Google Livres et la numérisation : quels impacts pour les bibliothèques numériques ?

Laurent Brosset

▶ **To cite this version:**

Laurent Brosset. Google Livres et la numérisation : quels impacts pour les bibliothèques numériques ?. Sciences de l'information et de la communication. 2016. dumas-01376071

HAL Id: dumas-01376071

https://dumas.ccsd.cnrs.fr/dumas-01376071

Submitted on 7 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.

Google Livres et la numérisation : quels impacts pour les bibliothèques numériques ?

BROSSET Laurent

UFR Langage, lettres et arts du spectacle, information et communication

Mémoire de master 2 recherche – Mention Information et communication

Parcours : Recherche et études en information-communication

Sous la direction de CLAVIER Viviane

Année universitaire 2015-2016

Remerciements :

Je tiens à remercier ma tutrice Mme Viviane Clavier qui m'a aidé quant à l'élaboration de mon plan, et m'a permis d'obtenir de précieuses références bibliographiques.

Je remercie également Céline Brosset, ma sœur, pour son soutien et son travail de relecture.

Résumé :

Résumé:

Depuis une vingtaine d'années, les bibliothèques accordent de plus en plus d'importance au numérique.

En France, la Bibliothèque nationale de France (BnF) lance, en 1997, Gallica, sa propre bibliothèque numérique. Différentes bibliothèques numériques voient le jour et se développent aux Etats-Unis et en Europe dans les années 90.

De nos jours, le numérique est une ressource stratégique pour les bibliothèques et ces dernières doivent s'adapter à ce nouvel environnement.

L'année 2004 marque l'arrivée de l'entreprise privée Google dans le monde de la numérisation. Avec son projet Google Livres, ce phénomène a accéléré, modifié, renforcé et dynamisé l'organisation des bibliothèques, ainsi que leur rapport avec le patrimoine numérique.

De par son statut d'entreprise privée, son capital financier (première capitalisation boursière des Etats-Unis en Février 2016), ses méthodes, les critiques à son égard (situation de monopole, pratiques anticoncurrentielles, non-respect de la vie privée…), et la finalité de son projet (devenir le plus grand index de livres numériques au monde), Google Livres ne cesse de susciter de nombreuses controverses sur le plan financier, juridique, idéologique et bibliothéconomique.

Nous nous intéressons ici aux liens forts entre Google Livres et les bibliothèques numériques, aux évolutions des bibliothèques numériques, mais également aux intérêts, plus ou moins visibles, que peut avoir l'entreprise privée, qui se donne pour mission « d'organiser l'immense volume d'informations disponibles sur le Web et dans le monde », à investir le secteur de la numérisation de livres.

Pour finir, la comparaison entre les sites de Google Livres et de Gallica nous permet de constater que les enjeux de Google, présentés comme philanthropes, sont finalement bien différents des enjeux patrimoniaux propres aux bibliothèques numériques.

Mots-clés: bibliothèque numérique, patrimoine, livre, numérisation, Google, médiation numérique, Gallica.

Abstract:

Since about twenty years, libraries are giving more and more importance to the digital technology.

In 1997, in France, The National Library of France launched Gallica, its digital library. During the 90'ies, various digital libraries were created, in Europe and in the United States.

Nowadays, the digital technology is a strategic resource for libraries, which have to adapt themselves to this new environment.

2004 represents the arrival, in the world of digitization, of the private company Google. With its Google Books project, this phenomenon has accelerated, modified, strengthened, and energized the organization of libraries. Moreover, it transformed their relationship to the digital heritage.

Due to its status of private company, which included its financial capital (first market capitalization of the United States in February 2016), its methods, its critics (monopolist position, anti-competitive practices, breach of privacy…), and the purpose of its project (to become the largest digital library of the world), Google Books doesn't stop arousing numerous controversies on the financial, legal, ideological and librarianship areas.

We are here, as interested in the strong links between Goggle books and digital libraries, as in the digital libraries developments. We are also concerned by the interests, more or less visible, that this private company can pursue, which gives itself for mission "to organize the huge amount of information available on the Web and in the world." Indeed, we can rightfully ask why Google wants to invest the digitalization of books's area.

To conclude with, the comparison between Goggle Books and Gallica's websites, allows us to notice that Google's issues, which are presented as philanthropists, are actually very different from the specific patrimonial stakes of digital libraries.

Key words: Digital library, patrimony, book, digitalization, Google, digital mediation, Gallica.

Table des matières

Introduction : 7

Partie I : Le développement de Google et de Google Livres 13

1. Le géant Google 13
   1.1 Qu'est-ce que Google ? 13
   1.2 Google et son mode de rémunération 15

2. Le développement de Google Livres 18
   2.1 Au commencement était Google Livres 18
   2.2 La librairie Google Play Livres 22

3. Un projet très controversé 23
   3.1 De nombreux démêlés avec la justice 23
   3.2 Le cas des œuvres orphelines 30
   3.3 Quelques critiques célèbres et alternatives 31

Partie II : Google Livres face aux bibliothèques numériques 36

1. Bibliothèques numériques, quelle(s) définition(s) ? 36
   1.1 Les bibliothèques numériques : une notion encore floue 36
   1.2 Trois modèles pionniers de bibliothèques numériques 39

2. L'impact de Google Livres sur les bibliothèques numériques 42
   2.1 Les contours des trois modèles pionniers redessinés 42
   2.2 L'arrivée de Google Livres : quel(s) impact(s) pour les bibliothèques numériques ? 44
   2.3 Google Livres est-il une bibliothèque numérique ? 46

3. Pourquoi Google a t il investit le secteur de la numérisation ? 51
   3.1 La numérisation : quel(s) intérêts(s) pour Google ? 51
   3.2 Google Livres, une composante supplémentaire de « l'empire Google » 55

Partie III : Analyse des sites Google Livres et Gallica 57

1. Pourquoi comparer Google Livres et Gallica ? 57

2. Analyse des sites 60
   2.1 La page d'accueil 60
   2.2 Les « espaces personnels » 63
   2.3 Les visualiseurs 67
   2.4 La fonction recherche 73

3. Gallica et l'éditorialisation 76
   3.1 Les collections Gallica 76
   3.2 Gallica et les réseaux sociaux 79
   3.3 Gallica et le site « data.bnf » 81

4. Conclusion de l'analyse 82

Conclusion générale 84

Bibliographie 88

Annexes 93

Introduction :

L'apport des nouvelles technologies, et en particulier du numérique, modifie notre perception du document.

Le numérique représente, depuis une vingtaine d'années, un enjeu majeur pour les institutions patrimoniales. Il s'agit, désormais, d'une activité stratégique pour les bibliothèques. La numérisation du patrimoine apparait, d'abord, comme une technique de reproduction de documents imprimés, puis d'œuvres d'art.

En ce sens, la Bibliothèque Nationale de France pratique la numérisation depuis 1992. Avec le temps, le perfectionnement des techniques de numérisation et son essor, les enjeux de la numérisation du patrimoine se sont peu à peu dessinés et transformés. Un tournant majeur est né dans le monde des bibliothèques avec l'arrivée de Google Livres.

En 2004, Google fait part de sa volonté de s'impliquer dans la numérisation de fonds patrimoniaux. L'objectif est colossal : devenir la plus grande bibliothèque numérique du monde. Cette aspiration grandiose se situe dans la lignée directe des rêves du Belge Paul Otlet, né en 1868, et connu pour être le père de la Classification Décimale Universelle (avec Henri La Fontaine, prix Nobel de la Paix en 1913).

Dès 1895, Paul Otlet souhaite réunir les notices de la totalité des ouvrages publiés sur l'intégralité des sujets de chaque époque, afin d'élaborer le Répertoire Bibliographique Universel (RBU). Dans le but de partager la connaissance, et de permettre aux hommes de se connaitre et de vivre en harmonie, Paul Otlet imagine répertorier l'ensemble des œuvres produites par l'humanité. De cette manière, il aspire à constituer une mise en réseau mondiale des bibliothèques. Plus tard, en 1934, Otlet imagine une forme de moteur de recherche analogique avant l'heure, le Mundaneum, qui contiendrait le Répertoire Bibliographique Universel (RBU), et serait capable de répondre aux requêtes des utilisateurs : « on peut

imaginer le télescope électrique, permettant de lire de chez soi des livres exposés dans la salle *teleg* des grandes bibliothèques, aux pages demandées d'avance. Ce sera le livre téléphoté.[1]».

Otlet, en envisageant un monde sans livres matériels, et en préfigurant également d'une certaine manière l'Internet et les moteurs de recherche, dessine, dès 1934, la bibliothèque à vocation universelle qu'aspire à devenir Google de nos jours.

Afin d'alimenter son projet de numérisation, la société réputée pour son moteur de recherche du même nom, a rapidement signé des accords avec plusieurs grandes bibliothèques Américaines (Michigan, Harvard, New York Public Library …), puis Européennes. Google, grâce à son capital technique, intellectuel et financier considérable, se positionne aujourd'hui comme l'« index de livres complets le plus fourni au monde » (mention qui apparait sur la page d'accueil de Google Livres).

Le champ d'activité de Google Livres s'est progressivement élargi, avec notamment le développement d'innombrables partenariats, (à l'histoire mouvementée), conclus avec de nombreux auteurs, éditeurs et institutions. Cette extension s'est renforcée grâce à l'ouverture d'une librairie numérique via la boutique Google eBookstore aux Etats-Unis, en 2010, qui, en 2012, deviendra Google Play.

De nos jours, Google Livres est un service hybride qu'il est difficile de qualifier, se situant entre un entrepôt, une bibliothèque numérique, et une librairie.

En effet, l'outil permet la lecture intégrale et le téléchargement d'ouvrages libres de droit, la consultation d'extraits (snippets) de livres sous droits, l'achat de livres via une boutique (Google Play), mais également la possibilité de faire des recherches-intra texte, et la constitution de collections personnelles.

Le développement de Google Livres est source de nombreuses interrogations. Parmi elles, on retrouve des questions sur les enjeux de politiques culturelles, ainsi que concernant les nouvelles formes de socialisation de l'information.

[1] http://usbek-et-rica.fr/paul-otlet-linventeur-du-google-de-papier/ [Consulté le 08/06/2016]

Avec le numérique, le patrimoine change de nature. En effet, la numérisation fait de l'objet patrimonial un « objet informationnel, trivial, dont la circulation nécessite une indexation, voire des traductions et autorise des adaptations »[2]. Dans le cadre de Google Livres, la notion de transformation de l'objet livre peut s'avérer pertinente, et ce notamment au vu des fonctionnalités que propose Google Livres. Grâce à ses algorithmes et ses modalités de traitement automatique (calculs statistiques etc), ces dispositifs éloignent l'utilisateur d'une expérience de lecture « traditionnelle ».

Au-delà de la question de la nature du document, se pose également la problématique économique. Le « raz-de-marée » Google, comme nous le verrons par la suite, a été vecteur de bouleversements majeurs vers une économie au sein de laquelle le partenariat privé – public (jusqu'alors peu pensable dans le domaine des bibliothèques, qui était avant tout une histoire de service public), domine. Ce phénomène nous amène à nous demander dans quelle mesure il s'agit d'un véritable partenariat, et non pas d'une l'hégémonie d'un groupe privé sur un marché[3] ?

L'arrivée de Google dans le monde des bibliothèques numériques a soulevé de très nombreuses controverses et batailles juridiques.

Si les moyens financiers et techniques dont dispose la fameuse société de Mountain View peuvent être perçus, de prime abord, comme une aubaine pour les bibliothèques, leur permettant ainsi d'entretenir, de mettre en valeur et réactualiser leurs fonds, nous verrons que les critiques à l'égard de Google Livres concernent donc plusieurs pans : juridiques (territorialité du droit d'auteur, violation du droit d'auteur, appropriation des œuvres orphelines), techniques, financières, idéologiques et bibliothéconomique

Malgré ces réactions, nous constatons que le projet Google Livres n'a pas cessé d'être l'acteur d'une multitude d'accords et de partenariats conclus avec des bibliothèques, des éditeurs, ainsi qu'avec des auteurs, et ce partout dans le monde. On notera que les clauses de

[2] DUFRÊNE Bernadette et IHADJADENE Madjid et DRUCKMANN Denis (dir), *Numérisation du patrimoine, Quelles médiations ? Quels accès ? Quelles cultures ?,* Herman Editeurs, Paris, 2013, p8.

[3] *Ibid.*

ces contrats sont, pour la plupart, privées, ce qui semble constituer la « norme » avec la société Google, qui aime entretenir le « culte du secret » qui l'entoure.

Les nombreuses batailles juridiques qui ont largement résonné dans les médias internationaux pendant 10 années paraissent apaisées concernant le volet Google Livres. Le groupe Google devrait par ailleurs recevoir d'ici l'été le verdict de la Commission de Bruxelles d'ici l'été 2016, après six d'enquêtes et de négociations. La presse parle déjà d'une amende de trois milliards d'euros pour abus de position dominante[4] tandis que d'autres enquêtes sont toujours ouvertes sur les dessous d'Android et sur les pratiques publicitaires de l'entreprise.

Dans ce cadre, il nous semble donc intéressant, de nous interroger sur les intérêts que peuvent avoir les bibliothèques à pactiser avec l'entreprise privée, mais également, et surtout, quant à l'existence même du projet Google Livres.

Fréquemment accusée d'optimisation fiscale, de pratiques anticoncurrentielles et de gestion opaques des données, la société Google se qualifie pourtant de philanthrope. Elle n'en demeure pas moins une entreprise commerciale, dont l'emprise sur la population mondiale, par le biais de ses différents services, ne cesse de grandir.

Quels sont les enjeux derrière le projet Google Livres pour l'entreprise ayant la première capitalisation boursière des Etats-Unis en Février 2016 ?

Nous verrons que Google Livres constitue, entre autre, un espace de choix pour du placement publicitaire (principale source de revenus de Google), mais aussi un outil idéal pour renvoyer vers sa librairie (Google Play Livres), ou vers ses partenaires. Google Livres représente également une énorme base de données multilingue, pouvant servir d'outil de travail pour l'amélioration des algorithmes relatifs à d'autres services de la société (moteur de recherche, outil de traduction).

Plus largement, la politique de Google Livres s'ancre dans le développement et l'agrandissement de « l'empire Google », et de sa mission autoproclamée « d'organiser la connaissance du monde ».

[4] NORA Dominique, « Le vrai pouvoir de Google en France», Le Nouvel Observateur, Juin 2016, p37.

Ainsi, pour englober toutes ces interrogations, nous proposons la problématique suivante : Dans quelles mesures l'arrivée de Google dans le secteur de la numérisation d'ouvrages a impacté le monde des bibliothèques numériques à des niveaux techniques, financiers, idéologiques et bibliothéconomiques ?

Nous émettons également les hypothèses suivantes :

- L'arrivée de Google dans la numérisation a amené les bibliothèques numériques à évoluer, se modifier mais aussi à affirmer leur identités.
- Si Google Livres a eu et a toujours un impact déterminant dans le développement et l'évolution des bibliothèques numériques, le projet ne peut pas pour autant être qualifié de « bibliothèque numérique », son organisation et ses objectifs étant bien distincts de ceux des institutions patrimoniales citées précédemment. L'analyse son site et de celui de Gallica, bibliothèque numérique de référence, permettra de mettre en lumières les motivations divergentes des deux projets. De cette hypothèse découle deux sous-hypothèses :

    a) Google Livres cherche à renvoyer le plus rapidement possible l'utilisateur vers sa librairie Google Play Livres.
    b) Google Livres ne possède pas de politique documentaire, contrairement aux bibliothèques numériques, qui, à l'heure du web 2.0, des moteurs de recherche et de la dématérialisation des documents, trouvent leur légitimité et leur autorité dans leur travail de médiation numérique et d'éditorialisation.

- La numérisation est pour Google une source de rémunération évidente. En plus de la librairie Google Play Livres, l'entreprise trouvent d'autres intérêts moins visibles dans son projet de numérisation, bien éloignés ses discours utopiques et philanthropes. De cette hypothèse découle deux sous-hypothèses :

    a) Google avait dès l'origine du projet en 2004, l'idée de lancer sa librairie numérique Google Play Livres.
    b) Les intérêts de Google pour la numérisation se situent dans la droite lignée des stratégies commerciales de l'entreprise (soft power, capitalisme politique, analyse de données personnelles etc…).

Dans une première partie, nous nous intéresserons à Google, son histoire, son modèle financier et sur le développement du projet Google Livres, des origines au lancement de la librairie Google Play Livres, en passant par les nombreux démêlés avec la justice et les controverses qu'a rencontré le projet au cours de son évolution.

Nous nous focaliserons ensuite sur les bibliothèques numériques, en tentant de les définir, et en positionnant le statut de Google Livres comparé à ces dernières. Google Livres est-il une bibliothèque numérique ? Nous verrons que non, mais le projet de l'entreprise privée et intrinsèquement lié aux évolutions des bibliothèques numériques qui ont vu leurs contours se redessiner avec l'arrivée de Google dans le secteur de la numérisation. Nous nous interrogerons également sur les intérêts de Google à investir ce secteur.

Enfin, dans une troisième et dernière partie, nous comparerons les sites de Google Livres et de Gallica, la bibliothèque numérique de la BnF pour mettre en lumière les enjeux divergents des deux projets.

Partie I : Le développement de Google et de Google Livres

1. Le géant Google

1.1 Qu'est-ce que Google ?

La société Google a été fondée en 1998 par Lawrence E. Page (Larry Page) et Sergey M. Brin, dans la Silicon Valley à Mountain View (Californie). Principalement connu pour son moteur de recherche du même nom, l'organisme a largement diversifié ses activités et a multiplié ses acquisitions au fil des années.

En complément du fameux moteur de recherche Google réputé pour sa position monopolistique, la société étend rapidement ses activités et investie le monde des technologies en général, via le développement et le rachat de nombreux logiciels, sites web, applications et technologies. Parmi les évolutions marquantes qui ont contribué à la reconnaissance de Google, nous pouvons citer les lancements de Google News en 2002 (qui alimentera de nombreuses polémiques), et du système messagerie électronique Gmail le 1er Avril 2004 (aux capacités de stockage bien plus importantes que celles des concurrents de l'époque), mais aussi l'inauguration de l'arrivée de Google dans le monde de la cartographie numérique avec Google Maps aux Etats Unis, en Grande Bretagne et au Canda le 08 février 2005, et qui s'étend au reste du monde, et qui connait par la suite de nombreuses mises à jour.

On peut également ici évoquer le lancement de Google Traduction en 2006, ou encore le rachat et le lancement, du système d'exploitation Android qui est, en 2015, le plus utilisé dans le monde avec plus de 80 % de parts de marché dans les smartphones.

Depuis le 04 aout 2015, suite à une restructuration de Google, les différentes sociétés précédemment détenues par cette dernière sont regroupées à la manière d'un conglomérat de filiales, sous une entreprise dénommée Alphabet.

De cette manière, la société Google est devenue une filiale du groupe Alphabet, se focalisant uniquement sur les activités concernant Internet et l'informatique (le moteur de recherche Google, YouTube, Android, Google Play, Google Livres, Google Ads etc...). Les différentes autres filiales d'Alphabet sont les filiales Nest (domotique et objets connectés), Life Sciences (Santé et lutte contre le vieillissement), Fiber (fournisseur d'accès à Internet), X (laboratoire multiprojets à long terme), Capital (finance), et Ventures (investissements).

Alphabet est l'une des premières entreprises américaines et mondiales en termes de valorisation boursière. Google fait partie du « Big Four » (les « quatre gros ») des entreprises de technologie aux côtés de Facebook, Apple, et Amazon.com. On notera que l'expression « GAFA » est fréquemment utilisée pour décrire ces organisations et leur impact sur le monde des technologies. En Février 2016, Alphabet, la société mère de Google, dépasse Apple en bourse, et devient ainsi la première capitalisation boursière des Etats-Unis, avec un total de 550 milliards de dollars dispersés au travers de ses différentes catégories d'actions.

La mission que s'est donnée Google est d'organiser l'information à l'échelle mondiale, et de la rendre universellement accessible et utile (« Google's mission is to organize the world's information and make it universally accessible and useful »[5]). Cette volonté philanthropique autoproclamée est résumée dans le slogan de la société : « Don't be Evil » (ne soyez pas malveillant, ne faites pas le mal).

Outre les critiques en lien avec leur position dominante et leur capital financier (pratique anticoncurrentielle, optimisation fiscal), la société Google et ses activités sont aussi contestées principalement dans trois domaines :

- Le respect de la vie privée, mis à mal par la myriade de données récoltées sur les utilisateurs (sites consultés, scan des boites mail, termes des recherches, positions géographiques etc...)
- La censure, avec l'occultation de masses de données opérée suite à la demande de certains gouvernements comme la Chine sur le massacre Tian'anmen par

[5] http://www.google.com/about/company/ [consulté le 08/06/2016]

exemple.

- Le respect du droit d'auteur, qui constitue l'une des critiques majeures faites au projet Google Livres.

1.2 Google et son mode de rémunération

Comme nous l'avons vu précédemment, la valorisation boursière de Google est énorme.

Google, à l'instar des réseaux sociaux, des moteurs de recherches et des firmes multimédia comme Apple, est pleinement intégrée et même représentative de ce que l'on nomme le « capitalisme cognitif ». Le capitalisme cognitif s'éloigne de la sérialisation et de la reproduction industrielle pour accumuler des informations, de la connaissance et de la créativité, et donc des formes d'investissement immatérielles[6].

Selon Yann Mouliet Boutang, les principaux facteurs du développement de cette économie de la connaissance sont les suivants[7]:

- La révolution du langage informatique.
- La montée continuelle de la capacité de traitement et de stockage de données numérisées par les ordinateurs.
- La réduction drastique des coûts d'acheminement de l'information (fibre optique…).
- La miniaturisation (nanotechnologie).
- L'appropriation des TIC par un nombre croissant de cerveaux humains.
- L'existence de réseaux informatiques reliés entre eux par au moins un réseau des réseaux, l'Internet.

Le capitalisme cognitif se sert de la codification de la connaissance, de la gestion

[6] DUFRÊNE Bernadette et IHADJADENE Madjid et DRUCKMANN Denis (dir), *Numérisation du patrimoine, Quelles médiations ? Quels accès ? Quelles cultures ?,* Herman Editeurs, Paris, 2013, p32.

[7] *Ibid.*

d'informations dans des bases de données qui sont exploitées pour cibler des publics.

Google, par exemple, via la multiplicité de services gratuits qu'il fournit (Gmail, Google Maps…), récolte des données sur les utilisateurs de ces services (goûts, connaissances, réseaux de relation)[8]. Les entreprises sont friandes de ces informations qui leur donnent des repères sur des publics d'acheteurs potentiels. Ainsi, c'est dans ce contexte, que « Google devient le portique indispensable de la mise en marché de toute marchandise immatérielle »[9].On notera que c'est avant tout à travers la publicité, que Google exploite ces données personnelles.

