



HAL
open science

Éditorialisation des bibliothèques numériques : le cas des Essentiels de Gallica

Arnaud Laborderie

► **To cite this version:**

Arnaud Laborderie. 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 ”, Université Paris 8 (Laboratoire Paragraphe), Université Paul-Valéry (Laboratoire LERASS-Céric), Université Lille 3 (Laboratoire GERiiCO), Nov 2015, Montpellier, France. hal-01239425

HAL Id: hal-01239425

<https://bnf.hal.science/hal-01239425>

Submitted on 7 Dec 2015

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.

Éditorialisation des bibliothèques numériques : le cas des Essentiels de Gallica

Editorialization Digital Libraries: the case of Gallica Essentials

Arnaud LABORDERIE

Bibliothèque nationale de France, Chaire Unesco ITEN, Université Paris-VIII,
Laboratoire Paragraphe
arnaud.laborderie@bnf.fr

Résumé. L'éditorialisation désigne les pratiques de publication et d'accessibilité des contenus sur le web, lesquelles posent des questions épistémologiques sur l'authenticité et la véracité de l'information. Au-delà de ses techniques et de ses formes, l'éditorialisation interroge la fonction éditoriale et auctoriale. Sélectionner, structurer, hiérarchiser, documenter, donner du sens : dans ce processus, les bibliothèques ont une responsabilité au même titre que les éditeurs. Au-delà de la mise à disposition de leurs ressources numériques, elles sont engagées dans la production de contenus sur le web. Première d'entre elles, la Bibliothèque nationale de France (BnF) est doublement impliquée : d'une part avec des ressources accessibles et interopérables grâce à data.bnf.fr et des corpus structurés dans Gallica, d'autre part avec une médiation spécifique à destination des publics scolaires et des enseignants, « Les Essentiels de la littérature ». Il s'agit d'une interface éditorialisée à la bibliothèque numérique qui tient compte des pratiques anthologiques actuelles pour faire découvrir notre patrimoine littéraire à travers des parcours guidés dans les collections.

Mots-clés. Éditorialisation, bibliothèque numérique, médiation numérique, curation de contenu.

Abstract. The editorialization means publishing practices and accessibility of content on the web, posing epistemological questions about the authenticity and veracity of information. Beyond the techniques and forms, the editorialization questions the editorial and authorial functions. Select, organize, prioritize, document, make sense: in this process, libraries have a responsibility as well as publishers. Beyond the availability of their digital resources, they are engaged in the production of content on the web. First of them, the French National Library (BnF) has been doubly involved: on one hand with accessible and interoperable resources through data.bnf.fr and structured corpus Gallica, on the other hand with specific mediation for school audience and teachers, "The Essentials of Literature". This is an editorialized interface to the digital library that reflects current anthological practices to discover our literary heritage through guided paths into the collections.

Keywords. Editorialization, digital library, digital mediation, content curation.

1 Introduction

Le numérique bouleverse profondément notre rapport au savoir avec un document qui a changé de nature. La numérisation a transformé le document en pixels et octets. Il n'est plus qu'une longue chaîne de 0 et de 1. Dématérialisé, celui-ci se prête à toutes les manipulations. La preuve qu'il pouvait apporter est désormais relative, selon les contextes d'interprétation. Dans l'environnement numérique, son authenticité peut être remise en cause alors que disparaît la notion même d'*original* derrière des *copies* potentiellement infinies (Pédaque, 2006). Ces questions épistémologiques constituent une véritable rupture : le document n'est plus à considérer en termes de structure ou d'objet mais de flux. Hétérogène, il est de tout type : tout peut devenir document dès lors qu'il porte une information. Selon Pédaque, « l'affaire du document n'est ni sa matière, ni sa forme, mais son usage. » (*Ibid.*).

