par le biais de dispositifs dédiés comme les catalogues en ligne, etc. Amanda Rueda considère la médiation et le dispositif comme indispensables à la transmission du contenu et à son appropriation par l'utilisateur (Rueda, 2010). Elisabeth Caillet et Daniel Jacobi, mettent l'accent sur l'interaction entre les acteurs et les dispositifs techniques utilisés dans la création de l'objet artistique et les formes de sa revalorisation pour une réappropriation par le public (Rueda, 2010). Malgré la richesse des approches, on note la réémergence de la médiation des contenus avec le web de données. Entendu comme le « [...] web des objets documentaires communicants » (Boustany et al, 2013), le web de données vise la mise en relation d'un ensemble d'objets et de contenus au profit de l'internaute et en fonction de ses besoins. L'idée est de parvenir à l'appropriation de ces ressources par l'utilisateur. Selon Bertrand Calenge, les bibliothèques voudraient infirmer le discours évoquant leur désintermédiation. (Calenge, 2015).

La médiation et la notion de l'espace

La médiation est en interaction avec l'espace dans lequel elle s'opère. La notion de spatialisation est très abordée par les professionnels de l'information et de la communication. Elle désigne le lieu physique et l'espace numérique du web. La médiation des bibliothèques physiques est différente de celles du web. Elle nécessite l'utilisation de stratégies ciblant le repérage géographique du lieu et la valorisation des collections auprès des lecteurs (Bermès et al, 2013). L'espace numérique se distingue par la diversité de ses contenus et la multiplicité des dispositifs d'accompagnement aux usagers, ce qui lui alloue le statut de dispositif médiateur (Fabre, 2013). C'est un espace qui médiatise le savoir et assure le développement des institutions culturelles et des collections patrimoniales (Rizza, 2014). Yves Jeanneret, considère la médiation comme le produit d'un espace « dense » de structures et de « constructions » reliant les acteurs à leurs objets de communication

(Bonaccorsi, 2009). C'est le cas du web où l'on retrouve la notion du « réseau » et de mise en relation des acteurs aux objets et aux contenus communiqués.

Problématique et hypothèses

Si la médiation est bien ancrée, pourquoi ré-émerge-t-elle sur le web de données ? Est-elle le nouveau paradigme des bibliothèques sur le web ? Quelle offre de service fournira-t-elle aux usagers du web ? Si les bibliothèques visent la médiation cela sous entend une visibilité peu conséquente de leur patrimoine. Cette recherche analysera l'appropriation des contenus puis l'enrichissement du patrimoine en fonction du besoin de l'utilisateur et conclura par l'étude du web de données comme un espace de mise en relation des acteurs du patrimoine et de leurs contenus.

Méthodologie

L'enquête sociologique et l'approche ethnographique sont les principales méthodes de cette recherche. N'ayant pas d'échantillon statistique nous avons analysé le comportement des enquêtés face à notre objet d'étude : le web de données. Or, comme le mentionne Bourdieu, le choix d'un sujet est rarement neutre car il est souvent lié au vécu du chercheur. Notre méthodologie est un croisement entre l'ethnographie réflexive et analytique car elle recoupe nos connaissances et prénotions professionnelles sur la médiation des bibliothèques à notre processus de recherche sur un objet qu'on maîtrise peu : le web de données.

Profils des enquêtés

Cette enquête s'est déroulée entre février-avril 2014. Elle a ciblé 9 professionnels qui se répartissent comme suit :

- ✓ Un chef de projet (BnF)
- ✓ Une responsable du service prospective et services documentaires (BnF)
- ✓ Une responsable au Département de l'information bibliographique et numérique - coordination des produits et services bibliographiques (BnF)
- ✓ Un chargé d'études documentaires et chef de projet - Ministère de la Culture et de la communication
- ✓ La responsable de l'édition de systèmes d'information documentaire - réseau CANOPE Académie de Poitiers
- ✓ La responsable du Webmagazine chez Bibliothèque publique d'information - Centre Pompidou
- ✓ Un membre du conseil d'administration de l'agence française d'informatique, gérant d'une entreprise spécialisée dans les solutions libres des systèmes intégrés de gestion des bibliothèques

- ✓ Un directeur des offres clients et responsable des produits d'une base de données bibliographique française
- ✓ Un enseignant-chercheur expert dans les technologies du web de données dans l'institution culturelle