Les revenus de Google proviennent en grande majorité de la publicité, et plus précieusement de la publicité contextuelle[10] (liens sponsorisés dans les résultats de recherche, facturation au « coût par clic » pour les annonceurs etc...). Le modèle économique de Google repose sur la gratuité de plusieurs de ses services pour les internautes.

Le système AdWords, lancé en 2002, permet à Google de « vendre aux enchères » des mots clés à des annonceurs. AdWords permet la création d'annonces personnalisées (titre, description) répondant à certains mots clés définis. Lorsqu'un internaute fait une requête sur un moteur de recherche se rapprochant d'un des mots-clés enregistrés par l'annonceur, une annonce publicitaire est susceptible d'apparaitre dans les résultats de la recherche, dans un encart spécifique dédié aux liens commerciaux. Si l'internaute clique sur l'annonce en question, l'annonceur reverse une certaine somme à Google. Il s'agit du « coût par clic ». L'ordre des liens commerciaux affichés est déterminé par plusieurs facteurs : les budgets alloués aux annonces par les utilisateurs, et le niveau de qualité de l'annonce (déterminé par les algorithmes secrets de Google, qui prennent probablement en compte des critères comme le taux de clics courant de l'annonce, la pertinence du texte de l'annonce par rapport au mot-clé et à la page de destination, l'historique de performance du mot-clé, l'historique de

[8] DUFRÊNE Bernadette et IHADJADENE Madjid et DRUCKMANN Denis (dir), *Numérisation du patrimoine, Quelles médiations ? Quels accès ? Quelles cultures ?,* Herman Editeurs, Paris, 2013, p36.

[9] *Ibid.*

[10] JACQUESSON Alain, *Google Livres et le futur des bibliothèques numériques*, Editions du Cercle de la Librairie, Paris, 2010, p13.

performance du compte etc…[11]).

Si AdWords permet d'acheter du trafic, Google a également mis en place un système permettant de vendre du trafic : Google AdSense. Il s'agit d'une solution stratégique forte pour élargir son domaine : un usager passant en moyenne 5% son temps sur les moteurs de recherche contre 95%[12] sur les sites de contenu, Google a vite souhaité investir ces derniers.
Le système AdSense for Search permet à un site web d'accueillir un moteur de recherche offrant la technologie de Google et, par extension, les publicités AdWords suivant les mots-clés sur ses pages web. Google reverse une partie de ses gains, toujours au coût par clic, à ce site web, le partage étant de 51 % pour l'éditeur, et 49 % pour Google. De même, le système AdSense for Content permet à un éditeur d'afficher les encarts publicitaires de Google sur son site web (ces annonces sont ciblées en fonction du contenu de la page sur laquelle se trouve l'encart, mais aussi en fonction de ce que Google a enregistré du comportement de l'internaute). Les revenus générés sont répartis à hauteur de 68 % pour l'éditeur, et de 32 % pour Google[13].

En 2007, Google a acquis la régie publicitaire Doubleclick pour 3,1 milliards de dollars. Il s'agissait alors de l'une des plus importantes entreprises mondiales en matière de vente d'espaces publicitaires, d'infrastructures, et de services pour les annonceurs. Le but de la manœuvre était de renforcer les activités de Google de vente directe et d'intermédiation sur le marché de la publicité non liée à la recherche.

Google souhaitait ainsi toucher également les agences de publicités, en leur donnant la possibilité de vendre à leurs clients la conception et la diffusion d'annonces sur les réseaux d'annonceurs partenaires de Google, et ce à grande échelle. Avec DoubleClick, Google est

[11] http://www.journaldunet.com/ebusiness/crm-marketing/adwords-les-bonnes-feuilles/encheres-automatiques-et-encheres-manuelles.shtml [Consulté le 08/06/2016]

[12] http://www.journaldunet.com/ebusiness/crm-marketing/adwords-les-bonnes-feuilles/encheres-automatiques-et-encheres-manuelles.shtml [Consulté le 08/06/2016]

[13] http://www.inaglobal.fr/communication-publicite/article/publicite-en-ligne-google-peut-il-cumuler-les-metiers-8236 [Consulté le 08/06/2016]

devenu fournisseur d'infrastructures (DoubleClick Ad Exchange et le réseau Google Display[14], constitué de sites web ayant formé un partenariat avec la régie Google (AdSense) pour afficher, avec contrepartie monétaire, des publicités ciblées sur leurs pages), mais également intermédiaire commercial, en effectuant le lien entre éditeurs et publicitaires.

De nos jours, Google est présent (depuis le début des années 2010), dans tous les secteurs du marché de la publicité en ligne : liens sponsorisés, publicités contextuelles, bannières, vidéo (YouTube), vente directe et intermédiation, ordinateurs, mobile et tablettes[15]. L'une de ces manifestations est, par exemple, la présence de publicités contextuelles sur la messagerie électronique Gmail.

Si la publicité génère une grande partie des ressources de Google, l'entreprise dispose également d'autres services marchands, comme des espaces de stockage en ligne sur les serveurs de Google, la version payante de Google Apps, ou encore le Google Play Store, un magasin en ligne faisant entre office de librairie (Google Play Livres), point sur lequel nous reviendront par la suite.

2. Le développement de Google Livres

2.1 Au commencement était Google Livres

Google a manifesté très tôt un vif intérêt pour les bibliothèques. Lorsque nous consultons l'onglet « à propos » du site Google Livres, la première phrase est en effet « In the beginning, there was Google Books »[16] (au commencement, était Google Livres). Si cette

[14] https://www.google.fr/ads/displaynetwork/ [Consulté le 08/06/2016]

[15] http://www.inaglobal.fr/communication-publicite/article/publicite-en-ligne-google-peut-il-cumuler-les-metiers-8236 [Consulté le 08/06/2016]

[16] http://books.google.com/googlebooks/about/history.html [Consulté le 08/06/2016]

phrase d'accroche n'est pas tout à fait exacte (le texte se poursuit par « well not exactly »), l'intérêt de Sergey Brin et Lary Page pour les bibliothèques se manifeste très rapidement. Au cours de leur carrière universitaire à l'université de Stanford (1996-1997), les fondateurs de Google avaient déjà travaillé sur la gestion et la constitution de fonds numériques pour une bibliothèque (Stanford Digital Library Technologies Project).

Le moteur de recherche Google, en projet depuis 1996 et initialement sous le nom de BackRub, dont l'objectif était d'analyser les différentes relations entre les sites web afin d'offrir des résultats plus pertinents que ceux affichés par les moteurs de recherche de l'époque (Altavista par exemple), est lancé en 1998, et connaît une ascension fulgurante. En marge de cela, une équipe de Google travaille, dès 2002, sur un projet tenu secret, sobrement intitulé « Books ».

Souhaitant optimiser grandement le temps et la qualité de la numérisation d'ouvrages, l'équipe s'inspire de projets existants, tels que le Projet Gutenberg de Michel Hart, l'American Memory de la Bibliothèque du Congrès, ou encore le Million Book Project de l'Internet Archive, pour mieux appréhender leurs différentes méthodes de travail. Les ingénieurs de l'équipe sont également chargés de trouver une méthode de numérisation puissante, efficace, ne détériorant pas l'état des ouvrages, pouvant prendre en compte les diverses tailles, polices de caractères et imperfection que présentent les livres imprimés, et ce dans 430 langues différentes.

Parmi l'équipe « Books », se trouve Larry Page. Ce dernier a pris contact avec la bibliothèque universitaire du Michigan, son ancienne université. Page affirme souhaiter numériser sept millions d'ouvrages sur une période de six ans, en prenant en charge le financement du projet, à condition d'intégrer le contenu des livres aux index de Google. Il s'agit là, du premier contact de Google avec une bibliothèque. D'autres accords ont été conclus avec la bibliothèque de Stanford (projet de numériser 9 millions d'ouvrages), d'Harvard (la Harvard University Librairy, la plus grande bibliothèque universitaire du monde), la Bibliothèque bodléienne de l'Université d'Oxford (pour numériser une collection d'un million d'ouvrages du XIXème siècle tombés dans le domaine public), et la New York Public Library (pour numériser 16 millions de livres).

Conformément à la politique commerciale de Google, ces partenariats font preuves d'une grande confidentialité. En effet, les entreprises traitant avec Google sont obligées de signer un accord de non divulgation (Non Disclosure Agreement), rendant ces accords secrets, ce qui constitue une pratique inhabituelle dans le milieu des bibliothèques.

Autre point crucial, les bibliothèques ont dû accepter une clause sur l'exclusivité commerciale de Google concernant les fonds numérisés (limitée à une certaine durée), ainsi que sur l'exclusivité d'indexation. Cette restriction est importante pour Google, car elle empêche les moteurs de recherche concurrents (Bing, Yahoo!) de pouvoir indexer ces contenus. L'exclusivité donne également à Google le choix de ce qu'il diffuse ou non. Il s'agit d'un pouvoir potentiel sur l'accès au savoir : « elle ouvre [l'exclusivité] la possibilité de pratiquer, comme peut le faire un monopole, des tarifs qui excluraient certains de l'accès à ce patrimoine numérisé. Ce qui est à l'opposé de la tradition de la lecture publique »[17].

Ces clauses sont ainsi formulées dans les contrats entre Google et les bibliothèques de l'Université de Michigan et de l'Université de Californie[18] :

« University shall implement technological measures (e.g. through use of the robots.txt protocol) to restrict automated access to any portion of the University Digital Copy or the portions of the University Website on which any portion of the University Digital Copy is available».

Certaines bibliothèques, comme la Bibliothèque Municipale de Lyon, ont réussi à négocier ces clauses.

Google va également aller à la rencontre d'éditeurs. En effet, Larry Page et Sergei Brin ont loué, en Octobre 2004, un stand à la Foire du livre de Francfort en Allemagne (Buchmesse)

[17] DUFRÊNE Bernadette et IHADJADENE Madjid et DRUCKMANN Denis (dir), *Numérisation du patrimoine, Quelles médiations ? Quels accès ? Quelles cultures ?,* Herman Editeurs, Paris, 2013, p20.

[18] https://scinfolex.com/2012/12/15/numelyo-la-bibliotheque-numerique-de-lyon-exister-a-lombre-de-google/ [Consulté le 08/06/2016]

- qui est le plus grand rendez-vous annuel mondial sur le thème de l'édition -, dans le but de présenter leur projet Google Prints for Editors. Google souhaite « aider les éditeurs à vendre leurs ouvrages en ligne en indexant gratuitement le contenu de ceux-ci et en les incluant dans leurs services en ligne »[19].

Le 14 décembre 2004, Google annonce publiquement, par le biais d'un communiqué de presse, l'existence de son projet Google Print for libraries. L'entreprise déclare ainsi, dans un premier temps, la numérisation d'ouvrages issus des fonds des cinq bibliothèques partenaires (Harvard, Michigan, Stanford, Oxford, New York Public Library), entre 6 et 15 millions sur une période de six ans. Les utilisateurs pourront faire des recherches sur l'intégralité du corpus via son moteur de recherche.

Les retombées médiatiques sont, pour la plupart, très dithyrambiques et la nouvelle est largement traitée dans les journaux du monde entier. Certaines réactions sont beaucoup plus sceptiques, mais nous y reviendront plus tard.

Par la suite, d'autres bibliothèques s'intègrent au projet Google Print for libraries en concluant divers accords. En 2005, la Bibliothèque du Congrès a notamment reçu un don de 3 milliards de dollars de la part de Google, pour aider à la numérisation de ses collections. En 2007, des bibliothèques européennes signent également avec Google, comme la Bibliothèque de Gand (Belgique), la Die Bayerische StaatsBibliiothek (Allemagne), la Bibliothèque Cantonale et Universitaire de Lausanne (Suisse), ou encore la Bibliothèque municipale de Lyon (France). En 2008, 28 bibliothèques sont partenaires du projet Google Print, maintenant renommé Google Books (Google Livres en français).

Bien que l'annonce de Google Print n'ait pas fait beaucoup de bruit à la Foire du livre de Francfort de 2004, Google a annoncé, en Septembre 2009, que 25 000 maisons d'éditions avaient signé avec lui des accords dans le cadre de Google Books[20].

[19] JACQUESSON Alain, *Google Livres et le futur des bibliothèques numériques*, Editions du Cercle de la Librairie, Paris, 2010, p44.

[20] JACQUESSON Alain, *Google Livres et le futur des bibliothèques numériques*, Editions du Cercle de la Librairie, Paris, 2010, p99.

D’innombrables nouveaux accords ont alors été conclus avec des bibliothèques, des éditeurs et des auteurs. Aujourd'hui, sur la page d'accueil de son site, Google Livres annonce « Effectuez des recherches dans l'index de livres complets le plus fourni au monde ».

2.2 La librairie Google Play Livres

Le 15 octobre 2009, à l’occasion de la 61ème Foire de Francfort, Google annonce son intention de devenir éditeur, et de commercialiser directement des livres électroniques pouvant être lus sur tout appareil disposant d’un navigateur web. Ce phénomène marque le début de Google Editions, vite renommé Google eBookstore (lancé d’abord aux Etats-Unis le 06 décembre 2010), puis Google Play Livres (le 6 mars 2012), qui se place ainsi en concurrent direct d’Amazon et de son Kindle, et d’Apple et son iBooks.

Google Play Livres désigne donc une librairie numérique s’intégrant dans Google Play, une boutique en ligne créée le 6 mars 2012 par Google, et qui résulte d’une fusion entre les services Android Market, Google Movies, Google ebookstore, et Google Music. Elle se compose d’une boutique d'application pour le système d'exploitation Android, d’une boutique de location de films et de séries télévisées, d’une boutique d'achat de musiques, de livres, de magazines, et de la gamme de smartphones et tablettes « Nexus ». Une adresse Gmail (le service de messagerie gratuit proposé par Google), est nécessaire pour utiliser le service Google Play.

Google Play Livres propose un catalogue de plus de 5 millions d’ouvrages, fournis par des éditeurs commerciaux ou tombés dans le domaine public (qui sont alors gratuits), disponibles dans la langue du pays concerné, ainsi qu'en anglais. Les livres achetés sont lisibles via les navigateurs Windows, Linux et Mac OS X, ou via l’application Play Livres sur les smartphones et tablettes (Android 2.2 minimum), et sur les appareils Apple (iOS 4.0 minimum). En plus de cela, les utilisateurs peuvent également télécharger les ouvrages achetés au format PDF ou ePub (acronyme de Electronic Publication, un format ouvert

standardisé pour les livres numériques).

L'application Play Livres propose différentes fonctionnalités pour optimiser les conditions de lecture. La fonction « texte flottant » permet de personnaliser la forme du texte en proposant différents choix comme les modes « jour » (texte noir sur fond blanc), « nuit » (texte blanc sur fond noir) ou « sepia » (texte noir sur fond beige). On y trouve également le réglage de la luminosité, le choix de la police, de l'alignement du texte, de la taille de la police et de la hauteur de ligne (espace interligne). A contrario du « texte flottant », le choix « pages d'origine » ne permet pas ces ajustements (à l'exception des modes jour, nuit et sépia). En septembre 2012, une mise à jour a ajouté de nombreuses fonctions comme le surlignage, l'ajout et l'édition de notes, la synchronisation des lectures entre plusieurs terminaux, ainsi que l'accès à Google Maps (si l'application est installée sur l'appareil), en sélectionnant une ville ou un lieu, avec la possibilité d'accéder à son article Wikipédia, ou d'effectuer une recherche sur internet via Google.
L'utilisateur peut, de cette manière, ajouter des annotations, obtenir une définition, une traduction, ou encore effectuer une recherche en lien avec les mots ou phrases sélectionnées dans le livre, web ou wiki. Les livres contiennent peut aussi, parfois, posséder un contenu audiovisuel.

3. Un projet très controversé

3.1 De nombreux démêlés avec la justice

Si l'ampleur du projet Google Livres et son expansion sont impressionnants, il est important de préciser qu'il s'agit de l'un des pans les plus controversés de Google. En effet, son histoire est assez mouvementée. Elle a donné lieu à de nombreuses controverses et à de nombreuses batailles judiciaires, aux Etats-Unis surtout, mais aussi partout dans le monde (notamment en France).

Sans ici prétendre à une exhaustivité totale, il nous semble nécessaire de mentionner, dans les grandes lignes, ces polémiques, et particulièrement aux États-Unis, car elles ont, pour la plupart, abouti à des négociations et sont donc indispensables à notre questionnement sur les enjeux et les relations qu'entretient Google Books avec les bibliothèques numériques, les éditeurs et les auteurs.

Dans son projet massif de numérisation, Google a numérisé différents types d'ouvrages : des livres tombés dans le domaine public disponible, des livres couverts par le droit d'auteur mais qualifiés d'« orphelins » (nous y reviendront par la suite), mais aussi des livres couverts par le droit d'auteur versé par les éditeurs aux auteurs, ou aux ayants droits. Google intègre également des livres existant originellement sous forme numérique, couverts par des licences numériques, les « Creative Commons », qui autorisent la reproduction sous certaines modalités.

Google a numérisé une quantité très importante d'ouvrages sous droit d'auteur : plus de 8 millions en Janvier 2010. Seulement, cette numérisation s'est faite sans demande d'accord des ayants droits, et sans même les en informer. Cette méthode a, bien entendu, soulever de nombreuses critiques et attaques en justice. Google a certainement agit de la sorte afin d'éviter de « contacter tous les éditeurs du monde entier, qui eux-mêmes auraient dû refaire un contrat d'édition avec tous les auteurs pour une nouvelle publication destinée au support numérique[21] ». En effet, les coûts et délais auraient été imaginables. Google a donc choisi de tout numériser sans accord préalable, en se proposant de supprimer un ouvrage à la demande de l'ayant droit (opération « Opt-out »). De nombreux éditeurs et auteurs ont donc demandé à Google de retirer leurs ouvrages de Google Books. A l'inverse, d'autres ont souhaité négocier pour obtenir des dédommagements pour les œuvres que Google s'est approprié : il s'agit de l'opération « Opt-in ».

Dans le cadre de son projet de numérisation, Google a estimé être couvert par le « fair use » prévu par le droit américain. Cet « usage loyal », ou « usage raisonnable » apporte des

[21] JACQUESSON Alain, *Google Livres et le futur des bibliothèques numériques*, Editions du Cercle de la Librairie, Paris, 2010, p125.

limitations et des exceptions aux droits exclusifs de l'auteur sur son œuvre[22].

Le fair use concerne notamment la nature commerciale ou non de l'utilisation de l'œuvre, la quantité de la partie utilisée, et les conséquences sur les utilisations potentielles commerciales ultérieures.

Le fair use apporte cinq exceptions au copyright strict : le droit à opérer une copie privée, le droit de citation, le droit au pastiche ou à la caricature, l'usage pour l'enseignement, ainsi que l'usage pour la recherche[23]. C'est bien entendu le droit de citation qui intéresse Google.

Lorsque Google a signé ses premiers accords avec des éditeurs en Octobre 2004, ces derniers permettaient un affichage de 20 % de l'intégralité du texte. Google a vite imaginé un usage étendu à une toute autre échelle (l'ensemble des livres numérisés sous droit d'auteur), en se justifiant via le fair use. Or, une telle proportion (environ 20 % de l'œuvre) est-elle toujours couverte par le droit de citation du fair use ?

Dans le cadre de la numérisation de livres, l'interrogation se pose également. Google considère, bien entendu, son projet comme étant de nature essentiellement philanthropique. Il est bon de noter qu'au lancement de Google Books, sa visée commerciale n'était pas encore clairement identifiable.

Enfin, Google Livres se heurte au droit d'auteur et à sa territorialité. Google est une entreprise américaine, ce qui signifie que le droit américain s'applique au projet Google Livres. Il concerne également des ouvrages étrangers déposés dans des bibliothèques américaines, mais toujours protégés par le droit d'auteur de leur pays d'origine[24]. Google Livres peut être consulté dans le monde entier, et le pays de consultation engendre automatiquement l'application du droit territorial.

C'est en fonction de ces statuts juridiques différents, que l'affichage des ouvrages dans Google Livres diffère. En effet, Google affiche le livre en intégralité lorsqu'il n'est plus couvert

[22] *Ibid.*

[23] PAPY Fabrice et LEBLOND Corinne, « Bibliothèque numérique : la nécessaire médiation », *Communication et Langage,* n°161, 2009, p39.

[24] JACQUESSON Alain, *Google Livres et le futur des bibliothèques numériques*, Editions du Cercle de la Librairie, Paris, 2010, p126.

sous droit d'auteur (date soumise à la territorialité du droit), ou lorsque l'accord des éditeurs ou des auteurs a été validé en amont (c'est le cas des Creative Commons). Dans d'autres configurations juridiques, Google affiche seulement un aperçu limité du livre (négocié avec les ayants droits, et dans l'intention de le rendre « feuilletable » et attractif à l'achat, tel un livre physique), de courts extraits (deux ou trois courts extraits de trois à quatre lignes souvent, cas dans lequel la puissance de Google est particulièrement saisissante, car si un utilisateur fait une recherche et que le terme en question est contenu dans un ouvrage qui affiche uniquement de courts extraits, les courts extraits en question afficheront le terme de la recherche). Dans certains cas, aucun extrait n'est disponible.

Le 20 septembre 2005, une première action en justice contre Google a été intentée aux Etats-Unis pour violation des droits d'auteur. Il s'agit de la « Class Action », un recours collectif de l'Authors Guild (la Guilde des Auteurs, qui comprend plus de 8000 membres), et de plusieurs auteurs indépendants.

Le mois suivant, l'Association of American Publishers (l'Association des éditeurs américains) poursuit également Google en justice pour le même motif de violation de droits d'auteur (il s'agit de deux cas séparés). Ces actions en justice ont été négociées au bout de 3 ans, Google proposant aux deux plaignants un « accord historique ».

Entamé à l'automne 2006 et validé une première fois en Novembre 2009 après de nombreuses concessions des deux parties, le Google Books Settlement (Règlement de Google Recherche de livres) est un document de 320 pages. Cet accord concerne les ouvrages enregistrés aux Etats-Unis et publiés en Australie, au Canada et au Royaume-Unis. Dans cet accord, de nombreux points sont abordés, et Google s'engage, entre autre, à verser 125 millions de dollars d'indemnités aux ayants-droit. L'entreprise s'engage également à financer un registre des droits du livre indépendant (mais administré par Google), afin de rechercher les ayants droits, et de créer une base de données recensant les droits d'auteurs de chaque livre indexé (pour un coût estimé à 34,5 millions d'euros). Pour finir, Google s'engage à rémunérer les auteurs et les éditeurs pour l'usage et la vente de leurs travaux. Le registre sera, ensuite, mis à jour, et financé par les revenus de Google Livres.

Cet accord n'a pas été accepté par toutes les parties, malgré plusieurs amendements. En effet, bien qu'un accord à l'amiable soit trouvé avec l'AAP en 2012 (n'ayant pas besoin de

l'aval de la justice pour être opérant), la Guilde des Auteurs poursuit, quant à elle, son attaque en justice.