Dans le même temps, le document numérique connaît une croissance exponentielle. Sa prolifération n'en menace-t-elle pas la survie ? N'a-t-on pas lieu de craindre, dans un monde post-numérique, alors que tout est potentiellement document, que rien ne le soit plus ? Déjà le document semble voué au fragment, sous le double coup du processus de discrétisation et des usages anthologiques. Le fragment n'est-il pas la seule trace à conserver dans des pratiques qui visent à faire disparaître les documents derrière des données qu'il s'agit de visualiser dans des représentations de masse interprétées par des machines ? Une trace à sauvegarder comme vestige de l'œuvre ? Ce qu'on peut craindre en effet dans un monde post-numérique, c'est la perte de l'original au profit des représentations.

L'usage crée une nouvelle « épistémè numérique » (Stiegler, 2012) et requiert de nouvelles compétences, une « digital literacy » (Doueïhi, 2008) : trier, hiérarchiser, fiabiliser, sourcer, corréler, etc. Il s'agit de former les esprits à trouver les bonnes informations, les bons documents, et savoir les assembler pour qu'ils fassent sens. Le lecteur est en effet constamment invité à construire du sens à partir des matériaux multiples et hétérogènes (Grafton, 2012).

Les conditions de conservation et de consultation ont également changé : conservé dans quelque *datacenter*, avec ses contraintes, ses fragilités et ses coûts, le document électronique n'est pas sans matérialité alors que le web donne l'illusion d'un document dématérialisé accessible en toute circonstance. En réalité, nombre de documents numérisés ne sont présents sur le réseau que derrière des systèmes de protection, des abonnements, sur des intranets, etc. Ces restrictions rendent difficiles d'accès certains documents des plus qualifiés, favorisant ceux disponibles sur internet sans filtre ni garantie. Leur fiabilité dépend en effet de l'instance de publication : la caution que l'éditeur apporte à la publication du document se délite sur le web où chacun peut publier des documents qui s'hybrident par nature. Des documents hétérogènes, multimédias, constitués de textes, d'images, d'audiovisuels, voire d'animations. Pluriels et dématérialisés, ils ne sont plus consultables qu'à travers un dispositif informationnel qui les rematérialise dans une interface qui doit restituer l'intégrité du document et permettre sa consultation, son exploitation par une communauté de lecteurs, son interconnexion avec d'autres documents. Tout l'enjeu de l'interface, c'est de restituer les trois dimensions du document selon Pédaque : forme, texte, médium.

Le web comme interface de médiation

L'usage n'impose-t-il pas le web comme interface de lecture et de transmission des connaissances en lieu et place du livre ? Le web se présente en effet comme une interface de médiation, de représentation du monde et d'organisation des savoirs. Alors

même que les éditeurs sont menacés par la remise en cause de leur modèle économique, il importe d'appliquer au web les principes du livre : il s'agit véritablement « d'éditer le web » (Bon, 2011). Quelle confiance peut-on conférer au dispositif informationnel considéré, logiciel de lecture, site web ou plateforme collaborative ?

Face à la « crise des médiateurs » sur le web, l'institution, conservatoire des collections patrimoniales, a un rôle déterminant : par la constitution de bibliothèques numériques et de bases de données, elle est garante de confiance et d'autorité. Investie d'une mission de service public, l'institution est une référence et s'impose comme un contrepouvoir face à des entreprises privées de numérisation des documents et de privatisation des usages (Jeanneney, 2007).

Pour autant, malgré l'optimisation des conditions d'accès aux documents, la croissance exponentielle des ressources numériques pose le problème d'interpréter la liste de résultats : difficulté à trier, hiérarchiser, identifier les documents pertinents pour un usage donné. Face à l'expansion des données, l'éditorialisation peut-elle guider et donner du sens ?