Déroulement de l'enquête

Cette enquête s'est fondée sur la conduite d'entretiens semi-directifs et compréhensifs. Ce sont les seconds, majoritairement présents (6 /9), qui nous ont permis de nous distancier de notre grille d'entretien. A la question « pensez-vous que le web de données est le média de valorisation des pratiques professionnelles des bibliothèques principalement celles des catalogueurs sur le web ? » nous étions confrontées à deux attitudes : celle qui recadre notre questionnement et rompt avec nos prénotions sur la médiation et celle qui crée une « enquête dans l'enquête » (Kaufmann, 2014). En d'autres termes nous étions face à une situation imprévue qui nous a mené à interagir avec nos enquêtés. Cette interaction se traduit par la mise en place de nouveaux questionnements. Concernant l'entretien semi-directif, il a suivi une grille d'entretien prenant en compte les problématiques du web de données en bibliothèques. On a dû réajuster cette grille en fonction du profil de l'enquêté. De visu, par téléphone ou sur Skype, ces entretiens ont duré entre une demi-heure et une heure. Pour une meilleure traçabilité et ré exploitation des données, nous avons enregistré nos entretiens après l'accord de nos enquêtés.

Création du corpus et analyse du contenu

Une fois enregistrés, ces entretiens sont retranscrits à l'aide d'une solution libre : Sonal, qui permet le découpage de l'enregistrement en plusieurs thématiques. L'intérêt de cet outil est qu'il favorise la transcription et l'analyse du contenu. Une fois les thématiques élaborées, on peut faire une requête sur un sujet donné au sein du corpus et importer la synthèse au format RTF ou Word.

Résultats et discussion

Résultats

Si la médiation ré-émerge en bibliothèques, c'est qu'il y a une visibilité et une réutilisation partielle du patrimoine numérique sur le web. Pour confirmer ou infirmer cette hypothèse, nous dédions cette partie à l'analyse des résultats et des verbatim de l'enquête.

De la recherche de l'information à l'appropriation des contenus

La recherche et l'accès à l'information ne sont plus les priorités des bibliothèques. Nos interviewés confirment l'atteinte de ces objectifs par d'autres dispositifs. « *De nos jours les gens n'ont pas besoin de consulter la bibliothèque pour avoir accès à l'information car il suffit pour cela de consulter le web pour trouver ce que l'on veut mais pas de manière certifiée comme le font les bibliothèques* » (B.B.,

UTC). Cette démocratisation du dispositif se heurte à la non-certification de leurs contenus et à la frustration de l'utilisateur en l'absence de l'objet recherché. « Dès qu'on est sur le web, l'utilisateur ou l'internaute ne voit pas l'intérêt de se limiter aux données des bibliothèques mais plutôt d'élargir son champ de recherche et de consultation, d'où ce sentiment de frustration quand il ne retrouve pas ce qu'il cherche vraiment. » (M.C., BnF).

Ce constat incite les bibliothèques à revoir leurs stratégies de visibilité en se mettant sur le « parcours critique » de l'utilisateur et en lui proposant des contenus authentiques non identifiés par les moteurs de recherche et susceptibles de l'intéresser (Azouz, 2014). L'idée est de lui fournir des contenus sans qu'il aille les chercher mais qu'il puisse se les approprier. L'appropriation est similaire à la lecture¹, elle n'est pas systématique et peut être substituée par une autre pratique. Elle ne se réalisera que lorsque les bibliothèques deviennent des lieux de savoir du web et non ses points d'accès. La réutilisation des données brutes et des référentiels de la Bibliothèque nationale de France confirme ce propos. A. Simon précise que certains usagers du data.bnf.fr sont parvenus à « la création d'un calculateur de données publiques » et que la plateforme du livre traduit *IF VERSO* utilise data.bnf.fr dans l'import des référentiels d'œuvres et de leurs traductions.

Le web de données et l'émergence d'un nouveau dispositif de médiation

L'enrichissement des contenus en fonction du besoin de l'utilisateur est une nouvelle stratégie de fidélisation et de revalorisation des contenus des bibliothèques. Elle permet à l'utilisateur d'affiner sa requête et d'avoir plus de probabilités de retrouver son objet de recherche. Le Réseau Canopé adopte cette démarche en intégrant les données du data bnf dans son portail e-sidoc pour répondre aux besoins de recherche des usagers : « Ce que je remarque au niveau des pratiques des enseignants et des élèves c'est que ces derniers utilisent beaucoup de contenus dans leurs établissements scolaires qui leur sont mis à disposition par de grandes institutions culturelle comme Gallica et tout ce que fait la BnF pour l'accompagnement des équipes éducatives ou des enfants » (CH.F- Réseau CANOPE)