L'accord a également été vivement critiqué par de nombreuses personnalités et associations. Le droit américain permet aux personnes concernées par le projet de transmettre une objection au tribunal, ou même de se porter plaignant. C'est ce qu'on fait d'innombrables associations.

Parmi elles, nous pouvons citer l'American Library Association (ALA), l'Association of College and Research Libraries (ACRL), ou encore l'Asssociation of Research Libraries (ARL). Ces trois associations, représentant 139 000 bibliothèques et 350 000 bibliothécaires, ont cosigné une lettre pour participer au Class Action, et s'opposer au monopole probable de Google, en demandant une révision du Google Book Settlement, en particulier sur le fait que Google puisse fixer les tarifs proposés aux bibliothèques publiques, ainsi sur le traitement de l'entreprise sur les œuvres orphelines. De nombreuses associations, américaines comme étrangères, ont également emboîté le pas de cette initiative visant à la révision du règlement.

Parmi ces opposants, se trouve l'Open Book Alliance, une association qui comprend des acteurs importants de l'édition et de la conservation du livre, ainsi que l'Internet Archive (organisation à but non lucratif consacré à l'archivage du Web). Les trois grands groupes que sont Microsoft, Amazon et Yahoo!, font également parti de l'OBA, et se plaignent, avant tout, au nom des lois AntiTrust. En effet, les accords entre Google et la guilde des auteurs et l'association des éditeurs américains représentent un frein important à la concurrence.

Le 14 novembre 2013, la Guilde des Auteurs a finalement été déboutée selon la décision rendue par le juge Denny Chin (New York), jugeant ainsi Google respectueux du Fair Use. La Guilde des Auteurs a ensuite fait appel, mais la US Supreme Court a rejeté ce dernier en Avril 2016.

Si l'ensemble des négociations et des procédures juridiques sont d'une énorme complexité, nous retenons que le modèle du copyright aux Etats-Unis diffère du droit d'auteur français, son but premier étant de « promouvoir le progrès des Sciences et des Arts utiles ».

La protection des auteurs n'est qu'un élément parmi d'autres d'une équation bien plus large[25]. Le juge Chin a donc pris largement en considération les bénéfices publics potentiels de Google Livres. Voici certains points mentionnés, qui ont contribué à la victoire de Google aux Etats-Unis[26]:

- « Google Books fournit une manière nouvelle et efficace pour les lecteurs et les chercheurs de trouver des livres » ;

- « Il rend des dizaines de millions de livres interrogeables par le biais de mots ou de citations » ;

- « Il fournit un index cherchable reliant chaque mot dans les livres aux ouvrages dans lesquels ils apparaissent » ;

- « Google Books est devenu un outil de recherche essentiel, qui aide les bibliothécaires à identifier et à trouver les sources pour la recherche » ;

- « Il rend le prêt entre bibliothèques plus efficace » ;

- « Il facilite la recherche et la localisation de citations » ;

- « Google Books étend l'accès aux livres » ;

- « En particulier, les populations mal desservies peuvent bénéficier d'un accès étendu à la connaissance et à davantage de livres » ;

- « Google Books a été intégré au système éducatif et son usage est enseigné au cours du cursus en culture de l'information que suivent les étudiants à tous les niveaux » ;

- « Il favorise grandement la fouille de données et de texte (data mining & text mining) » ;

[25] https://scinfolex.com/2013/11/15/verdict-dans-laffaire-google-books-une-grande-lecon-de-democratie/ [Consulté le 08/06/2016]

[26] *Ibid.*

- « Google permet à des personnes handicapées de chercher dans les livres et de les lire » ;

- « Google facilite la recherche pour des bibliothèques mal équipées, pour leurs acquisitions et le prêt entre bibliothèques » ;

- « Google Books aide à la conservation des livres et leur donne une nouvelle vie » ;

- « Google Books génère de nouvelles audiences et de nouveaux revenus pour les auteurs et les éditeurs ».

Cet accord, sur la base du fair use, est parfois perçu comme une chance pour favoriser l'émergence et le développement d'alternatives à but non lucratif de Google Livres[27] aux Etats-Unis comme, par exemples, les projets Internet Archive, Hathi Trust, ou encore la Digital Public Library of America.

On remarque que les opposants à cet accord n'ont pas été pas uniquement américains. En effet, nous pouvons mentionner L'UNEQ (Union des écrivaines et des écrivains québécois) au Québec, la compagnie Belge CopiePresse, ou encore la Ministre de la Justice Brigitte Zypries en Allemagne, qui a demandé le rejet de l'accord Google Book Settlement. La Commission européenne a également, de son côté, examiné le Google Book Settlement, et s'y est étrangement montrée favorable.

En France, nous pouvons évoquer le Syndicat National de l'Edition (SNE), le groupe La Martinière, la société des gens de lettres (SGDL), les éditeurs Gallimard, Albin Michel et Flammarion, mais aussi l'Agence France Presse, qui a, dès 2005, attaqué Google pour violations sur le droit d'auteur français. Les actions en justice ont été nombreuses et leurs débouchés variés. Parmi elles, et en réponse aux politiques de numérisation de Google, une « loi sur les livres indisponibles du XXème siècle », ainsi qu'un dispositif « base ReLIRE » gérés par la BnF, ont vu le jour.

Concernant les éditeurs, des accords, non-divulgués, ont finalement été signé avec

[27] https://scinfolex.com/2013/11/15/verdict-dans-laffaire-google-books-une-grande-lecon-de-democratie/ [Consulté le 08/06/2016]

chacun après plusieurs années d'attaque en justice.

Google a également perdu de nombreux procès auprès d'éditeurs, américains comme européens, qui sont parvenus à obtenir des dédommagements forfaitaires.

3.2 Le cas des œuvres orphelines

L'une des grandes controverses du projet Google Livres, concerne les œuvres orphelines et les ouvrages épuisés. En effet, cette numérisation massive a permis à Google de mettre la main sur de nombreuses œuvres orphelines (environ 2 millions d'ouvrages sur les 12 millions numérisés en Février 2010).

Les œuvres orphelines et épuisées sont toujours couvertes par le droit d'auteur. Concernant les ouvrages épuisés, l'auteur, l'éditeur et les ayants droits sont connus et identifiés, mais l'ouvrage n'est plus disponible dans le commerce (hormis d'occasion), et est seulement consultable chez les bibliothèques disposant d'un exemplaire. Les ayants droits restent maîtres de l'œuvre, et peuvent donc s'opposer à sa numérisation.

Le cas des œuvres orphelines est plus épineux. Une œuvre est considérée comme telle lorsqu'il est difficile, voire impossible, d'identifier et de retrouver clairement son auteur ou ses ayants droits. Les cas sont nombreux et variés. Il peut s'agir, par exemples, de la disparition d'un auteur (sans que l'on sache s'il est vivant ou non), de l'impossibilité de retrouver ses héritiers (tout comme il peut ne pas en avoir), ou encore parce que les droits sont détenus par une société disparue sans que l'on sache ce qu'est devenu son patrimoine (cessation de commerce, faillite). Une œuvre est donc qualifiée d'orpheline après l'échec d'une recherche d'un ou des ayants-droits, tout en étant clairement sous droit d'auteur (que personne ne peut donc faire valoir).

Grâce à l'option Opt-in, Google a donc mis gratuitement (les ouvrages ont été fournis gratuitement par les bibliothèques partenaires), la main sur plusieurs milliers d'œuvres orphelines en estimant que leur numérisation leur redonnerait une visibilité et une accessibilité. La question des œuvres orphelines, qui concernait alors environ 20% de Google Livres, a largement été débattue et révisée par le juge Denny Chin à New York. Dans la version de 2009 du Google Books Settlement, Google proposait, en plus de la création du registre évoqué précédemment (destiné à la recherche des ayants-droits d'œuvres orphelines), que 25% des sommes perçues des œuvres non déclarées soient dédiées à la recherche des ayants droits pendant 5 ans. Après 5 années de recherche sans résultat, ces sommes seraient alors versées à des associations luttant contre l'illettrisme dans les 4 pays concernés (Etats-Unis, Royaume Unis, Canada, Australie).

Le statut des œuvres orphelines est donc complexe, et leur statut juridique numérique encore plus. « L'affaire Google Livres » a largement contribué à la mise en place de débats et de questionnements sur leur nature juridique.

3.3 Quelques critiques célèbres et alternatives

Le projet Google Livres soulève de nombreuses polémiques et controverses. Nous l'avons vu précédemment, les méthodes de « cow-boys »[28] employées par la société concernant la numérisation d'ouvrages sous droit d'auteurs ont soulevé de nombreuses levées de boucliers. Pourtant, force est de constater que la grande majorité des actions en justice sont aujourd'hui terminées, et Google a su lier des partenariats avec nombre de maisons d'éditions et d'auteurs qui l'attaquaient autrefois.

Les critiques adressées à Google ne sont pas toutes de nature juridiques. En France, la plus célèbre réaction au projet fut celle de Jean-Noël Jeanneney, alors Président de la Bibliothèque Nationale de France, en 2005.

[28] JACQUESSON Alain, *Google Livres et le futur des bibliothèques numériques*, Editions du Cercle de la Librairie, Paris, 2010, p99.

Le 24 janvier 2005, le directeur de la BnF publie, dans le journal *Le Monde*, un article intitulé « Quand Google défie l'Europe ». L'article pointe du doigt les enjeux culturels et politiques derrière le projet de numérisation de Google. L'un des grands dangers constitue l'écrasante vision anglo-saxonne du savoir, reflétée dans les collections des bibliothèques retenues, que l'auteur met en opposition avec la notion de quota imposée en France sur les productions audiovisuelles. Jean-Noël Jeanneney, conscient aussi de l'écart de moyens financiers entre Google et Gallica par exemple, souhaite une réponse européenne au projet.

L'impact de l'article est énorme en France et à l'étranger, à tel point que le Président de la République Jacques Chirac exprime le vœu de rencontrer son auteur et de construire une stratégie alternative à Google. En mars 2005 est lancé le portail The European Library (TEL) permettant l'interrogation simultanée du catalogues informatisés de 19 bibliothèques partenaires et le 13 juillet 2005, le ministre français de la Culture met en place un "comité de pilotage en vue de la création d'une bibliothèque numérique européenne"

En Avril 2005, Jean-Noël Jeanneney publie un ouvrage « Quand Google défie l'Europe, Plaidoyer pour un sursaut », qui a été traduit dans une vingtaine de langues. Le livre appelle à une mutualisation des efforts, au travail en réseau, à une vaste collaboration pour numériser et mettre à disposition de tous, gratuitement, le fond des « autres bibliothèques »[29]. Il insiste sur la nécessité d'un « puissant effort financier sur fonds public »[30]. Ce sont les prémices d'Europeanna.

Aux Etats-Unis un autre article a rencontré un succès mondial. Rédigé par Robert Darnton, ancien directeur des bibliothèques de l'Université Harvard, et publié en Mars 2009, « La bibliothèque universelle de Voltaire à Google », est une mise en garde contre l'évolution du projet vers une visée plus commerciale. Sans être frontalement « anti-Google » comme l'ouvrage de Jean-Noël Jeanneney, l'article de Darnton insiste sur la nécessité de mettre en place une coopération équitable entre Google et les autres acteurs du livre numérique (éditeurs, auteurs, bibliothèques…). Publié à l'époque du Google Books Settlement, l'article souligne que Google Book Search est train de devenir la plus grande bibliothèque, et le plus

[29] JACQUESSON Alain, *Google Livres et le futur des bibliothèques numériques*, Editions du Cercle de la Librairie, Paris, 2010, p156.

[30] *Ibid.*

important magasin de livres de l'histoire. Pour Robert Darnton, c'est « un tournant majeur dans le développement de ce que nous appelons la société de l'information »[31], car « les intérêts privés pourraient bientôt l'emporter pour de bon sur l'intérêt public »[32]. Robert Darnton a été à la source d'un projet de bibliothèque numérique à l'accès gratuit dans le monde entier : la Digital Public Library of America (DPLA). Lancée en 2010, la bibliothèque a ouvert le 18 Avril 2013, et se veut une réponse à Google Livres, au service de tous et sans but lucratif.

Ces deux réactions englobent les principaux enjeux culturels, industriels, et politiques sous-jacents au projet Google Livres.

Si la numérisation d'une ampleur inédite de Google Livres, encore inimaginable il y a quelques années peut, de prime abord, renvoyer à l'idée grandiose d'une bibliothèque universelle, nous sommes cependant loin des idéaux de Paul Otlet et de ses désirs de partage de connaissances, de paix et d'égalité. Derrière les utopies culturelles que peut véhiculer un projet comme Google Livres, se trouvent de nombreuses tensions. La position dominante d'une entreprise privée sur la numérisation de livres, qui était encore jusqu'à peu l'affaire de politiques publiques, pose bien entendu la problématique du monopole industriel de Google, et de ses conséquences sur le secteur de la numérisation.

Ces critiques nous amènent à nous interroger sur l'existence de projets similaires, ou d'alternatives, à Google Livres.

Après avoir pris connaissance du projet de numérisation de Google, Microsoft va tenter de se placer en concurrent direct, et lance, en Décembre 2006, Microsoft Lives Book Search. Malgré un investissement fort de 5 millions de dollars, le projet est stoppé en 2008, et Microsoft a remis gratuitement 300 000 ouvrages numérisés à l'association Internet Archive.

Fondé en 1996, l'Internet Archive est un organisme américain à but non lucratif,

[31] DARNTON Robert, « La bibliothèque universelle, de Voltaire à Google », Le Monde Diplomatique, Mars 2009, p25.

[32] *Ibid.*

consacré à l'archivage du Web. Le projet consiste en une vaste bibliothèque numérique contenant du texte, des archives sonores, visuelles, etc. Parmi elles, on trouve des clichés instantanés (copies de pages prises à différents moments, de sites internet à diverses étapes de leur existence, et qui peuvent être recherchées grâce à un moteur de recherche spécifique : the wayback machine). Ce dernier propose de chercher l'histoire de plus de 487 billions de pages internet (« Search the history over 487 billion pages on the Internet »)[33].

L'Internet Archive est également membre d'un consortium international d'organisations, qui vise à la constitution d'archives publiques et gratuites contenant des livres et du matériel multimédia appartenant au domaine public ou numérisés, avec l'accord des ayants droits[34] : l'Open Content Alliance. Il regroupe des universités américaines et canadiennes, des industriels (Microsoft, Adobe, Yahoo !), mais également des bibliothèques (publiques, nationales).

L'Open Book Alliance (OBA), est un ensemble d'associations et de compagnies privées visant à s'opposer au monopole de Google Livres. On retrouve, parmi ses membres, Yahoo!, Amazon, Microsoft, ou encore l'Internet Archive.

Le HaithiTrust est une bibliothèque numérique mettant en commun des contenus de diverses bibliothèques numériques d'Europe et des Etats-Unis. Parmi elles, nous retrouvons Internet Archive, et même Google Livres. Il s'agit d'un projet sans but lucratif, le but étant de « collecter, d'organiser, de préserver, de communiquer et de partager les savoirs humains conservés en bibliothèques »[35].

HathiTrust a été fondé en Octobre 2008 par les 13 universités du CIC (Comimittee of Institutional Cooperation), ainsi que par l'Université de Californie. En Octobre 2015, HathiTrust contenait plus de 13,7 millions ouvrages (contre 25 millions pour Google Livres à la même période). La bibliothèque a également été attaquée en justice par la Guilde des Auteurs pour

[33] https://archive.org/index.php [Consulté le 08/06/2016]

[34] JACQUESSON Alain, *Google Livres et le futur des bibliothèques numériques*, Editions du Cercle de la Librairie, Paris, 2010, p176.

[35] JACQUESSON Alain, *Google Livres et le futur des bibliothèques numériques*, Editions du Cercle de la Librairie, Paris, 2010, p77.

violation de copyright en 2011, mais la justice américaine a tranché en faveur d'HaithiTrust, en invoquant le fair use (2014).

Ces deux cas d'énormes projets de nature non lucrative sont, avec, dans une moindre mesure, Europeanna que nous aborderons par la suite, les deux projets qui proposent un contenu comparable en termes d'échelle avec Google Livres, même si ce dernier reste largement dominant.

Devant cette multiplication de projets, de nature parfois assez hétéroclites, il nous semble bon de nous intéresser, désormais, à la définition que nous pouvons donner au terme « bibliothèque numérique ».

Partie II : Google Livres face aux bibliothèques numériques

1. Bibliothèques numériques, quelle(s) définition(s) ?

1.1 Les bibliothèques numériques : une notion encore floue

Avant de tenter de définir ce qu'est une bibliothèque numérique, il nous semble bon de préciser le terme de numérisation. Employé sans retenu, ce terme recouvre en fait des pratiques hétérogènes, appartenant parfois même à des métiers différents.

La numérisation concerne trois types de documents : les documents produits, créés sous forme numérique (digital natives), la reproduction de documents existants, en mode image ou texte (c'est le cas de la numérisation d'ouvrage papiers), et enfin la conversion de catalogues de bibliothèques (imprimés ou sur fiches) en catalogues informatisés en lignes (numérisation de catalogues). Ces trois types de numérisation comportent des différences en termes de production, de diffusion, d'éditorialisation, et de réalité économique. En effet, la reproduction de documents numériques concerne particulièrement des institutions patrimoniales comme la BnF, tandis que les textes produits sous formes numériques sont très utilisés dans les milieux universitaires. Les réalisations opérationnelles de dépôts organisés de contenus numériques, que sont HAL, Persée, ArchiveSIC, CAIRN, revues.org, Cerimes, Archives Audiovisuelles de la Recherche, Erudit etc…, en témoignent.

Une bibliothèque est le lieu au sein duquel est conservée et lue une collection organisée de livres. Elle propose, la plupart du temps, un accès à d'autres types de documents patrimoniaux comme des journaux, des périodiques, des enregistrements sonores, audiovisuels, des cartes, des plans, des partitions, etc… Tout aussi important que son fond, une bibliothèque répond à des objectifs de pérennisation du patrimoine, de diffusion, d'accessibilité et de partage.

L'expression bibliothèque numérique, équivalent français de « digital library », est une expression technique qui s'est rapidement banalisée et développée avec l'essor d'Internet. Elle fait la jointure entre deux mondes différents : celui des bibliothèques, et celui de l'informatique étendu aux Techniques de l'Information et de la Communication (TIC)[36]. Le terme a été utilisé pour la première fois en 1988, dans un rapport de la Corporation for National Research Initiatives[37].

Le terme « patrimoine numérique », quant à lui, a été employé pour la première fois en 2003 dans un texte de l'Unesco, la Charte sur la conservation du patrimoine numérique. Ce texte, proposait alors une « feuille de route pour les Etats et leurs institutions mémorielles, les bibliothèques nationales au premier rang, pour « élaborer des stratégies et des politiques [...] pour protéger le patrimoine numérique »[38].

En pleine expansion depuis les années 1990, les bibliothèques numériques sont complexes à définir. La pluralité des projets et les nombreuses évolutions technologiques qui

[36] PAPY Fabrice et LEBLOND Corinne, « Bibliothèque numérique : la nécessaire médiation », *Communication et Langage,* n°161, 2009, p33.

[37] http://www.cnri.reston.va.us/kahn-cerf-88.pdf [Consulté le 08/06/2016]

[38] BEQUET Gaëlle, *Trois bibliothèques européennes face à Google : aux origines de la bibliothèque numérique [1990 – 2010]*, Ecole des Chartes, Paris, 2014, p9.

les caractérisent en font un objet qui à la définition mouvante : « ces dispositifs [technologiques] ont, certes, apporté une contribution significative au concept de « bibliothèque numérique », mais en ont également brouillés les contours, en introduisant ce lien de subordination aux TIC et par transitivité immédiate à leur dynamique d'innovation perpétuelle »[39].

Quelques définitions, apportent toutefois un éclairage partiel sur leur nature.

La Digital Library Federation propose la définition suivante : « les bibliothèques numériques sont des organisations qui offrent des ressources, y compris en personnel, pour sélectionner, structurer, offrir un accès intellectuel, interpréter, distribuer, conserver des documents sous forme numérique. Une bibliothèque numérique garantit un accès sur la durée aux œuvres électroniques afin qu'elles soient aisément, et à un moindre coût, accessibles par un ou plusieurs publics spécifiques »[40].

En 2005 Carl Lagoze, professeur à la School of Information de l'Université du Michigan, soutient qu'une « une collection numérique devient une bibliothèque numérique lorsqu'elle répond aux quatre conditions suivantes :

- Elle peut être créée et produite dans un certain nombre d'endroits différents mais elle doit être accessible en tant qu'entité unique ;
- Elle doit être organisée et indexée pour un accès aussi facile que possible ;
- Elle doit être stockée et géré de manière à avoir une existence assez longue après sa création ;
- Elle droit trouver un équilibre entre le respect du droit d'auteur et la diffusion libre du savoir ; »[41].

[39] DUFRÊNE Bernadette et IHADJADENE Madjid et DRUCKMANN Denis (dir), *Numérisation du patrimoine, Quelles médiations ? Quels accès ? Quelles cultures ?,* Herman Editeurs, Paris, 2013, 314p.

[40] BEQUET Gaëlle, *Trois bibliothèques européennes face à Google : aux origines de la bibliothèque numérique [1990 – 2010]*, Ecole des Chartes, Paris, 2014, p16.

[41] *Ibid.*

Ces premiers éléments nous permettent d'avoir un premier aperçu du rôle des bibliothèques numériques. Nous allons maintenant nous intéresser à l'émergence de trois modèles considérés comme pionniers dans leur développement.

1.2 Trois modèles pionniers de bibliothèques numériques

L'histoire tend à montrer que les grandes bibliothèques numériques naissent au sein des bibliothèques nationales[42] : c'est le cas de la bibliothèque du Congrès, la British Library, ou encore Gallica (bibliothèque nationale de France). Ces bibliothèques peuvent être qualifiées de pionnières (milieu des années 90). De ces premières bibliothèques numériques, qui sont donc antérieures à Google Livres, se dégagent différents archétypes.

D'après Jean-Yves Mollier, trois modèles de bibliothèques numériques semblent, dans un premier temps, se dégager, pour ensuite se rapprocher[43].

Le premier modèle est celui du musée, et trouve son exemple le plus évocateur avec la British Library, institution héritière du British Musuem. La bibliothèque a fait ses premiers pas dans la numérisation en adoptant une approche que l'on pourrait qualifier de muséale, en sélectionnant méticuleusement les contenus à numériser, et en mettant en lumière les documents rares et les pièces uniques.

En plus de privilégier la qualité et la rareté à la quantité, la British Library va mettre en valeur ses documents, en développant une scénographie numérique très travaillée, comme le prouvent les différentes expositions virtuelles, ou encore la online Gallery.

Vient ensuite le modèle des archives, qui s'incarne particulièrement bien avec l'exemple de la Bibliothèque du Congrès (qui assure la fonction de bibliothèque de recherche

[42] DUFRÊNE Bernadette et IHADJADENE Madjid et DRUCKMANN Denis (dir), *Numérisation du patrimoine, Quelles médiations ? Quels accès ? Quelles cultures ?,* Herman Editeurs, Paris, 2013, p155.