2 Pratiques et enjeux de l'éditorialisation

Le terme « éditorialisation » recouvre les pratiques de publication et d'accessibilité des contenus sur le web. Il a d'abord été employé par Bruno Bachimont pour désigner « le processus consistant à enrôler des ressources pour les intégrer dans une nouvelle publication. Sortant de la logique purement documentaire, l'éditorialisation est une exploitation des contenus se fondant sur la recherche d'information mais ne s'y limitant pas. L'éditorialisation est la conclusion logique du processus de numérisation des contenus. » (Bachimont, 2007). L'auteur insiste sur des pratiques qui consistent à discrétiser les ressources numériques en unités documentaires, en segments pouvant être utilisés dans n'importe quel contexte, sans lien avec la source d'origine et dans un autre sens. Ceci pose des problèmes épistémologiques quant à l'authenticité et la véracité des documents numériques ainsi produits. Plus largement, l'éditorialisation « met l'accent sur les dispositifs technologiques qui déterminent le contexte d'un contenu et son accessibilité. Éditer un contenu ne signifie pas seulement le choisir, le légitimer, le diffuser, mais lui donner son sens propre en l'insérant dans un contexte technique précis, en le reliant à d'autre contenu, en le rendant visible grâce à son indexation, à son référencement, etc. » (Vitalo-Rossati et Sinatra, 2014). Toutes les actions destinées à rendre accessible et visible un contenu sur le web relèvent donc de l'éditorialisation. « Même la recommandation d'un document dans un réseau social est une pratique d'éditorialisation. » (*Ibid.*)

Sur le plan technique, c'est d'abord un enjeu de structuration des contenus, à la fois par hiérarchisation, rubricage et maillage. L'indexation et les métadonnées assurent l'accessibilité et l'interopérabilité. Des solutions logicielles de type CMS (*Content Management System*) permettent de formater les contenus en prescrivant des pratiques éditoriales. Les CMS s'imposent comme des outils incontournables qui facilitent et encadrent le processus de production — c'est un avantage certain, non sans inconvénient car ces mêmes CMS formatent les contenus qui tendent vers toujours plus de standardisation. Il importe alors de savoir jouer des contraintes et d'être créatif. Éditorialiser, c'est aussi personnaliser sa plateforme de publication en termes de design et de mise en page, c'est encore sélectionner les ressources à mettre en avant. Le lecteur lui-même est invité dans le processus d'éditorialisation par la contribution, le partage ou la recommandation. Degré zéro de l'éditorialisation, le simple geste de *liker* a des conséquences parfois décisives en termes de visibilité des contenus. Les réseaux sociaux, parce qu'ils cherchent à garder le lecteur dans leurs écosystèmes, génèrent des

formes spécifiques d'éditorialisation, algorithmiques, par extraction et génération de contenus. Aussi faut-il distinguer l'éditorialisation qui procède de l'algorithme de celle procédant du lecteur.

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 procèdent du lecteur qui produit ainsi son propre contenu, personnalisé. Face à la surabondance des contenus sur le web, de telles pratiques de curation peuvent guider par une fonction éditoriale subjectivée.

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.

Ainsi, sur le web, l'édition se transforme en éditorialisation : cette pratique d'organisation et de structuration de l'information participe à la production et à la circulation du savoir. Individuelle ou collective, elle s'affirme comme une pratique auctoriale autant qu'éditoriale.

L'éditorialisation va-t-elle se substituer à l'édition ? Avec la révolution numérique, l'édition est en mutation à la fois dans son processus de production (la chaîne du livre) et dans son modèle économique (Benhamou, 2014). La facilité de publication sur le web et les pratiques d'autoédition questionne un modèle éprouvé depuis des siècles. Faut-il craindre une « édition sans éditeurs » (Schiffrin, 1999) quand des acteurs comme Amazon cherchent à imposer leur modèle ? Paradoxalement, on n'a jamais tant eu besoin des éditeurs pour guider le lecteur dans la surabondance des contenus. Sans doute ne faut-il pas craindre la mort de l'éditeur pas plus que la mort du livre. Demeure par nécessité le maintien d'une fonction. Ce qui risque de mourir, c'est la caution de l'éditeur, garant d'une ligne éditoriale, de la qualité des œuvres et de leur diffusion.