Caroline Demescence² évoque le cas d'un lecteur qui cherchait une adaptation d'une nouvelle rédigée par un auteur anglais mais qui ne la retrouve pas dans les catalogues de la bibliothèque. « Là-dessus data.bnf.fr est très pratique puisqu'on ne pouvait chercher d'adaptateurs dans les catalogues alors que data peut très bien le faire. La personne était étonnée du nombre de résultats présentées et ne s'était pas rendu compte de la richesse des ressources du catalogue » (C. D. BnF). Par ailleurs cette richesse ne compte plus en l'absence de l'objet recherché. M. Clatin rebondit en évoquant l'expérimentation des

¹ Bachimont B., entretien Mars 2014

² Responsable au Département de l'information bibliographique et numérique Coordination des produits et services bibliographiques

usagers de la Bibliothèque de Fresnes qui étaient contents de la richesse du catalogue mais frustrés et perdus quand il ne leur offrait pas des contenus authentiques. Il fallait selon elle cibler les centres d'intérêts des publics et choisir les jeux de données adéquats : *« C'est aux bibliothécaires qui connaissent très bien leurs publics de faire le tri des jeux de données qui leur conviennent au mieux et qui intéressent surtout leurs publics. C'est un choix qui s'effectue dans le cadre de la politique documentaire pour le public »* (M. C., BnF)

Le web de données : vers un espace d'interconnexion des contenus et l'émergence d'un nouveau paradigme de médiation

Le dernier aspect de cette médiation est le décloisonnement des contenus et la mise en relation des acteurs. Nos enquêtés précisent que cette nouvelle manière de revalorisation du patrimoine est conditionnée par le choix et la mise en relation de données « propres », c'est-à-dire des données structurées avec le moins d'erreurs possibles. *« C'est bien les métadonnées qui en termes de qualité vont nous permettre d'assurer cette interconnexion »* (L.D. Electre). Cette convergence peut se faire par le biais du lien sémantique. Prônée par des adeptes du web de données, cette stratégie est critiquée par d'autres professionnels : *« Quand il y a trop de liens cela pourrait perturber l'utilisateur qui ne sait plus où chercher...Je constate qu'en salle lorsqu'il y a trop de liens, les lecteurs sont agacés parce qu'ils ne savent plus où aller et ne savent plus ce qu'il y a derrière »* (M.C BnF)

Cette interconnexion, permet l'émergence d'un nouveau paradigme de médiation entre les données audio-visuelles, archivistique, muséales, de lecture publique, de l'enseignement qui s'interconnectent pour une meilleure exploitation du patrimoine. Le web de données émerge comme une technique de mise en relation des contenus et un média reléguant aux bibliothèques le statut de lieu de savoir et de médiateurs numériques du patrimoine. La transition des bibliothèques d'un espace matériel (le bâtiment) vers un espace mouvant, sans frontières avec des profils divergents d'utilisateurs (le web) incite ces institutions à réadapter leur offre aux besoins de l'utilisateur.

Discussion

Cette communication confirme notre hypothèse. En nous référant aux définitions du cadre théorique on remarque qu'elles relient la médiation à l'appropriation de l'objet (Rueda, 2010), au dispositif, aux acteurs véhiculant les objets et les contenus, à l'espace (Jeanneret et Fabre), à l'action culturelle et aux interventions permettant la revalorisation du patrimoine (Caillet, Jacobi). Ce sont des critères qu'on retrouve dans notre recherche puisqu'on évoque l'appropriation du contenu comme la valeur ajoutée des bibliothèques. L'enrichissement du contenu en fonction du besoin de l'utilisateur est abordé par Emmanuelle Bermès dans son ouvrage « Le web sémantique en bibliothèques » (Bermès et al., 2013) qui mentionne la médiation des contenus et les différentes stratégies garantissant sa réussite.

L'appropriation, la réutilisation et l'enrichissement des contenus ou des collections sont des pratiques qu'on retrouve dans l'espace physique des bibliothèques. En effet les lecteurs qui consultent, lisent, réfléchissent sur le contenu d'une ressource parviennent généralement à se l'approprier pour créer un nouveau contenu. De même, l'enrichissement se traduit en bibliothèques physiques par la politique d'acquisition qui se fait en fonction des besoins des publics et des politiques documentaires de l'institution. On relève un recoupement entre les deux types de médiation. Par ailleurs, Catherine Dhérent³ émet des réserves sur cette médiation pour les raisons suivantes. Elle pense que l'hétérogénéité des contenus, des structures et des provenances de données, entravent le processus d'enrichissement des contenus. Si l'utilisateur ne trouve pas ce qu'il cherche il est perturbé et c'est difficile pour les professionnels de répondre positivement à sa requête pour des raisons techniques dont la propriété des données et la divergence des fichiers d'autorité car les bibliothèques ne se réfèrent pas toutes aux autorités de la BnF. La question qu'elle pose : « *est ce qu'on ne va pas trop loin en essayant de mettre en relation des termes qui sont différents mais qui veulent au final dire à peu près la même chose pour l'indexation, est ce qu'on ne va pas trop loin, est ce que le web d'aujourd'hui n'est pas tout à fait efficace pour le public ?* »