[43] *Ibid.*

du Congrès des États-Unis et, de facto, constitue la bibliothèque nationale américaine).

La Bibliothèque du Congrès, notamment au travers de son projet le plus emblématique, le site American Memory, adopte une démarche plus archivistique. En effet, « le niveau de granularité retenu pour la numérisation est celui du fond plutôt que du document en lui-même »[44]. La Bibliothèque du Congrès met en avant ses collections à travers une politique d'éditorialisation forte des contenus, tandis que le traitement de l'information est plus faible. Les collections sont organisées par thèmes cohérents et exhaustifs, logique permettant ainsi aux publics de se retrouver, tout comme un fond d'archive pouvant servir de source d'études et de curiosités pour les chercheurs et les passionnés aguerris.

Enfin, le dernier archétype est celui de la bibliothèque comme le montre Gallica, la bibliothèque numérique de la Bibliothèque Nationale de France (BnF). Sa création fait suite au projet de créer une « bibliothèque d'un genre nouveau », selon le vœu de François Mitterrand en 1988. En effet, les années 80 sont marquées par un fort questionnement sur le lien entre technologies numériques et bibliothèques en France[45] : développement du Minitel et des banques de données à vocation nationale, du catalogue collectif des périodiques CCN-PS, de la banque nationale des thèses Théléthèses et Pancatalogue, du système Foudre visant à moderniser le prêt entre bibliothèques universitaires grâce aux TICs ou encore du projet Transdoc. C'est également à cette période que la Bibliothèque nationale de France a informatisé ses catalogues (base BN Opal est opérationnelle en 1987 par exemple). Le 14 juillet, François Mitterrand annonce la construction d'une « très grande bibliothèque d'un genre entièrement nouveau ». Ce sont les prémices de Gallica, bibliothèque qui « devra couvrir tous les champs de la connaissance, être à disposition de tous, utiliser les techniques les plus modernes de transmission de données, pouvoir être consultée à distance et entrer en relation avec d'autres bibliothèques européennes »[46].

[44] DUFRÊNE Bernadette et IHADJADENE Madjid et DRUCKMANN Denis (dir), *Numérisation du patrimoine, Quelles médiations ? Quels accès ? Quelles cultures ?,* Herman Editeurs, Paris, 2013, p155.

[45] BEQUET Gaëlle, *Trois bibliothèques européennes face à Google : aux origines de la bibliothèque numérique [1990 – 2010]*, Ecole des Chartes, Paris, 2014, p16.

[46] *Ibid*.

Gallica, la bibliothèque numérique de la BnF a été conçue pour être une bibliothèque de référence pour les chercheurs souhaitant exploiter les potentialités du numérique, proposant ainsi une large sélection d'œuvres dans plusieurs disciplines, le livre occupant une place majeure[47], le tout reconstitué selon une charte documentaire très détaillée, et faisant office de référence.

Ces trois types de bibliothèques illustrent différentes dynamiques puissantes qui animaient les bibliothèques numériques mais ne font pas office de modèles immuables, et n'englobent pas l'intégralité des projets de bibliothèques numériques.

Ainsi, nous constatons que les bibliothèques numériques répondent à des projets de valorisation du patrimoine sensiblement différents. Les trois modèles (musée, archive, bibliothèque) se retrouvent sur certains points cruciaux[48], qui se verront bousculer par l'arrivée de Google Livres :

- Un financement exclusivement public

- L'absence de volonté d'exhaustivité, qui parait alors impossible (contraintes financières et de temps)

- Le monopole du domaine public, les projets ne disposant que d'œuvres hors droit, ce qui créé d'ailleurs une période creuse à partir du XXème siècle.

- L'indépendance et l'isolement général des différents projets et des différentes bibliothèques, au niveau international mais aussi national.

Ces points communs illustrent la situation dans laquelle se trouvaient les bibliothèques

[47] DUFRÊNE Bernadette et IHADJADENE Madjid et DRUCKMANN Denis (dir), *Numérisation du patrimoine, Quelles médiations ? Quels accès ? Quelles cultures ?,* Herman Editeurs, Paris, 2013, p155.

[48] DUFRÊNE Bernadette et IHADJADENE Madjid et DRUCKMANN Denis (dir), *Numérisation du patrimoine, Quelles médiations ? Quels accès ? Quelles cultures ?,* Herman Editeurs, Paris, 2013, p159.

numériques avant Google Livres. L'arrivée de l'entreprise de Mountain View dans le monde de la numérisation va significativement bouleverser ces acquis. Le poids de Google Livres a suscité une forte émulation, qui a amené ce secteur documentaire à se transformer et à se développer de manière considérable. Sans prétendre à l'exhaustivité, nous tenons à citer quelques nouvelles impulsions qui témoignent de l'impact qu'a Google Livres sur les bibliothèques numériques.

2. L'impact de Google Livres sur les bibliothèques numériques

2.1 Les contours des trois modèles pionniers redessinés

En premier lieu, le postulat d'un financement exclusivement public n'est plus nécessairement vrai, Google ayant, depuis, signé des accords avec de nombreuses bibliothèques nationales comme celles d'Autriche, d'Italie et des Pays-Bas. En France, c'est la Bibliothèque Municipale de Lyon qui a été la première, le 11 Juillet 2008, à signer un accord avec Google. Ce dernier a d'ailleurs soulevé de nombreuses critiques.

Les dynamiques de numérisation vont également largement s'accroitre, et la question du nombre va être considérée autrement. Les prouesses de Google dans la numérisation (rapidité, quantité), interrogent les bibliothèques. Sans aller jusqu'à une volonté d'exhaustivité comme Google, il est désormais question de passer d'une numérisation très sélective, à une numérisation plus conséquente. Reste la question du financement, qui fera progressivement du concurrent Google, une sorte de partenaire avec qui les bibliothèques vont tenter de définir des règles communes de partenariats, tout en évitant les situations de monopole (comme, par exemple, le CENL, le Consortium Européen des Bibliothèques Nationales).

La question de la quantité et des financements va également amener les bibliothèques à chercher à se distinguer encore davantage de Google, en mettant l'accent sur des offres autres que « l'objet livre ». Les bibliothèques vont donc se mettre à valoriser des documents rares, des manuscrits, des cartes, divers documents audiovisuels…etc, les trois différents modèles précités se rapprochant ainsi les uns des autres.

Afin de combler l'absence de contenu issu du XXème siècle (les œuvres étant toujours sous protection du droit d'auteur), Gallica va également proposer un accès à des titres récents disponibles dans le commerce (environ 20 000 en 2012). Les utilisateurs peuvent parcourir des extraits de l'œuvre, ou l'acheter auprès d'un distributeur web (Gallica renvoie vers le distributeur).

Le projet Europeanna, d'abord pensé comme une forme d'alternative européenne à Google Livres[49], pour ensuite vouloir faire office de vitrine de la culture européenne numérisée, montre une volonté de coopération entre différentes institutions, et rompt ainsi avec la solitude des précédents projets de bibliothèque numérique. Impulsé par la France et lancé par la Commission Européenne en 2008, Europeanna propose différentes ressources numériques (livres, audiovisuel, photographies, documents d'archive…etc), issues de différentes bibliothèques des 27 états membres du projet. Europeanna n'archive pas les œuvres sur son site, mais fonctionne davantage comme un catalogue de recherche. De nombreux projets collaboratifs ont été créés, l'un des plus significatifs étant Europeanna 14–18, destiné à constituer une bibliothèque virtuelle sur la Première Guerre Mondiale.

Le modèle Europeanna, différent des trois précités, témoigne ici d'une volonté de création d'une culture européenne à partir d'un agrégat, tandis que le projet American Memories (Bibliothèque du Congrès) cherche à afficher un pan précis de l'identité nationale

[49] DUFRÊNE Bernadette et IHADJADENE Madjid et DRUCKMANN Denis (dir), *Numérisation du patrimoine, Quelles médiations ? Quels accès ? Quelles cultures ?,* Herman Editeurs, Paris, 2013, p163.

linguistique, (c'est le cas d'autres bibliothèques comme Blegica[50], ou encore Hispanica[51]). Si Europeanna reste le projet le plus emblématique en terme de développement de dynamiques d'alliances entre différentes institutions, nous pouvons également mentionner d'autres initiatives comme le Réseaux Francophone Numérique (16 pays donnant un accès à de la presse francophone numérisée par pays, appartenant à l'organisation internationale de la francophonie).

2.2 L'arrivée de Google Livres : quel(s) impact(s) pour les bibliothèques numériques ?

L'arrivée de Google dans le monde de la numérisation a profondément impacté les bibliothèques et les politiques de numérisation. Les trois modèles pionniers, qui donnaient alors un aperçu des différentes dimensions des bibliothèques numériques, se voient modifiés par l'arrivée du géant industriel dans ce nouveau secteur.

Nous l'avons vu précédemment, Google multiplie les partenariats avec les bibliothèques, et ce, malgré de nombreuses controverses. Nous allons maintenant essayer de déterminer en quoi conclure des partenariats avec une entreprise privée peut-être bénéfique pour une bibliothèque.

La numérisation du patrimoine recouvre différents enjeux pour les bibliothèques : préservation du patrimoine (en particulier les documents fragiles), mise en valeur de fonds peu consultés, ou encore la volonté d'étendre leur portée et d'atteindre un public plus large.

[50] http://belgica.kbr.be/fr/accueil_fr.html [Consulté le 08/06/2016]

[51] http://www.bne.es/fr/Catalogos/BibliotecaDigitalHispanica/Inicio/ [Consulté le 08/06/2016]

Le potentiel de la numérisation des fonds patrimoniaux est donc énorme.

Cependant, plusieurs problématiques se posent[52] :

- Les questions juridiques sur les documents, comme le droit de reproduction, la question du « trou noir du XXème siècle », la viabilité et l'exploitation numérique des œuvres orphelines.
- L'évidente question des coûts financiers. La numérisation implique de nombreuses dépenses techniques, matérielles et humaines.
- La question de la temporalité de réalisation de la numérisation et de sa mise en ligne.
- La question technique de l'indexation et de la qualité des métadonnées.
- La conservation à long terme de ces données.
- Le coût et la maintenance des équipements informatiques nécessaires à leur stockage, diffusion, maintenance.
- La qualification des équipes, du personnel.

Pour les bibliothèques publiques, la numérisation rétrospective de leurs fonds pour permettre la consultation à distance, tout comme l'archivage d'une masse de données toujours plus grande, représentent un véritable défi. Il est aussi clair que ces missions n'ont pas entrainé de dotations budgétaires nouvelles[53]. Les entreprisse privées, et particulièrement Google, avec leurs fonds financiers, se posent comme un intermédiaire de choix.

Ces questions ont amené les bibliothèques à envisager des partenariats avec des entreprises privées, permettant ainsi une réduction des coûts de production, des délais

[52] DUFRÊNE Bernadette et IHADJADENE Madjid et DRUCKMANN Denis (dir), *Numérisation du patrimoine, Quelles médiations ? Quels accès ? Quelles cultures ?,* Herman Editeurs, Paris, 2013, p56.

[53] DUFRÊNE Bernadette et IHADJADENE Madjid et DRUCKMANN Denis (dir), *Numérisation du patrimoine, Quelles médiations ? Quels accès ? Quelles cultures ?,* Herman Editeurs, Paris, 2013, p56.

diffusion, mais également de disposer de garanties sur le professionnalisme des techniques utilisées, et la qualité des résultats[54]. En Europe, nous pouvons citer le partenariat avec IBM concernant la numérisation des Archives Séville. En France, nous pouvons évoquer l'implication de BULL et Kodak dans la numérisation des manuscrits enluminés de la Bibliothèque Nationale de France.

Si de tels accords contribuent à améliorer largement la diffusion et la visibilité du patrimoine numérisé, et s'ils semblent, de prime abord, providentiels, il est nécessaire, pour les bibliothèques, de s'interroger et de réfléchir sur ces partenariats publics/privés, et de définir en amont une stratégie claire, pour ainsi s'assurer de pouvoir en maitriser les conséquences.

2.3 Google Livres est-il une bibliothèque numérique ?

Si Google Livres joue un rôle décisif dans l'évolution des bibliothèques numériques, nous ne pensons pas pouvoir qualifier cet outil de bibliothèque numérique pour autant. En effet, Google ne revendique aucune politique documentaire, et peut être comparé à un « gigantesque entrepôt de données qui servent de support à une politique de service »[55].

A l'inverse, une bibliothèque comprend des enjeux d'accessibilité, de diffusion, ainsi que des enjeux de politiques culturelles, et de de socialisation de l'information.

Cela se caractérise notamment, au-delà de la mise à disposition de leurs ressources numériques, par un engagement dans l'éditorialisation de contenus sur le web. Les institutions patrimoniales doivent être sources de confiance, d'autorité, et de crédibilité. Leur

[54] *Ibid.*

[55] DUFRÊNE Bernadette et IHADJADENE Madjid et DRUCKMANN Denis (dir), *Numérisation du patrimoine, Quelles médiations ? Quels accès ? Quelles cultures ?,* Herman Editeurs, Paris, 2013, p56.P 160

rôle est déterminant : « investie d'une mission de service public, l'institution est donc une référence, et s'impose comme un contre-pouvoir face à des entreprises privées de numérisation des documents et de privatisation des usages »[56]. Cette vocation de sources de références amène les bibliothèques qui investissent le numérique, à s'interroger quant à l'éditorialisation de leurs contenus.

Le terme « éditorialisation » recouvre les pratiques de publication et d'accessibilité des contenus sur le web. Selon Bachimont, il s'agit de la conclusion logique du processus de numérisation des contenus. De manière plus concrète, l'éditorialisation englobe un ensemble d'actions ayant pour but d'apporter du sens au document. Éditer un contenu ne signifie pas seulement le choisir, mais lui « donner son sens propre en l'insérant dans un contexte technique précis, en le reliant à d'autres contenus, en le rendant visible grâce à son indexation, à son référencement, etc. »[57].

Pour donner une définition claire de l'éditorialisation et de ses enjeux, nous reprendrons les trois déclinaisons que propose Arnaud Laborderie[58].

- Sur le plan technique, elle correspond d'abord à un enjeu de structuration des contenus, à la fois par hiérarchisation, rubricage et maillage. L'indexation et les métadonnées assurent l'accessibilité et l'interopérable.

[56] JEANNENEY Jean-Noël, *Quand Google défie l'Europe : Plaidoyer pour un sursaut*, Mille et une nuits, Paris, 2007, p46.

[57] EBERLE-SINATRA Michaël et VITALI-ROSATI Marcello (dir), *Pratiques de l'édition numérique*, Parcours numérique, Montréal, 2014, p126.

[58] LABORDERIE Arnaud, « Éditorialisation des bibliothèques numériques : le cas des Essentiels de Gallica », *Colloque International sur le Document Electronique : « Documents et dispositifs informationnels à l'ère post-numérique »*, 2015, p4.

- En tant que pratique individuelle, l'éditorialisation peut s'inscrire dans une démarche de veille informationnelle. Il s'agit, pour l'internaute, de sélectionner des ressources, et de se les approprier. On peut distinguer trois étapes dans ce processus :

    - Rééditer du contenu en l'adaptant aux contraintes éditoriales du web.
    - Contextualiser ce contenu pour qu'il fasse sens.
    - L'enrichir, lui apporter une valeur ajoutée, un point de vue, un angle singulier.

Éditer, contextualiser, enrichir : ces actions peuvent également concerner le lecteur qui peut parfois, lui aussi, créer du contenu, l'organiser, le personnaliser (c'est le cas des étagères de lecture de Google Livres, où l'utilisateur peut ainsi classer et ranger sa collection de livres comme il le souhaite).

- En tant que pratique documentaire, l'éditorialisation est au service du sens : elle offre un appareillage herméneutique en permettant l'interconnexion des documents, en distinguant les sources primaires ou secondaires, en contextualisant. Attachée au document ou à l'ensemble documentaire, elle rend possible des effets de zoom informationnel. C'est une nécessité face à la diversité de ressources hétérogènes classées par type, ou cloisonnées par disciplines. L'éditorialisation offre alors la possibilité de réordonner les documents, et de construire du sens.

La notion d'éditorialisation nous semble être le terme clé pour représenter la mission des bibliothèques numériques. L'accessibilité, les relations de sens entre les documents, la notion de parcours, d'organisation du savoir et de médiation sont ce qui donne corps, autorité, et légitimité aux bibliothèques.

Le rôle premier des bibliothèques numériques ne diffère pas fondamentalement de celui des bibliothèques traditionnelles : elles investissent des missions de sélection et d'acquisition, d'interprétation et de description, de mise à disposition et de préservation[59]. Gallica nous

[59] JACQUESSON Alain, *Google Livres et le futur des bibliothèques numériques*, Editions du Cercle de la Librairie, Paris, 2010, 165p.

semble être, à ce titre, un exemple particulièrement représentatif de la mission des bibliothèques numériques.

Une bibliothèque n'est pas simplement une réserve de documents, elle a aussi des responsabilités. En plus de mettre à disposition des documents, elle se doit de les organiser pour qu'ils fassent sens. Gallica, à travers ses collections « Les Essentiels de la littérature » (des corpus multimédias sur des grands thèmes et auteurs), favorise la découverte et l'appropriation chez les publics scolaires et enseignants, à travers des « chemins d'accès, des parcours de connaissances, qui modélisent les usages et organisent les ressources »[60]. Les contenus sont choisis et organisés méticuleusement, afin de proposer des repères sur des grands mouvements littéraires et artistiques. Ces contenus, de natures diverses (imprimés, manuscrits, images, enregistrements sonores et vidéos, corpus…), sont donc mis en relation et organisés via des parcours guidés, afin d'être le plus accessibles et le plus logiques possible.

Une recherche avec le terme « Diderot » sur Gallica (option « dans tout Gallica ») affiche 13 476 résultats étendus, sur 899 pages. Nous comprenons là très vite la nécessité de hiérarchiser, et d'organiser les contenus.

La masse de documents numérisés, toujours plus importante, rend les informations plus complexes à trouver. Face à l'abondance de contenus, les modalités d'accès aux ressources sont déterminantes. Ainsi, les « métadonnées affectées à chaque document, les architectures catégorielles et ontologiques, les index, les listes »[61] et les collections, sont de multiples outils pour circuler au sein de la masse d'informations que contient une bibliothèque, et propose des alternatives à la simple fonction « recherche », qui offre quasiment systématiquement une liste de résultats impressionnante, qui peut décourager le grand public (qui n'est pas doté des mêmes compétences que l'expert, qui sait manier les filtres, les mots clés, l'option « recherche avancée etc).

[60] LABORDERIE Arnaud, « Éditorialisation des bibliothèques numériques : le cas des Essentiels de Gallica », *Colloque International sur le Document Electronique : « Documents et dispositifs informationnels à l'ère post-numérique »*, 2015, p12.

[61] LABORDERIE Arnaud, « Éditorialisation des bibliothèques numériques : le cas des Essentiels de Gallica », *Colloque International sur le Document Electronique : « Documents et dispositifs informationnels à l'ère post-numérique »*, 2015, p16.

Fabrice Papy, lorsqu'il tente de définir ce qu'est une bibliothèque numérique, précise bien le rôle centrale de la médiation : « concept clé de l'identité des bibliothécaires, dans toute l'étendue de sa polysémie, la médiation, souvent soule épargnée par les discours empreints de fatalisme ou l'expression de doutes sur l'avenir, se situe au cœur même de l'évolution de ce métier au sein d'un environnement documentaire hybride »[62].

D'après Michael Keller, le directeur des bibliothèques de l'Université de Standford, par le biais d'un article intitulé « What's a library anymore, anyway ? »[63], aucune collection n'est parfaite ou complète. Dans ce cadre, il constate qu'il est donc nécessaire, pour une bibliothèque, de pouvoir accéder à d'autres fonds possédés par diverses institutions, via le prêt entre bibliothèques, par exemple.

De plus, il est bon de préciser qu'une bibliothèque numérique est aujourd'hui faite d'éléments variés. Gallica, en plus de contenir de nombreux livres numérisés, contient des documents patrimoniaux divers, comme des numéros de presse, des revues, des images, des cartes, des vidéos, des partitions, des manuscrits, des livres, des objets, ou encore des enregistrements sonores. Ce n'est pas le cas de Google Livres, qui se concentre uniquement sur le support livre, et à moindre mesure, le magazine. Les bibliothèques numériques s'intéressent aussi aux éléments propres au Web : les blogs, bouquets de liens, données brutes.

Malgré sa quantité d'ouvrage sans commune mesure (Google Livres aime rappeler qu'il est « l'index de livres complets le plus fourni au monde »), l'absence totale de politique de médiation fait qu'il est difficile de considérer Google Livres comme une bibliothèque numérique. Nous l'avons vu, son fond documentaire s'apparente à un large entrepôt de données, sans forme d'éditorialisation. Au mieux, Google Livres ne revêt donc « qu'une des composantes d'une bibliothèque numérique telle que les professionnels la fabriquent ou du

[62] PAPY Fabrice et LEBLOND Corinne, « Bibliothèque numérique : la nécessaire médiation », *Communication et Langage,* n°161, 2009, p51.

[63] http://pear.accc.uic.edu/ojs/index.php/fm/article/view/1053/973 [Consulté le 08/06/2016]

moins en rêvent »[64]. Pourtant, son impact dans l'univers de la numérisation est énorme, et a amené les bibliothèques à se « réinventer ». Il existe donc un lien fort entre l'entreprise de Mountain View, et les bibliothèques numériques.

Nous pouvons donc nous demander quel est l'intérêt de Google pour la numérisation.

3. Pourquoi Google a t il investit le secteur de la numérisation ?

3.1 La numérisation : quel(s) intérêts(s) pour Google ?

Derrière le projet Google Livres, se cachent de nombreux enjeux, autres que la philanthropie dont se réclame l'entreprise. En 2004, lors du lancement de Google Print, le géant de Mountain View annonçait : « L'objectif du projet est simple : aider à conserver la prééminence du livre et des bibliothèques au sein d'une culture toujours plus centrée sur l'internet en plaçant leur information au sein du monde en ligne. Nous espérons que cela conduise plus de monde vers leur bibliothèque locale, que cela conduise à la numérisation des fonds de quelques-unes des plus grandes institutions de recherche, que cela rende tout ceux-ci accessible »[65].

Aujourd'hui, Google Livres contient une réserve impressionnante de livres libres de droit, comme sous droits d'auteurs, et les diffuse à une échelle mondiale, celle des grandes industries culturelles. Il est évident que la gratuité apparente du service a également, et

[64] JACQUESSON Alain, *Google Livres et le futur des bibliothèques numériques*, Editions du Cercle de la Librairie, Paris, 2010, 116.

[65] JACQUESSON Alain, *Google Livres et le futur des bibliothèques numériques*, Editions du Cercle de la Librairie, Paris, 2010, p53.

surtout, une finalité marchande.