Face à ces bouleversements, une institution telle que la BnF s'impose comme garant de qualité, de véracité et d'authenticité : c'est pourquoi elle assure une fonction éditoriale et investit pleinement l'éditorialisation, en tenant compte des usages émergents.

Actualiser des pratiques ancestrales

Ces pratiques d'éditorialisation qui nous paraissent si actuelles n'en sont pas nouvelles pour autant : elles puisent à une longue tradition d'édition des textes et de perfectionnement des usages multipliant les modalités d'accès aux œuvres. Prenons l'exemple des concordances, déjà présentes dans la tradition livresque juive ainsi que dans la bibliothèque d'Alexandrie. Elles peuvent nous apparaître comme des formes d'éditorialisation qui seront reprises par les bibliothèques chrétiennes dès le II^e siècle de notre ère. En établissant des correspondances entre les Évangiles avec un principe d'indexation et en mettant en place tout un dispositif de production des textes, Eusèbe de Césarée n'a-t-il pas, au tournant des III^e et IV^e siècles, en quelque sorte « éditorialisé » le Nouveau Testament ? Chaque Évangile fut en effet découpé en

segments, classés dans dix tables auxquelles le lecteur est invité à se reporter, pour découvrir les passages semblables, par un numéro indiqué dans la marge (Grafton, 2012). Ces correspondances préfigurent les hyperliens. Plus tard, en accueillant la glose dans l'intérieur même du texte, les copistes médiévaux créaient des paratextes offrant, dans des mises en pages sophistiquées, des parcours de lecture et d'interprétation véritablement « hypertextuels ».

Mise au point au XVI^e siècle, la roue à livres permettait une lecture comparée de plusieurs livres simultanément. N'est-ce pas déjà du multitâche à fenêtres multiples ? Avec un savoir médiéval circonscrit qui brusquement s'ouvrait aux savoirs antiques, lesquels semblaient infinis, les Humanistes se sont confrontés, dans un contexte différent, à une situation identique à la nôtre : de la rareté à l'abondance, l'arbre du savoir médiéval plongeait dans l'océan des savoirs antiques. Constitués sur plus d'un millénaire, les savoirs de l'Antiquité excédaient tout ce que les scolastiques avaient accumulé en quelques siècles. Des tels savoirs soudain révélés par l'Orient dépassaient l'entendement. L'homme de la Renaissance s'est employé à mettre en place des méthodes et des formes d'accès, d'apprentissage, de transmission, qui permettaient d'embrasser ces savoirs immenses : en développant toute une épistémologie du fragment à travers les florilèges, centons et lieux communs. Constitués sur une trentaine d'années, les quelques quatre mille adages d'Érasme embrassent ainsi la totalité de la littérature antique par une sélection de citations, structurées et commentées. Érasme use de la métaphore de l'abeille « volant autour de toutes les espèces de fleurs, d'herbes, d'arbrisseaux » transformées en elle par un processus digestif : « tu n'y reconnaîtrais la saveur ou l'odeur ni de la fleur ni de l'arbrisseau butiné, mais la production de la petite abeille composée de tous ces éléments. » (*Ciceronianus*, 1528). Un tel « butinage », déjà en usage sous l'Antiquité, est théorisé par les Humanistes et pratiqué comme une méthode, un moyen de transmission des connaissances, et même un modèle rhétorique (Moss, 2002). Cette manière devait trouver son couronnement dans les *Essais* de Montaigne. Toute une méthode de lecture, d'apprentissage et de transmission se fonde sur un principe de résumés, de citations, d'anthologies et d'appropriation individuelle dans des carnets de lieux communs. Des lieux communs, antérieurs à toutes considérations de propriété intellectuelle, qui résonnent singulièrement avec la notion émergente de « bien commun numérique ».