Conclusion

Le web de données est en train de changer le rôle des bibliothèques. Il n'est plus question de promouvoir l'information professionnelle standardisée (notices bibliographiques, catalographiques et d'autorité) mais de fournir une information publique. On relègue à l'internaute le statut de principal interlocuteur du web, d'où sa convoitise par les professionnels de l'information. Bibliothèques, éditeurs de contenus et institutions culturelles sont amenés à créer des passerelles entre leurs

³ Conservateur général du patrimoine, chef de projet Livre et lecture à Lille Métropole

données. La réussite de cette médiation dépend de l'évolution des pratiques professionnelles : le bibliothécaire doit dépasser la culture du catalogue pour se mettre au service de l'utilisateur, de ses attentes et de ses besoins sur le web. Les bibliothécaires doivent cerner les pratiques actuelles des internautes et comprendre les motifs qui les incitent à passer par un moteur de recherche pour parvenir sur le site d'une institution culturelle ou patrimoniale. La valorisation des contenus ne consiste plus, selon ces professionnels, à mettre en exergue des outils de recherche classiques (notices, catalogues, etc.) mais à mutualiser et à rendre interopérables un ensemble de ressources authentiques et intéressantes dans des dispositifs ou des réseaux dédiés aux pratiques et centres d'intérêt des internautes.

Bibliographie

- Azouz K. (2014) : Les bibliothèques et la visibilité des données culturelles sur le web : réalité d'un défi. *Actes du 6e colloque spécialisé en Sciences de l'Information - Information, communication, documentation : les nouvelles utopies*, Institut d'Administration des Entreprises et Université de Poitiers, Université de Moncton, Canada, 17-18 juin, pp. 16-29.
- Bermès E., et al. (2013). *Le web sémantique en bibliothèques*. Paris : Collection Bibliothèques, éditions du Cercle de la Librairie
- Bonaccorsi J.(2009), *Le devoir de lecture. Médiations d'une pratique culturelle*. Paris, Lavoisier-Hermès, coll. Communication, médiation et construits sociaux,, 218 p.
- Boustany J., Broudoux E., Chartron G (2013). La médiation numérique : renouvellement et diversification des pratiques. Actes du colloque Document numérique et société, Zagreb, 2013. *collection Information & stratégie*. Editions de Boeck. Bruxelles, 2014. p.7
- Calenge, B. (2015). *Les bibliothèques et la médiation des connaissances*. Paris : Éd. Du Cercle de la librairie, 147 p
- Galaup X. (2012). Développer la médiation documentaire numérique .Presses de l'enssib
- Ihadjadene M., Chaudiron S. (2008). Quelles analyses de l'usage des moteurs de recherche ? *Questions méthodologiques. Questions de Communication*, Presses universitaires de Nancy, pp.17-32.
- Davallon, Jean, (2003). La médiation : la communication en procès ? *Médiation et information (MEI)*. N° 19. pp. 38-59
- Dufrène B., et al. (2013) *Numérisation du patrimoine: Quelles médiations? Quels accès ? Quelles cultures ?* Paris : Collection Cultures numériques, éditions Harmattan.
- Fabre, I. (2013). L'espace documentaire comme lieu de médiations. Esquisse. ISSN 1925-4873
- Kaufmann J.C. (2014). *L'enquête et ses méthodes : l'entretien compréhensif*. Paris. Armand Colin
- Liquète V.,et al (2010). Faut-il reconsidérer la médiation documentaire?, *Les Enjeux de l'information et de la communication*, pp. 43-57.
- Paugam S. (2010). *L'enquête sociologique*. Paris, PUF.
- Rueda A. (2010). Des médias aux médiations : quelles médiations, quels objets, quels enjeux?», *Les Enjeux de l'information et de la communication*, pp. 88-103.
- Vitalli Rosatti M.(2011), La philosophie du numérique : lecture de pour un humanisme numérique de Milad Doueïhi, *Sens Public*, pp.1-8