En effet, Google est une entreprise privée, cotée en bourse, employant environ 50 000 salariés, et dont l'objectif concret est, avant tout, de faire du profit.

Nous avons évoqué précédemment le fonctionnement financier de Google, nous allons maintenant essayer de comprendre en quoi Google Livres est source de valeur pour l'entreprise.

Premièrement, la librairie numérique Google Play Livres est la source de rémunération la plus visible et identifiable pour l'entreprise. Concurrente directe de l'iBookstore d'Apple et du Kindle d'Amazon, Google Play Livres peut, en rendant disponible un grand nombre de livres gratuitement, « guider les clients vers l'offre commerciale, et les fidéliser à la plateforme »[66]. Le fait d'être englobée dans une boutique plus large (Google Play) peut également avoir le même impact. En effet, un utilisateur satisfait et « fidélisé» de Google Play Music ou de Google Play Films & TV, sera probablement plus enclin à se diriger vers Google Play Livres, plutôt que vers les services concurrents.

Google Livres représente aussi, bien évidemment, un moyen d'accéder à de nombreuses informations sur les utilisateurs du service, utiles pour améliorer les systèmes de publicités contextuelles qui se trouvent sur les différents services Google, et sur tous les sites partenaires.

En nous intéressant au modèle financier de Google, nous avons vu que la publicité était sa principale source de rémunération. Chaque page web constitue donc un ensemble d'« espaces » potentiels, au sein desquels de la publicité peut être affichée. Aujourd'hui, Google Livres n'affiche étonnement aucune publicité sur son site. Cependant, on en trouve

[66] DUFRÊNE Bernadette et IHADJADENE Madjid et DRUCKMANN Denis (dir), *Numérisation du patrimoine, Quelles médiations ? Quels accès ? Quelles cultures ?,* Herman Editeurs, Paris, 2013, p70.

lorsque l'on consulte la page de présentation d'un livre sur la boutique Google Play Livres, qui contient un bandeau sur la droite présentant des « articles similaires », et « plus d'ouvrages du même auteur ». Le potentiel reste important.

Tout comme Google Play Livres est une composante de Google Play, nous constatons que Google Livres s'inscrit dans une large panoplie de services, souvent gratuits, que propose Google (Gmail, Google Maps, Google Scholar, Google Play, Google Traduction etc…). Google se présente en « guichet unique »[67], et cherche à ce que l'utilisateur navigue le plus longtemps possible sur son réseau (on pourrait parler « d'univers Google »). En plus d'engager les internautes, cela permet également à Google de réaliser des économies d'échelles. En effet, les différents services de l'entreprise utilisent les mêmes centres de données, et se partagent différents éléments techniques (algorithmes, base de données etc…). L'investiture de Google dans le monde livre constitue un agrandissement, et une continuité logique dans son développement. Nous rappelons que l'entreprise se donne pour mission « d'organiser la connaissance du monde », le monde du livre s'ancrant tout à fait dans cette logique.

Le nombre immense d'ouvrages que contient Google Livres représente aussi un fond de connaissances et de données propice aux opérations de calcul et aux traitements algorithmiques, utiles à l'amélioration d'autres services de l'entreprise. Ainsi, l'énorme corpus multilingue qu'est Google Livres constitue surement une aubaine pour améliorer l'outil Google Traduction de l'entreprise, d'autant plus lorsque l'on sait que la traduction automatique par voie statistique sur une très grande masse de données est plus performante qu'une approche par modélisation[68]. Dans la même optique, nous pouvons citer le projet de recherche Culturonomics de Google Labs qui, en utilisant le corpus de Google Livres, essaye de « mesurer » la culture, en analysant des évolutions culturelles, sociales, économiques ou encore politiques, à partir de fréquences de mots, ou de suites de mots. Google reste très discret sur les usages internes de la base de données qu'est Google Livres, mais nous pouvons

[67] *Ibid*.

[68] DUFRÊNE Bernadette et IHADJADENE Madjid et DRUCKMANN Denis (dir), *Numérisation du patrimoine, Quelles médiations ? Quels accès ? Quelles cultures ?,* Herman Editeurs, Paris, 2013, p71.

déjà émettre l'hypothèse que ces analyses computationnelles seront, un jour, (et si ce n'est pas déjà le cas), exploitées dans des buts commerciaux. Google Livres représente donc une aubaine pour des traitements de données simples et applicables à un nombre très conséquent de données (calcul de probabilités, comptage de mots etc…).

Les pratiques des utilisateurs et leur interaction avec les livres et l'interface Google Livres sont également sources de données. La création « d'étagères » grâce auxquelles l'utilisateur organise ses collections (catalogage), peut être un indicateur de proximité entre les livres (thématique). Ainsi, les « différents gestes de recherche produisent des listes et des profils qui représentent des trajectoires de signification, c'est-à-dire des relations (statistiques) entre des mots et des livres qui sont dérivés de pratiques réelles et non seulement des contenus et des traitements de texte »[69]. Le projet Google Flu Trends[70], a montré le potentiel analytique, et même prédictif, du traitement automatique des données d'un très grand nombre de requêtes.

Google détient aujourd'hui le fond de livres numérisés le plus important du monde. Dans de nombreux cas, il dispose également d'exclusivités concernant leur exploitation commerciale ou leurs métadonnées. En plus de cela, il est le moteur de recherche le plus utilisé à l'échelle planétaire. L'entreprise dispose donc d'une position de choix concernant la recherche et le tri des contenus, et peut ainsi proposer des réponses très adaptées aux requêtes bibliographiques. Enfin, via sa plate-forme de vente (Google Play Livres) et ses partenariats, Google est dotée d'une place confortable pour inciter l'internaute à acquérir des documents.

Les clauses entre Google et les bibliothèques, dans lesquelles Google a, dès l'origine du projet, demandé des exclusivités concernant l'indexation et la commercialisation des copies numérisées (dans le but principal d'empêcher l'indexation sur les moteurs de recherche

[69] DUFRÊNE Bernadette et IHADJADENE Madjid et DRUCKMANN Denis (dir), *Numérisation du patrimoine, Quelles médiations ? Quels accès ? Quelles cultures ?,* Herman Editeurs, Paris, 2013, p72.

[70] https://www.google.org/flutrends/about/ [Consulté le 08/06/2016]

concurrents), laissent à penser que l'entreprise possédait déjà, même aux prémices de Google Livres, des objectifs commerciaux bien définis.

3.2 Google Livres, une composante supplémentaire de « l'empire Google »

Enfin, en plus de la publicité, des données recueillies, et de l'inter-connectivité des services Google, de manière plus large et pour revenir à l'origine du projet, les démarches stratégiques de Google sont souvent rattachées au « soft power », un concept inventé au début des années 1990 par Joseph Nye, pour qualifier le rayonnement culturel et diplomatique des Etats-Unis, par opposition au « hard power » militaire de l'après-guerre.

Les premiers contacts de Google avec de très grandes bibliothèques américaines emblématiques, dont la plus grande bibliothèque universitaire du monde (Harvard), s'inscrivent dans cette offensive de « charme » et de « puissance douce », qui contribue à soigner l'image de l'entreprise (même si, bien entendu, de nombreuses critiques ne tarderont pas à être émises dans le monde entier avec la numérisation de livres sous droits d'auteur). Google Livres est donc également un « argument » de la part de l'entreprise, pour développer des partenariats avec des acteurs de la culture.

Là encore, nous constatons que Google Livres s'imbrique dans les logiques stratégiques de l'ensemble de la société, Google multipliant les financements de start-up (le 20 avril 2016, l'ONG Emmaus a reçu un don d'un million de dollars de la part de Google pour financer sa start-up WeTechCare), les dons, les créations d'évènements de prestiges (relatifs aux les fameux Zeitgeist), les rapprochements avec des politiques (Barrack Obama est souvent considéré comme le premier lobbyiste de Google[71]), dans le but d'étendre son image de « bienfaiteur » et ses activités dans de nombreux domaines et pays.

Avec le projet Google Livres, l'entreprise met à disposition du monde entier des

[71] NORA Dominique, « Le vrai pouvoir de Google en France», Le Nouvel Observateur, Juin 2016, p35.

millions d'ouvrages consultables et téléchargeables gratuitement, et facilite donc l'accès à la culture. Cette vision de Google Livres se situe dans la parfaite continuité de la mission principale dont se targue l'entreprise : « organiser la connaissance du monde ».

Partie III : Analyse des sites Google Livres et Gallica

1. Pourquoi comparer Google Livres et Gallica ?

Dans cette dernière partie, nous allons comparer le fonctionnement des sites Google Livres et Gallica. Nous pensons aujourd'hui que Google Livres est organisé de manière à amener le lecteur sur la plateforme Google Play Livres.

Gallica, développe une éditorialisation très poussée, représentative des enjeux patrimoniaux et communicationnels des bibliothèques numériques. Nous n'analyserons pas l'intégralité des deux sites, celui de Gallica étant très vaste, mais nous nous focaliserons sur les fonctionnalités les plus significatives (page d'accueil, fonction recherche, visualiseur de documents, espace personnel etc...)

Comme nous l'avons vu précédemment, l'arrivée de Google dans le monde de la numérisation a profondément affecté ce secteur, en le dynamisant et en amenant les bibliothèques à se dynamiser et se transformer.

Gallica, la bibliothèque numérique de la Bibliothèque Nationale de France et de ses partenaires, a été mise en ligne en 1997. Sa création fait suite au projet de créer une « bibliothèque d'un genre nouveau », selon le vœu de François-Mitterrand, en 1988. En effet, est envisagé, avec l'essor des technologies et du multimédia, un projet de bibliothèque virtuelle accessible via des postes de lecture dédiés à cet usage (les PLAO, postes de lecture assistés par ordinateur), afin de proposer aux lecteurs un ensemble de contenus et documents en mode image, en libre consultation, qu'il est entre autres possible d'annoter. Les premières numérisations d'ouvrages libres de droit commencent en 1992. La bibliothèque est, à ce titre, pionnière et novatrice. La numérisation en mode « texte » est prise en compte à partir de 1993, grâce notamment à un partenariat avec le CNRS.

Lorsqu'elle s'ouvre au public en 1997, la bibliothèque Gallica propose un accès aux œuvres et à des dossiers documentaires (textes et images principalement) francophones du XIXème siècle. Petit à petit, s'intègrent la prise en compte d'autres formats documents en mode texte, mais également la création de parcours structurés, ainsi que des dossiers thématiques. Nous constatons l'élaboration d'une véritable politique d'éditorialisation et d'agencement des contenus. Une deuxième version du site est mise en ligne en 2000. De nouveaux documents sont numérisés (grand projet de numérisation de la presse quotidienne française par exemple). Le mode image reste largement dominant.

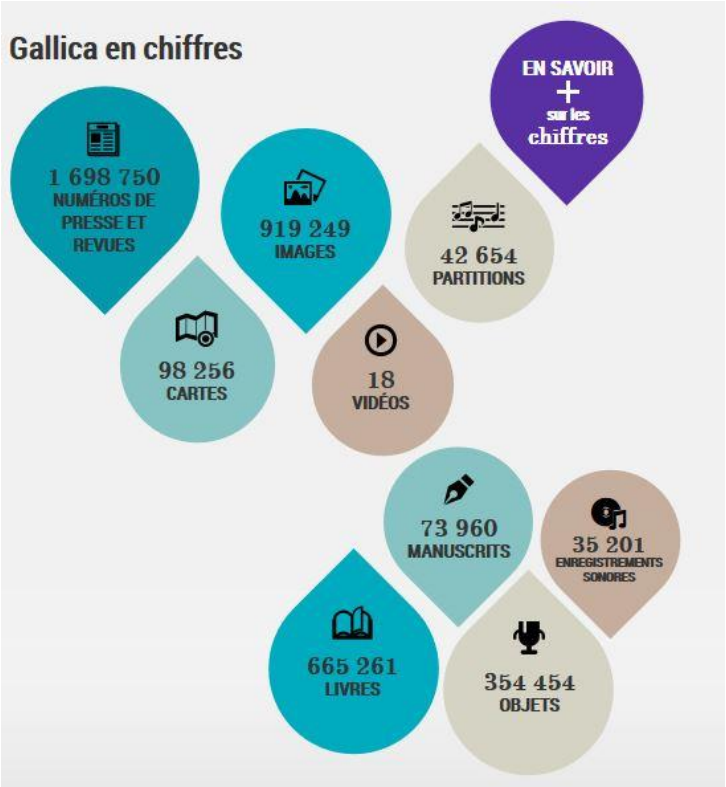
Avec l'arrivée de Google dans le monde de la numérisation et l'appel de Jean-Noël Jeanneney (à cette époque directeur de la BnF), Gallica s'engage dans la numérisation de masse, et se met à utiliser la technique de l'OCR (Optical Character Recognition). Ce procédé permet de situer et reconnaître les chaînes de caractères dans une image, et donc d'opérer la conversion des mots qui peuvent ensuite être utilisés pour différents usages, dont la recherche plein texte qu'offre la bibliothèque numérique Gallica. Il s'agit d'un traitement automatique fait par logiciel. L'OCR est également utilisé par Google Livres.

Gallica 2, une nouvelle version de Gallica, est lancée en Novembre 2007. Le site s'appuie sur l'algorithme du moteur de recherche Lucene pour organiser ses résultats.

En réponse à Google, des œuvres sous droits sont intégrées au site, grâce à des partenariats entre la Bibliothèque Nationale de France, la Direction du livre et de la lecture, le Centre national du livre, et le Syndicat national de l'édition. Le moteur de recherche donne accès à la fois aux documents libres et aux documents sous droits, les diffuseurs restant libres de décider des conditions de consultation du texte (accès intégral, extraits « snippets » similaires à ceux de Google Livres …). Gallica permet également l'accès aux ressources de bibliothèques virtuelles partenaires via le protocole OAI-PMH.

Toujours en réponse à Google, Gallica signe, le 15 Janvier 2013, des accords de numérisation avec des sociétés privées : la société américaine ProQuest, pour la numérisation de livres anciens, ainsi que les sociétés Believe Digital (France) et Memnon ARchiving Services (société Belge rachetée par Sonny Europe LTD en 2015), pour numériser des disques vinyles.

Gallica propose aujourd'hui un accès à de nombreux documents de nature très variée.



Constitution du fond documentaire de Gallica tel qu'il est présenté sur la page d'accueil du site. La capture d'écran date du 05 juin 2016.

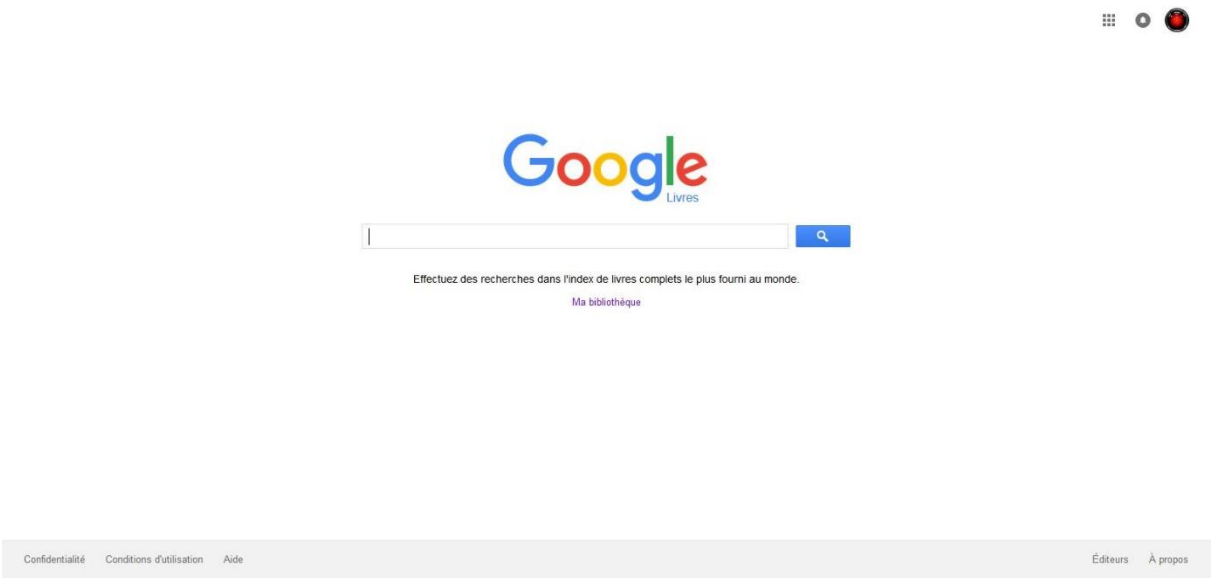
La bibliothèque a également largement investie les réseaux sociaux (Facebook, Twitter, Printerest, blogs).

Ce bref historique, qui se focalise uniquement sur certains points cruciaux, nous permet de justifier le choix de la bibliothèque numérique Gallica pour l'analyse de contenu. Pionnière dans le domaine de la numérisation, cette dernière a opéré de nombreux changements avec l'arrivée de Google, en accélérant massivement son rythme de numérisation, en améliorant ses composantes techniques, en signant des partenariats publics/privés, en intégrant un espace personnel pour les utilisateurs, en structurant son minutieusement son fond documentaire et ses chemins d'accès, mais aussi en proposant des services marchands.

2. Analyse des sites

2.1 La page d'accueil

La page d'accueil de Google Livres est particulièrement saisissante. Elle s'ancre parfaitement dans la continuité et l'identité visuelle du moteur de recherche de Google. Le fond de la page est blanc, on ne peut plus simple et épuré, et on retrouve le fameux logo de Google, enrichi de la mention « livres ». Juste en dessous, une simple barre de recherche sous laquelle est précisé : « Effectuez des recherches dans l'index de livres complets le plus fourni au monde. ». Toujours sous la barre de recherche, on trouve un lien hypertexte intitulé « Ma bibliothèque ». Si l'on excepte les fonctionnalités propres à Google qui apparaissent en haut à droite de la page (applications Google, notifications Google et Compte Google+), et un bandeau discret en bas de page qui comprend les boutons « Confidentialité », « Conditions d'utilisations », « Aide », « Editeurs » et « A propos », la page d'accueil ne contient que deux fonctionnalités : la barre de recherche, et le bouton « Ma bibliothèque ». Dès la page d'accueil, Google montre ainsi sa confiance dans son moteur de recherche et ses algorithmes. Google, contrairement à Gallica, ne donne aucune information quant au nombre d'ouvrages numérisés accessibles sur le site.



La page d'accueil de Google Livres.

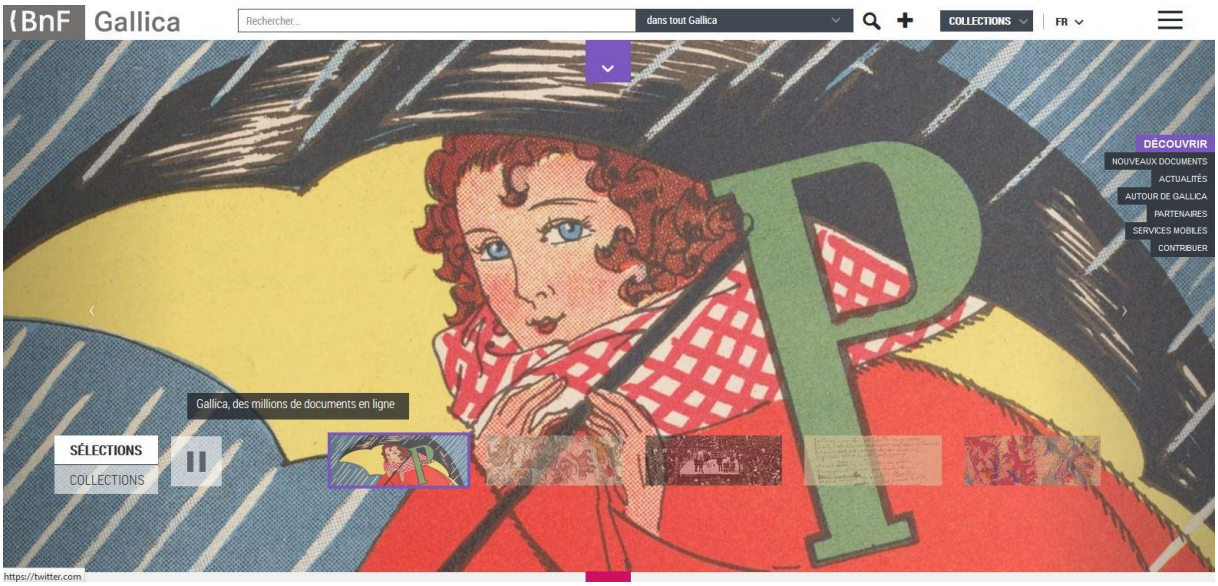
La page d'accueil de Gallica est, quant à elle, radicalement différente. En haut à gauche, nous trouvons un bouton qui permet de se rendre sur le site de la BnF, suivi d'un bouton qui permet d'accéder à la page d'accueil de Gallica. En haut au centre, on aperçoit une grande barre de recherche et un filtre permettant d'effectuer la recherche dans « tout Gallica », « les livres », « les manuscrits », « les cartes », « les images », « la presse et les revues », « les enregistrements sonores », « les partitions » et « les objets ». Juste à côté, un bouton permettant d'accéder à la fonction « recherche avancée », un bouton « collection », qui ouvre un filtre permettant de naviguer au sein des collections « par types de documents », «par thématique » ou « par aires géographiques », mais aussi une rubrique « langue » (qui n'est pas encore opérationnelle), et un onglet « accéder au menu ». Lorsque l'on clique sur ce dernier, un bandeau se déroule au-dessus de celui-ci, et permet d'accéder aux rubriques « en savoir + », « liens utiles », « services mobile ». Il contient également des liens vers les comptes Facebook, Twitter, et Pinterest, de Gallica, ainsi qu'un formulaire d'abonnement à la newsletter, un bouton d'accès à un « Espace personnel », et un dernier renvoyant à son « Panier ».

Au centre de la page, on trouve un espace « Découvrir », qui s'avère être un grand slide de quatre diapositifs. Un bouton « Collections », permet de remplacer ces diapositifs par sept

autres slides, cette fois-ci cliquables, permettant d'accéder à des collections déterminées (« La Première Guerre Mondiale », « Le Japon », « Les archives de la Bastille » …).

La page d'accueil est construite en longueur. Le menu déroulant permet d'accéder à d'autres rubriques, chacune possédant un fond différent et ayant son propre code couleur. Sous l'onglet « Découvrir », se trouve la rubrique « Nouveaux documents ». Celle-ci comporte un slide de documents récents. Un bouton « dernières mises en ligne » permet d'accéder à une recherche classée selon sa date de mise en ligne, et un onglet EPUB permet la même action concernant le format éponyme (format ouvert standardisé pour les livres numériques).