N'est-ce pas le propre du numérique de revisiter des pratiques ancestrales, de les hybrider, dans une tension entre ruptures et continuités ? Ces pratiques, prises aujourd'hui dans un autre contexte, avec une autre dimension, posent les mêmes problématiques et les mêmes enjeux de sélection, d'indexation, d'édition et de médiation. À tel point qu'on parle de notre époque comme d'une Renaissance et d'un humanisme numérique (Doueïhi, 2011).

L'apport des bibliothèques : indexation, métadonnées, pratiques documentaires

Le processus d'éditorialisation recourt à des pratiques bibliothéconomiques éprouvées, à la fois dans l'indexation, la constitution des catalogues et la description des ressources.

Le catalogage et l'indexation permettent d'identifier et de localiser les ressources dans les collections des bibliothèques, de tout type et support, voire d'y accéder directement dans le cas de ressources numérisées. Des notices bibliographiques contiennent des informations précisant, pour un livre par exemple, le titre, l'auteur, l'éditeur, la date, etc. Ces métadonnées sont structurées dans des formats standardisés, tel UNIMARC, le format officiel d'échange de l'information bibliographique en France. C'est à l'aide d'un format minimal de quinze éléments, le Dublin Core, que sont décrites

les entités documentaires sur le web. Ces métadonnées permettent de passer du web des documents au web de données, ou « web sémantique » dont l'enjeu consiste à « pouvoir déléguer à la machine une partie de l'interprétation des ressources du web ». (Bachimont, 2011)

Un modèle de données permet d'exprimer informatiquement des données hétérogènes, le RDF (Resource Description Framework), qui repose sur une structure logique de trois éléments, dite « triplet » : sujet, prédicat, objet. Ces triplets permettent de lier les informations entre elles et de les interroger.

L'enrichissement des métadonnées participe du processus d'éditorialisation en déterminant la visibilité des contenus. La qualité des métadonnées s'avère en effet stratégique dans un environnement informationnel à la croissance exponentielle.

À ces pratiques bibliothéconomiques s'ajoutent des pratiques documentaires de médiation : la gestion de l'information ne se contente plus d'une mise à disposition des ressources mais évolue vers une production propre qui tend à structurer les contenus en ligne. Les bibliothécaires pratiquent depuis longtemps des formes d'éditorialisation de contenus sous la forme de bibliographies, de « coups de cœur » ou de listes d'acquisitions, axées autour de la valorisation de leurs collections. Avec l'explosion des ressources en ligne, ils développent leur offre de contenu et investissent un nouveau rôle de « curateur » de l'information (Di Pietro, 2014). Ainsi le processus d'éditorialisation permet-il d'impliquer les métiers des bibliothèques et de les repenser, en interprétant le bibliothécaire comme « journaliste de ses collections » (Dujol, 2011).

3 L'éditorialisation à la BnF

La numérisation de masse des collections patrimoniales a révélé de nouveaux enjeux : alors que toute information, tout document, sont immédiatement disponibles sur le web, l'accroissement exponentiel des bases de données rend, paradoxalement, cette information, ce document, plus difficiles à trouver.

Plus de 4,2 millions de documents sont aujourd'hui accessibles sur Gallica : interroger « Jean-Jacques Rousseau », c'est obtenir plus de 13 000 résultats, pas moins de 6 200 résultats pour « Les Confessions ». Seul le chercheur, l'expert, parce qu'il manipule les outils de recherche fédérée, parce qu'il sait user des filtres qui lui permettent de trier les résultats, parce qu'il concatène des mots-clés pertinents, parvient à trouver le texte, l'édition, le manuscrit, l'image qui peuvent l'intéresser. Le grand public, lui, ne peut qu'être dérouteré par une telle abondance de la ressource et, dans bien des cas, se détourne de ces bases de données, pourtant au cœur même de la connaissance et du processus de construction des savoirs. Les modalités d'accès à ces ressources et l'éditorialisation deviennent alors déterminantes : c'est en effet là, dans la bibliothèque numérique, que se trouvent les sources primaires et les autorités qui fiabilisent l'information. La sémantisation des ressources, les métadonnées affectées à chaque document, les architectures catégorielles et ontologiques, les index, les listes etc. offrent de multiples manières de circuler dans ces bases, au-delà du champ « rechercher ». Mais comment naviguer au-delà de la liste de résultats et du filtrage à facettes ?