En « scrollant » encore, nous arrivons sur la rubrique « Actualités », qui propose des liens sur les derniers billets du blog de Gallica, et une infographie des contenus de la bibliothèque. Vient ensuite la catégorie « Autour de Gallica », dans laquelle se trouvent différents liens vers les comptes des réseaux sociaux investis par Gallica, ainsi qu'un encart « Du côté des Gallicanautes » (nous y reviendrons par la suite). En continuant de descendre sur la page d'accueil, nous découvrons les rubriques « Partenaires », « Services mobiles » (Gallica disposant d'une application et d'un site spécial pour les mobiles), et enfin « Contribuer ». Cette dernière permet de contacter Gallica (pour signaler une anomalie, obtenir de l'aide pour accéder à un document etc), et/ou « d'adopter » un document en finançant la numérisation d'un livre ou d'un enregistrement sonore. A noter que le don est ensuite déductible des impôts à hauteur de 66%. Enfin, l'internaute peut « en savoir plus », via l'onglet les « Amis de la BnF ».



Le haut de la page d'accueil de Gallica.

Cet aperçu des deux pages d'accueil témoigne de l'opposition entre la simplicité, la sobriété, et l'efficacité, dont souhaite faire preuve Google, en rendant son interface Google Livres très proche de celle de son moteur de recherche, et la structuration forte de Gallica, qui découpe sa longue page d'accueil en différentes rubriques, et qui joue sur de nombreux éléments visuels et codes de couleurs.

2.2 Les « espaces personnels »

Le second bouton de la page d'accueil Google Livres, sous la barre de recherche, se nomme « Ma bibliothèque ». Celui-ci permet de constituer une bibliothèque virtuelle et personnelle d'ouvrages complets (gratuits ou payants). Chaque utilisateur peut regrouper, et classer des livres entre eux, pour créer divers corpus grâce à la création « d'étagères ». Il s'agit donc d'un espace personnel.

Les utilisateurs doivent, en amont, disposer d'un compte chez Google. Pour ce faire, il est nécessaire de posséder une adresse mail, ainsi qu'un mot de passe. Le compte peut être enrichi par l'ajout d'informations personnelles (âge, loisirs, lieu de naissance, études,

profession etc…), sans que cela soit obligatoire. Cette dynamique s'ancre dans la volonté de Google, d'interconnecter ses différents services (Google +, Gmail, Google Play, YouTube etc…).

En haut de la page, on retrouve une barre de recherche similaire à celle de la page d'accueil. Cette dernière permet d'effectuer une recherche sur l'intégralité du site.

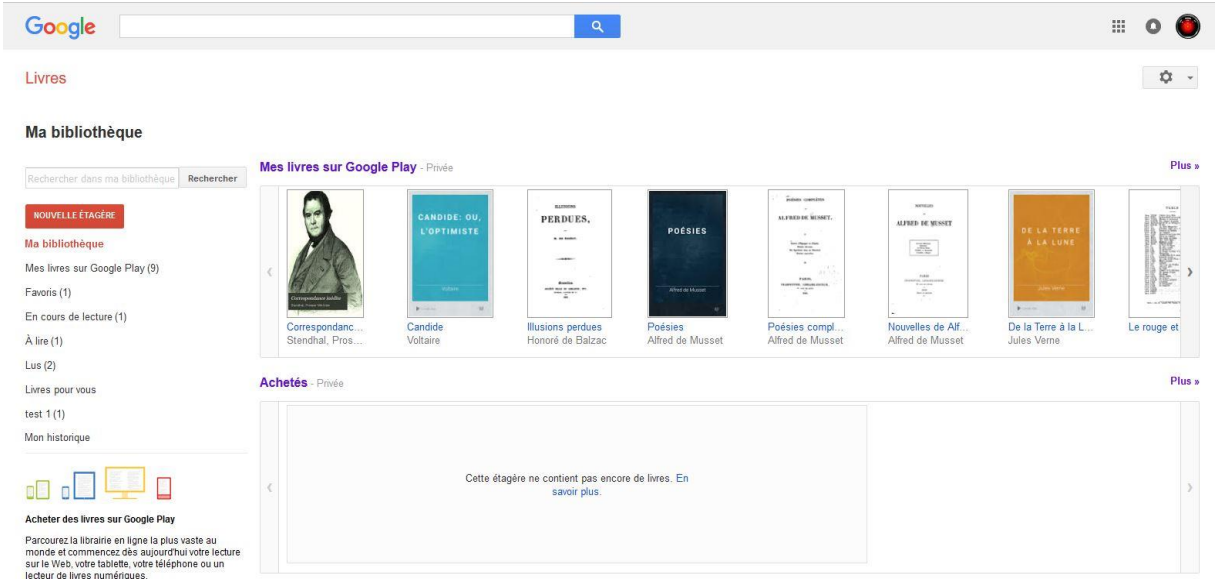
Sur la gauche, se trouve un encart avec le profil Google+ de l'utilisateur : un clic sur l'avatar renvoie sur le profil Google+ de ce dernier. Juste en dessous, une seconde barre de recherche permet d'effectuer des requêtes uniquement dans la section « Ma bibliothèque ». Nous y reviendrons par la suite. Plus bas, un bouton « Nouvelle étagère » permet de créer des étagères, puis une liste des étagères de l'utilisateur se termine par en encart « Acheter des livres sur Google Play ». Celui-ci renvoie évidemment sur le site de la fameuse librairie.

Au centre de la page, on trouve donc des « étagères», qui permettent à l'utilisateur d'organiser sa bibliothèque virtuelle. Une partie d'entre elles est déjà créée automatiquement. Ces notamment le cas de « Mes livres sur Google Play », « Favoris », ou encore « En cours de lecture ». Elles ne peuvent pas être supprimées. Les utilisateurs peuvent créer de nouvelles étagères, les nommer et renommer selon leur convenance, ajouter des descriptions, les rendre publiques ou privées, les trier, les vider, les supprimer, ou encore les exporter au format XML. Les étagères sont publiques ou privées. Dans le premier cas, elles ne sont visibles que par l'utilisateur, dans le second, elles peuvent être partagées via Google+, et donc être visibles par les contacts Google de l'utilisateur. Une étagère par défaut, intitulée « Livres pour vous », conseille des ouvrages en s'inspirant des achats et des historiques de lecture des internautes. Les étagères permettent donc à l'utilisateur d'organiser sa collection d'ouvrages à la manière d'un cabinet de lecture. Il est possible d'effectuer des recherches via la barre de navigation « Rechercher dans ma bibliothèque ». La recherche s'effectue uniquement, comme son nom l'indique, parmi les ouvrages présents dans la bibliothèque de l'utilisateur, et non sur l'ensemble de la base de données de Google Livres. Les résultats affichent également, par défaut, les livres présents dans l'historique de navigation, probablement dans une visée de captation (objectif marchand). Il est toutefois, par la suite, possible d'affiner la recherche pour les évincer. Il est important de noter que la première étagère qui apparait dans la bibliothèque s'intitule « Mes Livres sur Google Play ». Cette

dernière n'est pas supprimable, et apparait systématiquement en amont des autres étagères. Un clic sur le nom de l'étagère renvoie sur le site de la librairie. Nous pensons que ce choix est fait dans le but de renvoyer l'utilisateur vers la librairie Google Play, et ce le plus rapidement possible.

En haut à droite de la page, nous retrouvons, dans un bandeau, les icones qui apparaissent sur différents services (applications Google, notifications Google et Compte Google+). Juste en dessous, on aperçoit un bouton en forme d'écrou, qui contient trois liens : « Ma bibliothèque » (qui renvoie vers ma bibliothèque… qui est donc la même page), « Aide », et « Recherche avancée de Livres ».

Enfin, Google Livres permet d'importer des livres dans « Ma bibliothèque ».



Capture d'écran de la partie « Ma bibliothèque » de Google Livres.

Gallica dispose d'un service similaire dénommé « Espace personnel ». Comme mentionné lors de la description de la page d'accueil, il faut, afin de se rendre sur l'espace personnel, dérouler le bandeau « accéder au menu ». Il est nécessaire de créer un compte pour profiter de cet espace. L'espace personnel s'ouvre sur une nouvelle page, et se décline en parties distinctes pour gérer différentes activités relatives à la BnF. Gallica présente les différentes fonctionnalités de l'espace personnel ainsi :

- Utilisez-le pour marquer des pages sur Gallica, enregistrer vos recherches favorites, classer vos notices sur le Catalogue général,

- Acheter des reproductions en ligne,

- Réserver vos places et documents,

- Renouveler votre carte en ligne en bénéficiant d'une réduction de 10%.

La partie de l'espace personnel qui nous intéresse ici, se nomme « Mon Gallica », et recense les documents que l'utilisateur a ajouté à son espace personnel, et/ou ceux qu'il a « marqué » à l'aide de l'option « marque page ». Ces opérations s'effectuent dans l'espace de visualisation de documents (sur lequel nous reviendrons par la suite). Tous les documents sont ensuite répertoriés dans l'espace « Mon Gallica ».

L'espace « Mon Gallica » est assez similaire à celui de « Ma bibliothèque », proposé par Google Livres. En effet, on retrouve une colonne sur la gauche, qui nous permet de naviguer entre les documents et de les classer, suivie de liens vers les autres fonctions de l'espace personnel (« Mon catalogue général », « Mes réservations & ma carte », « Mes achats »). Au centre, les documents sont répertoriés sous forme de liste. Chaque document est représenté par une image, son titre, son type (manuscrit, carte…), un accès vers les informations détaillées, un accès au document, et un accès pour le « feuilleter ». Il est aussi précisé dans quel dossier est rangé chaque document. En haut de cette liste, nous trouvons des boutons permettant de les supprimer, les copier, les déplacer, les envoyer par courriel, mais aussi de les trier selon la date d'édition, l'auteur, la date d'ajout, et le titre. C'est ici que l'utilisateur organise et classe les documents. Plus au-dessus encore, on trouve une barre de recherche (les requêtes concernent tout Gallica, contrairement à celle de Google Livres permettant d'effectuer des recherches uniquement sur les ouvrages de « Ma bibliothèque), suivie d'un lien pour effectuer une recherche avancée.

Une option « Gérer mes dossiers » permet de créer et renommer des dossiers, ainsi que de gérer leur arborescence (possibilité de constituer des sous-dossiers). La même fonctionnalité est disponible avec les marques pages (« gérer mes marques pages »). « Mes

préférences » permet de paramétrer les préférences d'affichage général, et de trier les services « mes documents » et « marques pages ».

Capture d'écran de l'espace « Mon Gallica ».

Nous constatons donc deux services similaires entre Google Livres et Gallica, à la différence notable que Google insère de la publicité ciblée (étagère « livres pour vous »), et propose des étagères (qui sont l'équivalent des dossiers sur Gallica) non supprimables. Parmi elles, on trouve l'étagère « Livres sur Google Play », positionnée en amont des autres. Google Livres permet également d'effectuer une recherche dans l'intégralité de la bibliothèque de l'utilisateur.

2.3 Les visualiseurs

Après avoir effectué une recherche et trouvé l'ouvrage souhaité sur Google Livres, l'utilisateur clique sur ce dernier, et arrive sur une nouvelle page. Sur la gauche, on aperçoit un lien pour acheter la version numérique du livre, une option qui renvoie vers des distributeurs physiques (« obtenir la version papier de ce livre »), et une qui propose de trouver ce livre dans une bibliothèque. En dessous, on trouve la première de couverture du livre en miniature, une notation sur 5 étoiles qui est une moyenne des avis des utilisateurs, la possibilité de rédiger également un avis, une fonction « recherche », qui permet une recherche sur l'ensemble de l'ouvrage, un lien « à propos de ce livre », un accès à « ma bibliothèque », à « mon historique », à Google Play, et aux conditions d'utilisation. En haut,

on peut utiliser diverses fonctions de mise en page (parfois non disponibles), qui permettent de zoomer, de passer d'un affichage « page simple », à un affichage « double page ». Viennent ensuite une fonction qui permet de copier le lien URL, un bouton visant à ajouter l'ouvrage à sa bibliothèque, puis un dernier qui permet de rédiger un commentaire. Sur la droite, deux boutons servent à tourner les pages. Un autre onglet permet d'accéder à l'aide, à la recherche, à « ma bibliothèque », et même, selon les cas, de télécharger l'ouvrage en version PDF ou EPUB.

Fonctionnalités du visualiseur de Google Livres

Le contenu du livre affiché dépend, bien sûr, des conditions juridiques établies au préalable (ouvrage complet, aperçu, courts extraits, aucun extrait…). L'affichage que nous propose Google Livres est ici très minimaliste, et offre assez peu d'interaction. Nous allons, à présent, nous rendre compte que Google cherche à attirer le lecteur vers sa librairie Google Play, plateforme sur laquelle les livres sont lus et bénéficient de fonctionnalités additionnelles.

Lorsqu'un utilisateur fait l'acquisition d'un livre, même lorsqu'il s'agit d'un livre gratuit, il est ajouté à la bibliothèque de ce dernier. Pour les consulter, Google fait naviguer l'utilisateur de sa bibliothèque Google Livres, à un visualiseur, qui est en fait celui de Google Play Livres. Beaucoup plus esthétique que celui de Google Livres, qui sert uniquement à feuilleter les ouvrages et à faire de brèves recherches, le visualiseur de Google Play Livres offre également de nombreuses nouvelles fonctionnalités, qui se situent en haut à droite de la page de visualisation d'un ouvrage, et qui se divisent en cinq différents pictogrammes.

Premièrement, l'accès à un sommaire permet d'avoir accès, comme son nom l'indique, au sommaire du livre, mais aussi à un sommaire des « passages favoris » de l'utilisateur, qu'il peut marquer avec une option « marque page ». L'utilisateur peut, en outre, se rendre vers un sommaire des notes, qu'il peut créer un sélectionnant du texte, et à des « fiches info », que seuls certains ouvrages contiennent.

Viennent ensuite les options d'affichages. Ici, il convient de distinguer les ouvrages disponibles en format PDF, EPUB ou, comme c'est généralement le cas, surtout avec les livres plus récents, les ouvrages disponibles aux deux formats.

Le format PDF est un format image. Il propose la visualisation du texte d'origine (appelé « page d'origine » par Google). Le format EPUB, quant à lui, permet de personnaliser l'aspect visuel du texte. L'utilisateur peut ainsi redimensionner le texte, modifier la police, la mise en page, l'interlignage, ou encore la taille des caractères. Particulièrement adapté pour les smartphones, les netbooks, et les appareils de lecture de livres numériques, l'EPUB permet également, sur ces plates formes, de modifier la couleur du fond et de la police utilisée. Pour passer d'une page à l'autre, l'utilisateur « tourne les pages » tactilement sur son écran, et le site (ainsi que l'application), simulent visuellement les pages, qui se courbent et se tournent. Lors d'une lecture sur ordinateur, ce dispositif visuel n'existe pas, et l'utilisateur passe d'une page à l'autre en cliquant sur des boutons prévus à cet effet, ou en utilisant les flèches de son clavier.

On trouve ensuite un bouton « recherche », qui permet encore une fois de faire une recherche au sein de l'ouvrage, ainsi qu'un bouton « aide », qui ouvre l'accès à différents liens hypertextes (« à propos de ce livre » que nous avons évoqué précédemment, « centre d'aide Google Play », « Utilisation du lecteur Web, « Signaler un contenu choquant » et « Signaler un problème »). Puis, un dernier dispositif sert à marquer des pages (toutes les pages marquées sont répertoriées dans le sommaire). Enfin, en bas de page, se trouvent les flèches permettant d'accéder à la page suivante ou précédente. Une barre permet également de visualiser le nombre de pages que contient l'ouvrage, et de voir où le lecteur se situe dans ce dernier.

Pour finir, une autre option de visualiseur, particulièrement intéressante, consiste en la possibilité, lorsque l'utilisateur met en surbrillance un mot ou une suite de mots, d'accéder à différentes fonctionnalités. La fonction « Définir » ouvre un petit encart au sein duquel est défini le mot, en se basant sur une recherche Google.

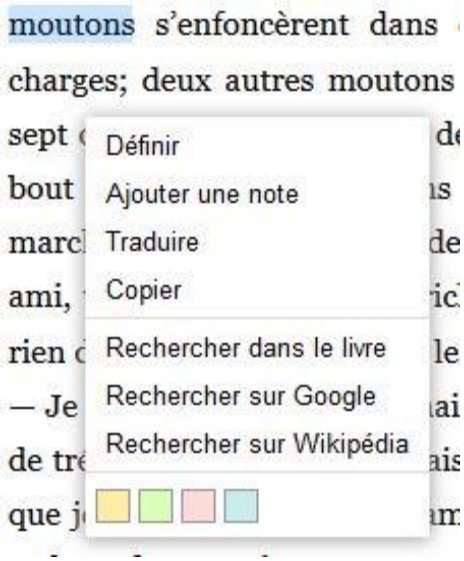
Le lien hypertexte « plus » permet d'ouvrir une recherche de ce mot sur une page web complète. « Ajouter une note », comme son nom l'indique, vise à ajouter des informations sur

le mot en question. Une fenêtre de texte s'ouvre, et l'utilisateur renseigne les informations qu'il souhaite (importance d'un personnage par exemple). Le mot reste ensuite surligné, pour que l'utilisateur puisse y revenir. Un code de quatre couleurs est disponible pour organiser les notes. Ces dernières sont retrouvables dans le sommaire.

Ensuite, l'option « Traduire » ouvre une fenêtre permettant la traduction du mot en surbrillance, dans une langue choisie. Il s'agit de l'intégration du module Google Play Livres avec Google Traduction. Un dispositif « Copier » permet de copier le mot. Enfin, trois fonctions « recherches » sont disponibles, toujours lorsque l'utilisateur met en surbrillance un mot, ou un ensemble de mots. « Rechercher dans le livre » ouvre une fenêtre sur la droite qui affiche tous les passages de l'ouvrage dans lesquels le mot, ou groupe de mots, est cité. « Rechercher sur Google » et « Rechercher sur Wikipédia », sont deux rubriques qui renvoient vers de nouvelles fenêtres. Nous signalons également que, bien que cette option ait l'air d'avoir été supprimée, Google permettait également d'interconnecter les services Livres et Maps, et ce dans le but de permettre l'affichage des lieux cités dans un livre sur la carte géographique de Google Maps… : un résultat « étonnant et convaincant, mais non pas sans erreurs »[72].

Si le visualiseur de Google Livres est très classique, celui de Google Play Livres, en revanche, est beaucoup plus développé. Tout en restant dans une esthétique très sobre (l'un des leitmotivs de la charte graphique de Google), il permet d'accéder à de nombreuses autres fonctionnalités. Parmi elles, on en retrouve identiques à celles que proposent le format EPUB, pour faciliter et personnaliser la lecture, ainsi que l'interaction entre différents services Google (moteur de recherche, traduction). Particulièrement adapté aux smartphones et aux tablettes, disponible à la fois en tant que site internet et application, le visualiseur Google Play est pensé comme un véritable support de lecture.

[72] Gioogle livres p87



Fonctionnalités du visualiseur Google Play Livres lorsque l'utilisateur met un mot en surbrillance

Le visualiseur de documents actuel de Gallica, est en ligne depuis 2015. Il s'agit d'une amélioration de l'ancien visualiseur, se voulant plus accessible. En effet, une enquête d'usage et de satisfaction menée auprès des utilisateurs a montré « la nécessité de simplifier les modalités de consultation des documents numérisés »[73].

Cette nouvelle version est organisée de manière assez classique. Si le bandeau de pictogrammes se situe en haut à droite et en longueur sur Google Livres, il est placé au centre à gauche, et en longueur sur Gallica. Le document consulté apparait au centre comme sur Play Livres. Les boutons pour passer d'une page à l'autre, sont situés dans un petit bandeau en longueur situé à droite de la page, et encadré par deux autres boutons qui permettent de se rendre, en un clic, aux extrémités des ouvrages.

Le bandeau positionné sur la gauche est composé de 12 pictogrammes. Lorsque l'utilisateur clique sur l'un des pictogrammes, un nouveau bandeau s'ouvre à gauche du

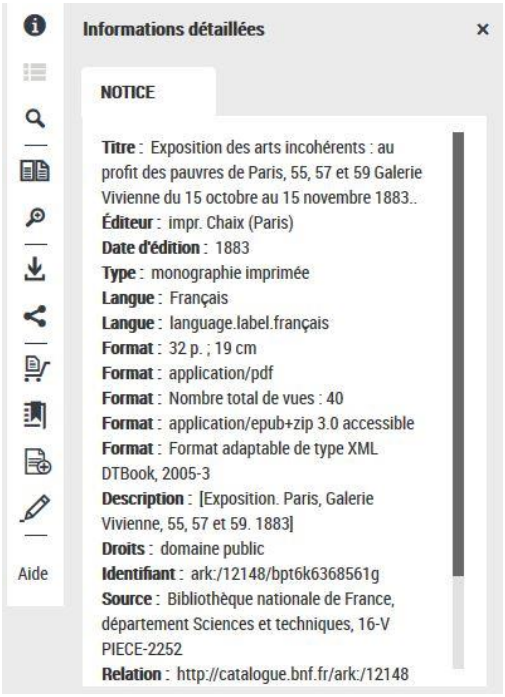
[73] http://blog.bnf.fr/gallica/index.php/2015/01/26/sur-gallica-labs-un-nouveau-visualiseur-de-documents/ [Consulté le 08/06/2016]

document. L'accès à toutes les informations et toutes les modifications peut donc se faire en restant sur la page du visualiseur.

Le premier permet d'afficher les informations détaillées concernant le document. Celles-ci apparaissent entre le bandeau et le document. Elles sont, la plupart du temps, très complètes (titre, éditeur, date d'édition, type, langue, format, description, droit, identifiant, source, relation provenance, date de mise en ligne), plus que sur Google Livres, et sont visibles à côté du document, tandis que Google les affiche en ouvrant une nouvelle page. Sous les informations détaillées, on trouve le bouton «table des matières », qui n'est pas disponible pour tous les documents. L'onglet « rechercher dans ce document » ouvre une barre de recherche pour effectuer des requêtes. Cette option est similaire à celle de Google Livres. La rubrique « Affichage », permet de choisir entre la simple page, la double page, le défilement vertical, la mosaïque, la superposition du mode texte et image (utile pour faciliter la lecture de manuscrits par exemple), et le mode vocal. La fonction « zoom » permet, bien entendu, de zoomer, mais également de passer en mode « plein écran », et d'appliquer des rotations à 90° (horaires ou antihoraires), au document. L'option « téléchargement et impression » propose plusieurs formats de fichiers (PDF, JPEG, TXT, EPUB, Format adaptable…), afin de télécharger une partie, ou l'intégralité, du document.

Ensuite, le bouton «partage et envoi par courriel » permet de copier l'URL précise du document entier ou de la page consultée, et de le partager sur les réseaux sociaux, ou par courriel. « Acheter une reproduction » vise l'accès à des fonctions pour acheter des reproductions d'une partie du document, ou pour acheter des réimpressions à l'identique (via deux sites partenaires de Gallica). «Marque page » et « Ajouter ce document » permettent, en un clic, de répertorier une page ou l'intégralité du document dans « Mon Gallica ». Pour finir, les options « Signalement d'une anomalie », et « Aide » fournissent la signification de chaque pictogramme et de chaque option.

Très abouti également, le visualiseur de Gallica est adapté à la consultation de documents variés (manuscrits, cartes, player intégré pour consulter des contenus audivisuels…). Contrairement à Google Livres, le document source est particulièrement mis en évidence et valorisé sur Gallica.



Fonctionnalités du visualiseur de Gallica.

Les deux visualiseurs sont simples d'utilisation, et permettent une consultation optimale des documents. Ils répondent toutefois à des objectifs différents. Le visualiseur de Google Play Livres est un véritable support de lecture, cherchant à optimiser au maximum le confort et l'expérience de lecture de l'utilisateur, tandis que celui de Gallica met avant tout en valeur le document en lui-même, l'expérience pour l'utilisateur se rapprochant donc davantage d'une véritable consultation virtuelle d'archives.

2.4 La fonction recherche

Nous avons effectué une recherche très classique au sein des deux bibliothèques. En rentrant le terme « Louis Aragon », fameux poète, romancier et journaliste français du XXème siècle, on trouve « environ 250 000 » résultats dans Google Livres, et 12 192 résultats dans Gallica.

Le premier lien que propose Google, est un lien vers la page Wikipédia de l'auteur. S'en suit alors une longue liste d'ouvrages, les résultats étant des œuvres d'Aragon (principalement

sa poésie dans les premiers choix), ou encore des textes et ouvrages traitant de ce dernier. Nous remarquons, avec étonnement, que son roman « Aurélien », qui est reconnu comme l'une des œuvres les plus significatives de l'auteur, n'apparait pas avant la 6ème page, et ce à la suite de plusieurs écrits concernant Aragon.

Dans Gallica, le premier résultat ne concerne pas l'auteur Louis Aragon, mais la bibliothèque du même nom, située à Amiens (il s'agit du catalogue de la bibliothèque léguée par M. Cozettte – 1844), l'auteur étant identifié comme « Bibliothèque Louis Aragon ». Les résultats qui suivent sont de nature très variée. En effet, on y trouve des extraits audio de poèmes de l'auteur, des cartes, des extraits de journaux, qui ne concernent pas tout le temps l'écrivain, mais on ne trouve aucun de ses ouvrages.

Cela témoigne, en premier lieu, du fameux « trou noir du XXème siècle », qui fait défaut aux bibliothèques numériques aujourd'hui, puis ensuite, pour les utilisateurs ayant un objectif de recherche précis, de la nécessité de perfectionner sa requête en employant des mots-clés, et des filtres adéquats. Un bandeau sur la gauche, permet d'ailleurs, d'affiner considérablement la recherche (type de document, site de consultation, auteur, date d'édition, thème, version, langue, mode de texte, type d'accès). Ceci témoigne également de l'importance de l'indexation et des métadonnées dans la recherche documentaire.

Arrêtons-nous maintenant sur la fonction « A propos de ce livre », et proposée par Google Livres après que l'utilisateur ait effectué une recherche et sélectionné un ouvrage. En plus de retrouver la possibilité d'acheter l'ouvrage et l'existence de la première de couverture, du titre, de la notation du livre et de la fonction « Recherche » à l'intérieur de l'ouvrage, nous constatons la présence des avis des internautes. Il s'agit, plus précisément, d'extraits de commentaires, issus du site de Google Livres, ou de celui des revendeurs partenaires qui vendent l'ouvrage en question. L'option « Consulter l'avis complet » renvoie d'ailleurs sur le site d'où provient l'avis. Sous ces avis, Google met l'ouvrage en lien avec d'autres livres, sous la forme d'un slide intitulé « Livres sur des sujets connexes ». Il s'agit en fait, principalement, d'autres ouvrages du même auteur. En adoptant la même mise en forme, Google Livres propose également des « Pages sélectionnées » (souvent la page de titre et la table des matières), qui constituent d'autres éditions du même ouvrage.

Pour affiner la présentation du livre, Google propose un nuage de mots, des « Expressions et termes fréquents », qui résulte d'un calcul de la fréquence des mots significatifs dans l'ouvrage (plus la police du mot ou de l'expression est grande, plus il est récurrent dans le livre). Lorsque l'on clique sur l'un des mots, nous sommes renvoyés vers les passages du livre contenant ce dernier. Cette fonction permet à l'utilisateur d'avoir un rapide aperçu des thématiques du livre. S'en suit un autre affichage, particulièrement intéressant, intitulé « Fréquemment cité ». En effet, Google affiche les passages du livre cités par d'autres ouvrages contenus dans sa base de données, créant ainsi, automatiquement, des relations de sens entre différents ouvrages (ouvrages cités par d'autres ouvrages, différentes éditions du même ouvrage). Très intéressante pour les curieux et les chercheurs souhaitant mettre en relations différents livres, cette fonctionnalité ne nous semble malheureusement pas assez exploitée, car elle se limite à dix passages cités au maximum. Cela reste toutefois un outil puissant, surtout lorsque l'on prend en considération la quantité de livres contenus dans Google Livres. Enfin, nous retrouvons des informations bibliographiques. Ces dernières nous informent quant à la provenance de l'original, ainsi que concernant sa date de numérisation et son éditeur.

Capture d'écran d'une partie de la page « A propos de ce livre » de l'ouvrage La Comédie Humaine d'Honoré de Balzac.

Enfin, pour conclure la partie sur les fonctionnalités de recherche, nous mentionnons la possibilité de faire des « recherches avancées » sur les deux sites. Sur Google Livres, la « Recherche avancée de Livres » n'est pas particulièrement mise en avant. Non disponible sur la page d'accueil, elle n'est accessible qu'après avoir fait une première recherche. De plus, il

faut, pour y accéder, cliquer sur le bouton « Options », qui se situe en haut à droite de la page. La recherche avancée de Google Livres est simple et puissante, à l'image de son moteur de recherche. Les fonctionnalités classiques de la recherche booléenne sont, bien sûr, applicables, et Google permet, entre autres, d'affiner ses recherches par langue, éditeur, auteur, titre, date de publication, numéro ISBN et ISSN.

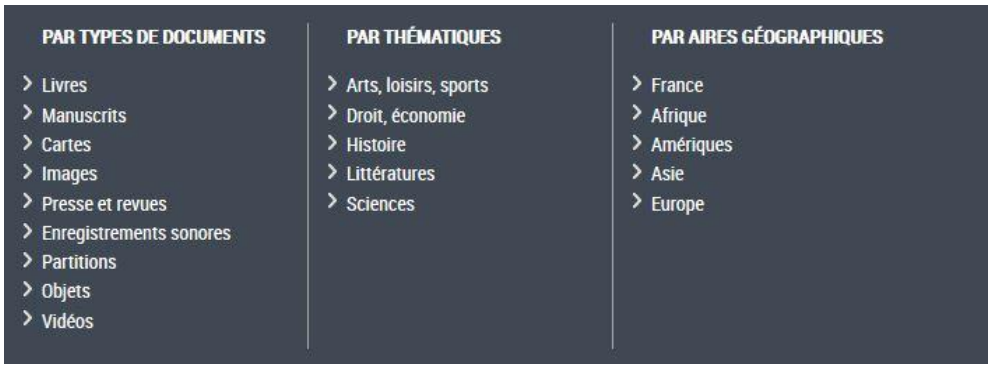
Exemple de paramètres de la fonction recherche avancée de Google Livres.

La fonction recherche avancée est disponible et visible dès la page d'accueil sur Gallica. Assez similaire à celle de Google, elle permet également de rechercher par type de documents, par thème, par bibliothèque et par format (EPUB, mode texte, 3D).

3. Gallica et l'éditorialisation

3.1 Les collections Gallica

La principale différence entre Googles Livres et les bibliothèques numériques, constitue, avant tout et selon nous, l'absence d'éditorialisation sur le site de l'entreprise Californienne. Gallica, au contraire, possède un modèle pertinent d'organisation, de structuration, et de mise en relation de contenus. Le site est énorme, et il peut s'avérer difficile de naviguer parmi les nombreux contenus disponibles. Gallica a donc regroupé différents documents sous forme de collections, afin de les articuler entre eux de manière logique. Comme mentionné lors de la description de la page d'accueil, un bouton « Collections » ouvre un encart permettant d'accéder à diverses collections, regroupées par leur type, leur thématique, ou leur aire géographique.



Le bandeau Gallica permettant de choisir entre les collections

On note que, pour l'instant, les collections par aires géographiques ne sont pas encore disponibles, à l'exception de la France et de l'Asie.

La thématique « Histoire », nous amène sur une nouvelle pag,e au sein de laquelle se trouve une présentation de la thématique, ainsi qu'un choix entre différentes sous-thématiques. Dans le cas de la catégorie « Histoire », les sous-catégories proposées sont : « La Première Guerre Mondiale », et « Ressources généalogiques ». La sous-catégorie « Première Guerre mondiale », se décline ensuite en différents sujets (« Chansons et musiques », « Discours d'hommes politiques », « Journaux de tranchées », « La guerre en images », « Les armées françaises dans la Grande Guerre », « Les bulletins de presse étrangère » et « Revues illustrées »). Ces derniers, peuvent également avoir des déclinaisons (pour l'exemple, « Journaux de tranchés » se découpe en Journaux de l'infanterie, de l'artillerie, de la cavalerie, du génie, des chasseurs etc…). Les contenus ainsi agencés en corpus, permettent à l'utilisateur de naviguer entre eux, et ce en suivant un parcours structuré, et largement pensé en amont.

Autre cas encore plus saisissant, Gallica propose également une collection intitulée « Les Essentiels de la littérature ». On accède à cette rubrique via les collections dédiées à la littérature de la bibliothèque. Cette dernière, a la particularité de posséder son site internet propre[74], et d'être pensée spécialement à destination des publics scolaires et des enseignants, afin d'offrir un accès simple aux « essentiels » du patrimoine littéraire français. Pour l'instant centré sur le XVIIIème siècle, l'objectif premier est donc de faire découvrir des auteurs

[74] http://gallica.bnf.fr/essentiels/ [Consulté le 08/06/2016]

(Rousseau, Du Chatelet, Voltaire, Diderot, Montesquieu), et des thèmes (Les femmes, L'utopie, L'esclavage, Le roman épistolaire, Les Lumières etc…), par le biais de parcours multimédia. Ces parcours mettent en relation différents documents, comme des corpus d'images, de manuscrits, des enregistrements sonores et visuels, ou encore des conférences en ligne.

La page d'accueil contient des liens pour accéder directement à certains thèmes et auteurs. En haut de la page, l'utilisateur peut choisir de se rendre sur sites de la BnF et de Gallica. Une « tirette » permet également d'afficher une barre de navigation, dans le but de faire des recherches dans tout Gallica. A côté, on trouve trois choix d'entrée pour explorer le site avec les boutons « chronologies », « thèmes », et « repères ». Ces derniers sont suivis d'une barre de recherche. Le centre de la page permet d'accéder à différents auteurs, thèmes, périodes et œuvres, via un slide et des boutons. Enfin, en bas de la page, on aperçoit deux liens renvoyant vers le site pédagogique de la BnF[75], ainsi que vers Eduthèque[76](celui-ci est lié au ministère de l'éducation nationale, et propose des ressources pédagogiques, culturelles et scientifiques pour les enseignants).

La page d'accueil du site 'Les Essentiels de la littérature » qui affiche quatre entrées différentes : le choix d'auteurs, d'œuvres, de périodes et de thèmes.

[75] http://classes.bnf.fr/index.php [Consulté le 08/06/2016]

[76] http://www.edutheque.fr/accueil.html [Consulté le 08/06/2016]

Chaque entrée du site permet de mettre en contexte des contenus variés. Ainsi, lorsque l'utilisateur se rend sur la partie « XVIIIème siècle » ; elle-même divisée ensuite en quatre périodes distinctes (-1685-1715 : Anciens contre Modernes, - 1715-1751 : L'âge des expériences, - 1751-1778 : Le triomphe des Lumières, - 1778-1800 : L'ère des révolutions), arrive sur une page très visuelle, sur laquelle les différentes œuvres et les différents auteurs sont organisés dans des tableaux chronologiques permettant une contextualisation forte (œuvres majeurs, repères historiques). Les contenus sont hiérarchisés. Les œuvres les plus significatives sont les plus visibles.

Capture d'écran du site sur la page 1751-1776.

3.2 Gallica et les réseaux sociaux

La bibliothèque Gallica investie largement les réseaux sociaux dans le cadre de sa politique d'éditorialisation. Cette présence, a pour but de promouvoir des contenus numériques et des services associés, et ce en tenant compte de l'évolution des pratiques du web, et en expérimentant de nouvelles formes d'interactivité avec les utilisateurs.

Gallica possède une page Facebook qui recueille plus de 115 350 mentions « j'aime ». Des documents y sont publiés à un intervalle assez régulier et continu (environ un post tous les 2

– 3 jours), afin de mettre en valeur la bibliothèque. Il peut s'agir d'un, ou plusieurs documents numérisés, de nature très variée. Le ton employé est assez libre, souvent humoristique, et tranche avec la froideur supposée que certains peuvent se faire du monde des bibliothèques. La plupart du temps, les publications se concluent par un lien vers le site de la bibliothèque, pour accéder à davantage de contenus.

Sur Facebook, Gallica met régulièrement en avant des documents en lien avec l'actualité, mais aussi d'autres dotés d'une nature plus insolite ou inattendue. La bibliothèque n'hésite pas à poser des devinettes aux internautes. Par exemple, on peut citer le « post » suivant : « Pour résoudre l'énigme du jour, il suffit d'identifier la langue dans laquelle ont été écrits ces manuscrits. Indiquez vos réponses dans les commentaires de chaque image ».

Exemple de publication sur la page Facebook de Gallica

Gallica surnomme les internautes du site les «Gallicanautes » (nous l'évoquions lors de la description de la page d'accueil de la bibliothèque), ce qui démontre d'une volonté de constituer une forme de communauté. Les réactions et commentaires des internautes sont très nombreux, et témoignent ainsi d'un véritable intérêt pour l'offre documentaire de Gallica, ainsi que pour ce mode de valorisation « sociale », qui permet de partager, discuter, et réagir à ces documents.

La bibliothèque numérique possède également, depuis Janvier 2010, un compte Twitter. Il comptabilise aujourd'hui 14 2000 tweets, et 41 400 abonnés. Tout comme pour Facebook, Gallica s'en sert pour mettre en lumière divers documents. En plus de cela, la bibliothèque retweet les « trouvailles » des internautes, leur réappropriation, et leur propre mise en valeur des contenus. De cette manière, ce compte valorise non seulement Gallica, mais aussi ses utilisateurs, et la manière dont ceux-ci mentionnent, citent, et commentent les documents numérisés par la bibliothèque. Nous notons la présence d'un hashtag #Gallicanautes.

Gallica possède également un compte Pinterest.

Enfin, nous mentionnons également le blog de Gallica, et plus particulièrement sa catégorie « Blog des Gallicanautes », qui donne la parole aux internautes. Ici, les utilisateurs présentent leurs projets ainsi que la manière dont ils utilisent la bibliothèque numérique. Ces présentations prennent la forme d'interviews.

Ce panorama de l'usage qu'a Gallica des réseaux sociaux, met en lumière sa stratégie de communication visant à renforcer sa visibilité et assurer la promotion de sa bibliothèque. Cette dernière a également pour dessein de favoriser la création d'une forme de communauté virtuelle. Ainsi, nous constatons de l'effort, appuyé, réalisé par Gallica en termes d'éditorialisation.

3.3 Gallica et le site « data.bnf »

Depuis 2011, la Bibliothèque Nationale de France a mis en ligne site data.bnf[77]. Celui-ci recense les pages de référence sur les auteurs, les œuvres et les thèmes de Gallica, ainsi que celles issues des différents catalogues bibliographiques. Cette forme d'éditorialisation permet d'accroitre la visibilité de ces données (faisant partie du « web profond »), en les rendant indexables par les moteurs de recherche. L'utilisateur peut se rendre sur le site dédié, et ainsi

[77] http://data.bnf.fr/ [Consulté le 08/06/2016]

accéder à ces données, sans passer par des catalogues ou portails spécifiques[78]. L'accueil du site précise que le projet « utilise les outils du Web sémantique, et s'inscrit dans une démarche d'ouverture des données ».

4. Conclusion de l'analyse

Cette analyse tend donc à montrer que les sites Google Livres et Gallica sont très différents, et répondent à des objectifs qui le sont tout autant.

La construction de Google Livres repose sur la même logique que celui du moteur de recherche Google, à savoir une interface minimaliste et épurée, qui se focalise avant tout sur la puissance de son moteur de recherche. Les requêtes dans « l'index de livres complets le plus fourni au monde » sont donc simples et effectives, et la recherche avancée permet de retrouver des ouvrages précis parmi l'immense corpus. L'espace « Ma bibliothèque » permet ensuite d'organiser et de structurer ses collections. La consultation d'ouvrages dans Google Livres est très rudimentaire, et nous comprenons vite, à la manière dont est agencé le site, que Google cherche à amener l'utilisateur le plus rapidement possible vers sa librairie Google Play Livres, qui contient, en plus de sa fonction de boutique numérique, un espace dédié à la lecture. Google permet, grâce, à ses algorithmes et à l'inter-connectivité de ses différents services, une multitude de fonctionnalités intéressantes visant notamment la création automatique de relations de sens entre différents ouvrages, ainsi que la personnalisation de l'expérience de lecture (notes personnelles, définition, traduction).

Le site de Gallica est, quant à lui, lui davantage représentatif des objectifs des bibliothèques numériques. Très grand et très fouillé, divisé en de nombreuses catégories bien distinctes, son rôle affiché est la mise en valeur des documents. Du visualiseur et sa portée

[78] LABORDERIE Arnaud, « Éditorialisation des bibliothèques numériques : le cas des Essentiels de Gallica », *Colloque International sur le Document Electronique : « Documents et dispositifs informationnels à l'ère post-numérique »*, 2015, p8.

expérientielle archivistique, jusqu'à sa structuration, en passant par ses services annexes (présence sur les réseaux sociaux, site data.bnf), le site de Gallica ne se contente pas de mettre en ligne des contenus, et se donne pour principale mission de les agencer et de les structurer de manière à les rendre accessibles et assimilables.

Conclusion générale

Derrière le projet de numérisation de Google, rappelant de prime abord les utopies culturelles d'une bibliothèque universelle et les rêves de Paul Otlet, l'émerveillement a rapidement fait place à la méfiance et à la critique devant des risques de monopoles industriels de plus en plus tangibles.

En douze années d'existence, Google Livres a eu, et possède toujours, un impact considérable sur le monde de la numérisation et sur celui des bibliothèques.

En effet, la numérisation est aujourd'hui une activité stratégique de premier plan pour les bibliothèques. Elle est vectrice de nombreux questionnements d'ordre techniques, financières, juridiques, idéologiques et bibliothéconomiques. L'arrivée de Google dans ce secteur en plein expansion a amené les bibliothèques à se modifier, se redéfinir et à se redynamiser. L'émergence de partenariats public/privé, l'accélération des processus de numérisation, la mise en relations entre bibliothèques à échelle nationale et internationale, le soin apporté à l'indexation et aux métadonnées, les nombreuses questions juridiques concernant la nature des documents numériques (œuvres orphelines), la prise en compte de documents de natures de plus en plus variées (manuscrits, cartes, partitions, documents audiovisuels…), la volonté de trouver des alternatives au « trou noir du XXème siècle » (ouverture de Gallica à l'édition contemporaine), ainsi que le soin apporté aux politiques d'éditorialisation sur le net, témoignent des nombreuses transformations des bibliothèques liées au numérique.

Les bibliothèques doivent s'adapter à ce nouvel environnement. Si une définition des bibliothèques numériques, phénomène encore en plein essor et dont les frontières sont encore floues et mouvantes, reste difficile à formuler, leur mission est de faire office de structures d'autorité et d'influence dans le milieu numérique. A l'heure où les moteurs de recherche et les algorithmes semblent régner en maîtres, les bibliothèques numériques trouvent leur essence dans la médiation et l'éditorialisation qu'elles effectuent sur leurs contenus. Elles doivent être pensées et structurées de manière à ce que l'accès aux documents

soit le plus cohérent, et pertinent possible. Les ressources doivent donc être agencées grâce à des relations de sens et des parcours informationnels structurés. Nous l'avons vu avec Gallica, modèle pionnier et représentatif des bibliothèques numériques, qui trouve ainsi sa légitimité et son autorité dans sa mission centrale de donner du sens à ses contenu.

En parallèle, le développement de Google Livres, des premiers partenariats avec de grandes bibliothèques américaines au lancement de la librairie Google Play Livres, en passant par les nombreuses batailles juridiques et idéologiques qu'a engendré la numérisation « sauvage » de livres sous droits, ont rapidement montré les véritables enjeux de ce projet pour l'entreprise californienne. Nous avons vu que, contrairement à Gallica, Google Livres ne dispose pas de politique documentaire, et devient donc difficilement qualifiable de bibliothèque numérique. Les objectifs de Google, d'abord présentés comme philanthropes sont, bien entendu, avant tout de nature économique. Adepte du « soft power », parfait exemple du capitalisme cognitif, l'entreprise a probablement largement pensé en amont le développement de son projet. S'il est impossible d'affirmer que Google avait, dès 2004, l'intention de créer sa librairie Google Play Livres, les exclusivités demandées dans les premiers contrats de numérisation avec les bibliothèques tendent à laisser plausible un tel scénario. Plus largement, Google Livres s'inscrit dans la stratégie financière et politique globale de Google, avec ses avantages, mais également, (et surtout ?), ses controverses bien connues (recueil de données personnelles, position largement dominante dans l'univers de la publicité contextuelle etc…).

Google Livres et le développement des bibliothèques numériques et de la numérisation sont donc intrinsèquement liés, bien que leurs objectifs soient différents.

Si nous nous sommes focalisés sur l'impact de Google Livres sur le monde des bibliothèques, il est également clair que le sujet touche également au monde de l'édition, et directement aux auteurs (les Programmes Partenaires de Google Livres auprès des auteurs et des éditeurs ne sont qu'un exemple), apportant ainsi de nouveaux questionnements relatifs à la rémunération des auteurs, aux canaux de diffusion, ou encore au développement des Creative Commons.

Afin d'apporter une continuité à ce travail, nous pourrions nous focaliser sur l'objet « livre » en lui-même. Avec le numérique, notre rapport avec le document a évolué. Google,

avec les fonctionnalités de Google Livres (structuration de collection numérique, recherche intra-texte au sein de l'ensemble de la collection, relations automatiques entre documents avec liens automatiques d'ouvrages cités par d'autres ouvrages), et du visualiseur Google Play Livres (inter-connectivité avec son moteur de recherche et Google Traduction, son système de note, de marque page, de recherche intra-texte, personnalisation de l'interface grace au texte flottant), modifie probablement l'expérience de lecture de l'utilisateur. Si nous manquons de données empiriques à ce sujet, nous constatons, en nous basant sur l'étude qualitative de Muriel Amar et de Cécile Toutoui concernant l'usage des corpus patrimoniaux numérisés à l'heure de Google Books[79] (étude menée entre fin 2010 et début 2011), que parmi les 15 interrogés, utilisateurs intensifs de bibliothèques numériques, tous avaient utilisé Google Livres au moins une fois, la plupart l'utilisant régulièrement, voire systématiquement.

Les avis concernant le projet sont très contrastés. En effet, certains déplorent l'absence de politique documentaire, tandis que d'autres le considère comme une nouvelle norme (« Il faut bien dire que tout le monde utilise Google Books. Maintenant ça devient la chose qu'on utilise le plus »[80]). Un interviewé reproche également à Gallica sa complexité, comparée à l'interface épurée et la quantité de documents très facilement accessibles sur Google Livres. Depuis cette étude, les bibliothèques numériques, tout comme Google Livres, ont évolué (Gallica possède d'ailleurs une nouvelle interface et un nouveau visualiseur depuis 2015, tandis que Google Play Livres et les fonctionnalités de son visualiseur évoquées précédemment ont été lancées en 2012).

Il nous semble intéressant de nous pencher sur les usagers de ces services. Qui sont-ils ? Des universitaires, des chercheurs, des passionnés ? Que cherchent-ils ? Comment naviguent-ils sur ces sites ? Comment utilisent-ils ces nouvelles interfaces de lecture et de consultation de documents ? Quelle méthodologie utilisent-ils ? Quelles fonctionnalités exploitent-ils et pourquoi ? Quelles valeurs attribuent-ils à ces documents numériques ? Gallica est-il considéré comme trop vaste et complexe à prendre en main ? Google Livres est-

[79] PAPY Fabrice (dir), *Recherches ouvertes sur le numérique*, Lavoisier, Paris, 2013, p117.

[80] PAPY Fabrice (dir), *Recherches ouvertes sur le numérique*, Lavoisier, Paris, 2013, p121.

il considéré comme pertinent pour les chercheurs et les universitaires ? Les fonctionnalités de Google permettant d'effectuer des recherches textuelles sur un corpus d'ouvrages et les relations entre ouvrages cités par d'autres ouvrages sont-elles intéressantes pour les chercheurs ou s'agit-il seulement de « gadgets » ? Les utilisateurs de Googles Livres sont-ils aussi utilisateurs d'autres bibliothèques numériques et inversement ? Les utilisateurs de bibliothèques numériques comme Gallica sont-ils « anti-Google Livres » ?

Ces questionnements nous permettraient de comprendre comment les usagers se familiarisent et s'adaptent à ces dispositifs numériques ; leurs pratiques et attentes contribuant probablement également aux transformations et aux évolutions de ce secteur encore en plein développement.

Bibliographie

- Articles en ligne :

LABORDERIE Arnaud, « Éditorialisation des bibliothèques numériques : le cas des Essentiels de Gallica », *Colloque International sur le Document Electronique : « Documents et dispositifs informationnels à l'ère post-numérique »*, 2015, Disponible sur le Web : https://hal-bnf.archives-ouvertes.fr/hal-01239425/document [Consulté le 08/06/2016]

LAGOZE Carl, KRAFFT B. Dean, PAYETTE Sandy et JESUGORA Susan, « Qu'est-ce qu'une bibliothèque numérique, au juste ? Au-delà des fonctions recherche et accès dans la National Science Digital Library », *AMETIST : Appropriation, Mutualisation, Expérimentations des Technologies de l'IST*, 2006, Disponible sur le Web : http://archivesic.ccsd.cnrs.fr/sic_00123423/document [Consulté le 08/06/2016]

SALAÜN Jean-Michel, « Bibliothèques numériques et Google-Print. Regards sur l'actualité », *La Documentation Française*, 2006, Disponible sur le Web : http://archivesic.ccsd.cnrs.fr/sic_00001576/document [Consulté le 08/06/2016]

WESTEEL Isabelle, « Le patrimoine passe au numérique », *Bulletin des bibliothèques de France [en ligne]*, n° 1, 2009, Disponible sur le Web : http://bbf.enssib.fr/consulter/bbf-2009-01-0028-003 [Consulté le 08/08/2016]

- Articles de revue :

BACHIMONT Bruno, « Enjeux et technologies : des données au sens », *Documentaliste-*

Sciences de l'Information, n°48, 2011 p1-14.

DINET Jérôme, « Pour une conception centrée-utilisateurs des bibliothèques numériques », *Communication et Langage*, n°161, 2009, p59-74.

JEANNENEY Jean-Noël, « Allégresse et vigilance », *Le Débat*, n°170, 2012, p192-201.

LUPOVICI Catherine, « De la bibliothèque classique à la bibliothèque numérique, continuité et rupture », *Documentaliste - Sciences de l'information*, n°37, 2000, p286-297.
MOATTI Alexandre, « Bibliothèque numérique européenne : de l'utopie aux réalités », *Annales des Mines - Réalités industrielles*, 2012, pp.43-46.

PAGANELLI Céline et CLAVIER Viviane, « Patrimoine et collections numériques : politiques, pratiques professionnelles, usages et dispositifs - Présentation du dossier 2015 », *Les Enjeux de l'Information et de la Communication*, n°16/2, 2015, p5-13.

PAPY Fabrice et LEBLOND Corinne, « Bibliothèque numérique : la nécessaire médiation », *Communication et Langage,* n°161, 2009, p37-58.

REBIMBEAU Gérard, « Du patrimoine aux collections numériques : pratiques, discours et objets de recherche », *Les Enjeux de l'Information et de la Communication*, n°16/2, 2015, p15-27.

- Ouvrages :

AMAR Muriel et MESGUISCH Véronique (dir), *Bibliothèques 2.0 à l'heure des médias sociaux*, Editions du Cercle de la Librairie, Paris, 2012, 224p.

BEQUET Gaëlle, *Trois bibliothèques européennes face à Google : aux origines de la bibliothèque numérique [1990 – 2010]*, Ecole des Chartes, Paris, 2014, 473p.

DOUEIHI Milad, *Pour un humanisme numérique*, Éditions du Seuil, Paris, 2011, 177 p.

DUFRÊNE Bernadette et IHADJADENE Madjid et DRUCKMANN Denis (dir), *Numérisation du patrimoine, Quelles médiations ? Quels accès ? Quelles cultures ?,* Herman Editeurs, Paris, 2013, 314p.

EBERLE-SINATRA Michaël et VITALI-ROSATI Marcello (dir), *Pratiques de l'édition numérique*, Parcours numérique, Montréal, 2014, 224p.

HENNO Jacques, *Silicon Valley / Prédateurs Vallée ? Comment Apple, Facebook, Google et les autres s'emparent de nos données*, Télémaque, Paris, 2011, 124p.

JACQUESSON Alain, *Google Livres et le futur des bibliothèques numériques*, Editions du Cercle de la Librairie, Paris, 2010, 224p.

JACQUESSON Alain*, L'informatisation des bibliothèques. Historique, stratégie et perspectives*, Editions du Cercle de la librairie, Paris, 1995, 362p.

JEANNENEY Jean-Noël, *Quand Google défie l'Europe : Plaidoyer pour un sursaut*, Mille et une nuits, Paris, 2007, 114p.

PAPY Fabrice (dir), *Les bibliothèques numériques*, Lavoisier, Paris, 2005, 226p.

PAPY Fabrice (dir), *Recherches ouvertes sur le numérique*, Lavoisier, Paris, 2013, 320p.

PERRI Pascal, *Google, un ami qui ne vous veut pas que du bien*, A. Carrière, Paris, 2013, 174p.

RACINE Bruno, *Google et le nouveau monde*, Plon, Paris, 2010, 149p.

- <u>Presse généraliste</u> :

DARNTON Robert, « La bibliothèque universelle, de Voltaire à Google », Le Monde Diplomatique, Mars 2009, p24-25.

NORA Dominique, « Le vrai pouvoir de Google en France», Le Nouvel Observateur, Juin 2016, p33-42.

- Rapports :

MOATTI Alexandre, TESNIERE Valérie, LESQUINS Noémie, « Bibliothèque numérique européenne remis le 31 janvier 2006 au Ministre de la Culture », Paris, 2006, 69p.

TESSIER Marc, « Rapport sur la numérisation du patrimoine écrit. Remis par Marc Tessier au ministre de la Culture et de la Communication le 12 janvier 2010 au Ministère de la Culture et de la Communication », Paris, 2010, 64p.

- Sitographie :

Article du juriste et bibliothécaire MAURIEL Lionel sur Google et la Bibliothèque municipale de Lyon : https://scinfolex.com/2012/12/15/numelyo-la-bibliotheque-numerique-de-lyon-exister-a-lombre-de-google/ [Consulté le 08/06/2016]

Article du juriste et bibliothécaire MAURICE Lionel sur les issues judiciaires de l'affaire Google Livres : https://scinfolex.com/2013/11/15/verdict-dans-laffaire-google-books-une-grande-lecon-de-democratie/ [Consulté le 08/06/2016]

Article sur le blog de la BnF concernant la mise en ligne du nouveau visualiseur : http://blog.bnf.fr/gallica/index.php/2015/01/26/sur-gallica-labs-un-nouveau-visualiseur-de-documents/ [Consulté le 08/06/2016]

Article du Journal du Net sur la publicité contextuelle de Google : http://www.journaldunet.com/ebusiness/crm-marketing/adwords-les-bonnes-feuilles/encheres-automatiques-et-encheres-manuelles.shtml [Consulté le 08/06/2016]

Article de Ina Global sur la publicité en ligne et Google : http://www.inaglobal.fr/communication-publicite/article/publicite-en-ligne-google-peut-il-cumuler-les-metiers-8236 [Consulté le 08/06/2016]

Page « présentation » du site Google Réseau Display : https://www.google.fr/ads/displaynetwork/ [Consulté le 08/06/2016]

Page « à propos » du site Google : http://www.google.com/about/company/ [Consulté le 08/06/2016]

Page « historique » du site Google Books : http://books.google.com/googlebooks/about/history.html [Consulté le 08/06/2016]

Site de l'Internet Archive : https://archive.org/index.php [Consulté le 08/06/2016]

Site d'Hispanica : http://www.bne.es/fr/Catalogos/BibliotecaDigitalHispanica/Inicio/ [Consulté le 08/06/2016]

Site de Belgia : http://belgica.kbr.be/fr/accueil_fr.html [Consulté le 08/06/2016]

Site Les Essentiels de Littérature : http://gallica.bnf.fr/essentiels/ [Consulté le 08/06/2016]

Site pédagogique de la BnF : http://classes.bnf.fr/index.php [Consulté le 08/06/2016]

Site Eduthèque : http://www.edutheque.fr/accueil.html [Consulté le 08/06/2016]

Site data.bnf de la BnF : http://data.bnf.fr/ [Consulté le 08/06/2016]

Annexes

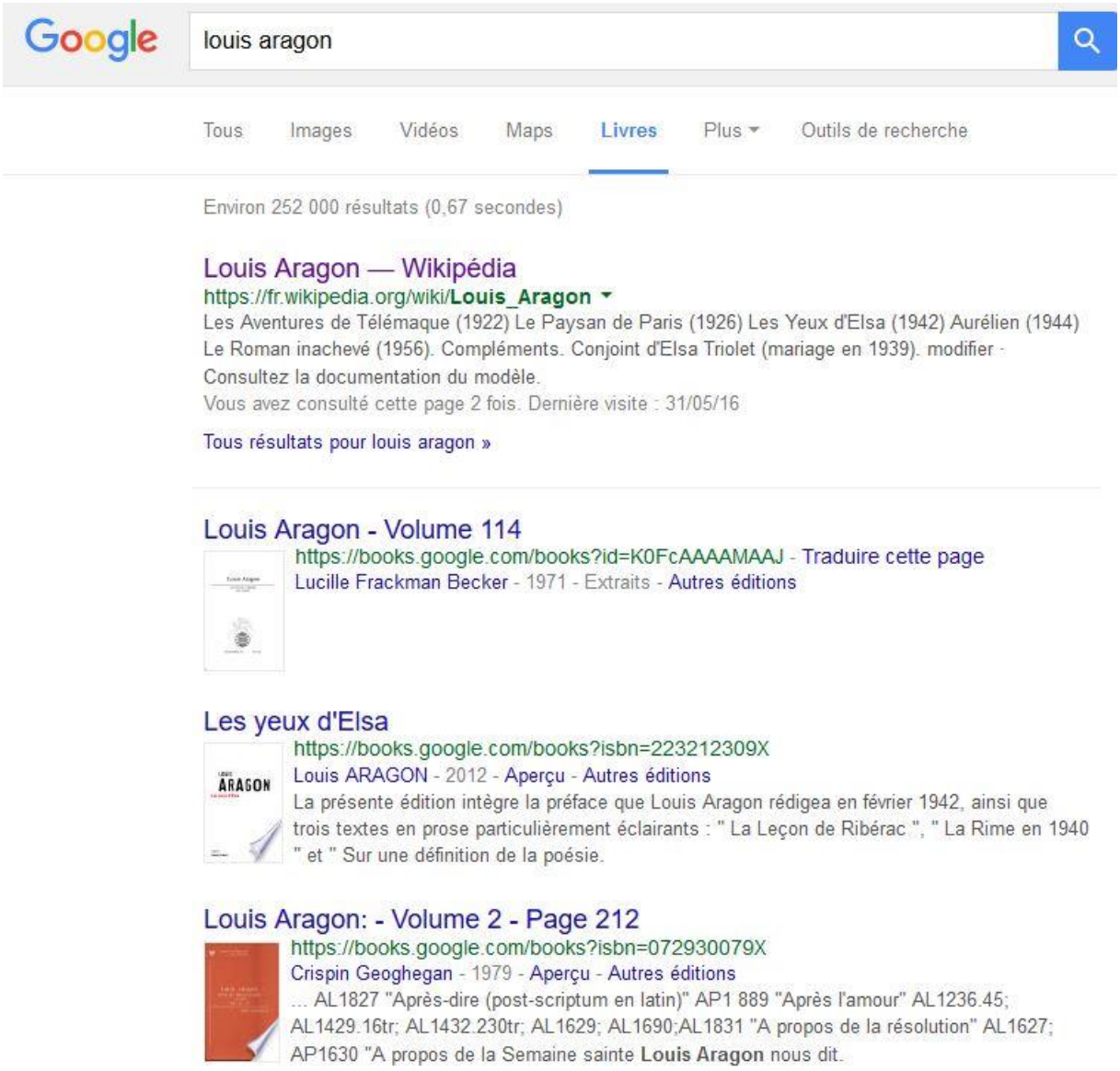
- Annexe 1 : Grille d'analyse du site Google Livres

Google Livres	
Page d'accueil	• Très épurée • 2 boutons : barre de recherche et ma bibliothèque Rappelle la page du moteur de recherche Google
Espace personnel	• L'utilisateur classe ses ouvrages par « étagères » • Certaines sont programmées par défaut et non supprimable • Renvoie vite vers Google Play Livres • Permet recherche plein texte dans l'ensemble des ouvrages de la bibliothèque. • Possibilité d'importer des livres.
Visualiseur	• Assez pauvre et peu esthétique pour Google Livres • Fonction « à propos de ce livre » avec informations intéressantes (fréquemment cités, nuages de mots…) • Les livres, même gratuits doivent « s'acheter » pour ensuite être répertorié dans « ma bibliothèque ». Ils se consultent ensuite via Google Play Livres. • Visualiseur Google Play Livres beaucoup plus riche et esthétique que celui de Google Livres : inter-connectivité avec le moteur de recherche et Google Traduction, notes, texte flottant pour personnaliser lecture. • Visualiseur Google Livres = feuiilletage, visualiseur Google Play Livres = lecture
La fonction recherche	• Très simple, fonctionne comme un moteur de recherche • Nombre de résultats très impressionnants • Ordre des résultats parfois surprenants • « Recherche avancée » peu visible, classique, puissante, efficace.
Conclusion	• Google Livres ressemble au moteur de recherche de Google (simplicité, esthétique). • Cherche à renvoyer sur librairie Google Play Livres. • Quelques fonctions très intéressantes (recherche plein texte sur ensemble bibliothèque, extraits d'ouvrages cités par d'autres ouvrages…).

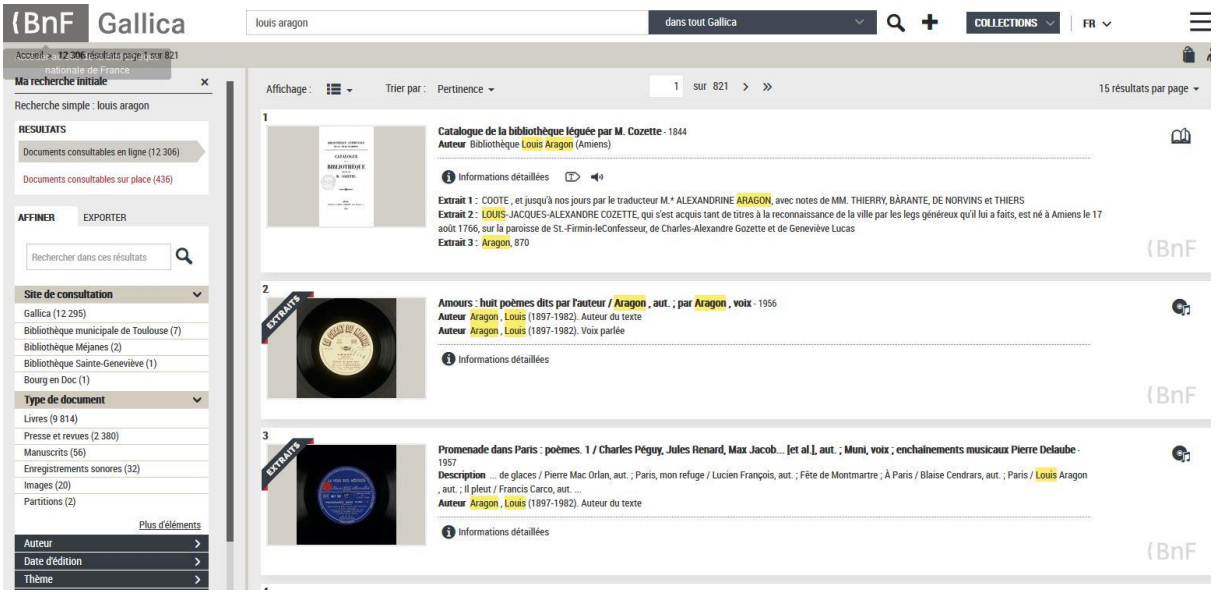
- Annexe 2 : Grille d'analyse du site Gallica

Gallica	
Page d'accueil	• Très fouillée, plusieurs catégories, boutons, bandeaux. • Esthétique très moderne (sliders, codes couleurs, site « en longueur »). • Version mobile du site disponible.
Espace personnel	• Assez similaire aux fonctionnalités de Google. • L'ajout de document et de marque page se fait en passant par le visualiseur. • Personnalisable intégralement.
Visualiseur	• Très accès sur la mise en valeur du document source. • Pas aussi adapté à la lecture que celui de Play Livres, angle différent : la consultation d'archives. • Efficace et complet, tout s'affiche dans la même page contrairement à Google Play Livres (ce qui est gênant pour les informations détaillées).
La fonction recherche	• Très visible, propose différents filtres et un bouton recherche avancée dès le haut de la page d'accueil. • Recherche avancée encore plus complète que celle de Google.
Editorialisation	• Documents classés en collections • Forte présence sur les réseaux sociaux • Prise en compte du web sémantique : data.bnf
Conclusion	• Médiation numérique et éditorialisation omniprésente • Site fouillé et moderne (dur d'accès ?) • Document particulièrement mis en avant avec le visualiseur

- Annexe 3 : Captures d'écran

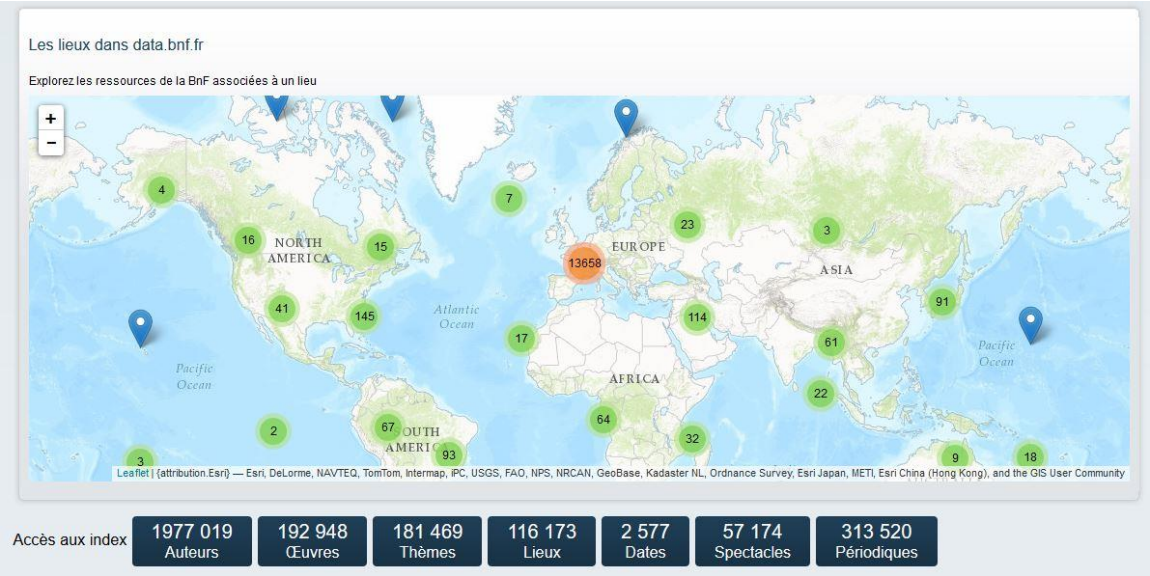


Recherche « louis Aragon » dans Google Livres



Recherche « Louis Aragon » sur Gallica.

Recherche du mot « terre » au sein de ma bibliothèque Google Livres



Site data.bnf

Page Twitter de Gallica.

- Annexe 4 : Déclaration anti-plagiat

UFR LLASIC
UNIVERSITÉ Grenoble Alpes

Déclaration anti-plagiat
Document **à scanner** après signature
et **à** intégrer au mémoire électronique

DECLARATION

1. Ce travail est le fruit d'un travail personnel et constitue un document original.
2. Je sais que prétendre être l'auteur d'un travail écrit par une autre personne est une pratique sévèrement sanctionnée par la loi.
3. Personne d'autre que moi n'a le droit de faire valoir ce travail, en totalité ou en partie, comme le sien.
4. Les propos repris mot à mot à d'autres auteurs figurent entre guillemets (citations).
5. Les écrits sur lesquels je m'appuie dans ce mémoire sont systématiquement référencés selon un système de renvoi bibliographique clair et précis.

NOM : BROSSET PRENOM : Laurent

DATE : 08/06/2016 SIGNATURE : L Brosset