Une nouvelle interface pour Gallica

Mise en ligne en 1997, Gallica, la bibliothèque numérique de la BnF, n'a cessé de se transformer pour prendre en compte la richesse et la diversité des contenus disponibles et pour répondre aux évolutions des usages du web. Le moteur de recherche Cloudview, développé en 2013 par la société Exalead, améliore la pertinence

des résultats et intègre des fonctionnalités d'aide à la recherche. Une nouvelle interface de Gallica¹ est ouverte au public depuis le 1^{er} octobre 2015 : nouveau graphisme, nouveau visualiseur et nouvelles fonctionnalités, comme l'intégration de la recherche avancée et la mise en place du téléchargement de documents aux formats jpeg et pdf. Un visualiseur unique permet de choisir différents modes d'affichage (simple page, double page, défilement vertical, vue d'ensemble mosaïque) et de manipuler plus facilement les documents. Des améliorations ont aussi été apportées aux listes de résultats de l'interface de recherche².

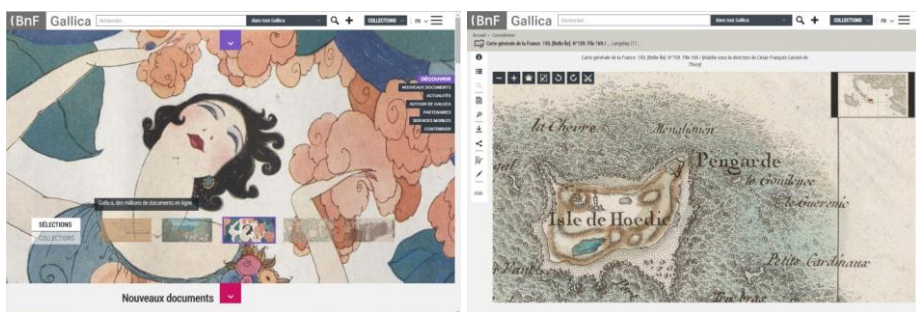


Fig. 1. Nouvelle interface de Gallica : page d'accueil et visualiseur

La bibliothèque numérique constitue le troisième espace que la BnF offre au public, avec sa bibliothèque d'étude et sa bibliothèque de recherche. Comment circuler dans cet espace virtuel ? Quels dispositifs de consultation pour ses lecteurs ? Quels chemins d'accès proposer à un plus large public au sein de ses collections ?

Data.bnf.fr : rendre accessibles les données de la BnF

Face à la difficulté de trier et hiérarchiser les résultats, la BnF a mis en place dès 2011 data.bnf.fr³ : ce sont des pages de référence sur les auteurs, les œuvres et les thèmes qui regroupent les informations de Gallica et celles issues des différents catalogues bibliographiques. Les usagers peuvent accéder à ces données qui décrivent les ressources sans passer par des catalogues ou des portails spécifiques. Data.bnf.fr couvre actuellement 63% des notices bibliographiques du Catalogue général, pour 890 000 auteurs et quelque 8 millions de documents liés. Les données des catalogues de la BnF font partie de ce qu'on appelle le « web profond ». Une telle éditorialisation accroît leur visibilité en les rendant indexables par les moteurs de recherche. L'objectif est de valoriser sur le web la richesse des fonds et de servir de pivot entre les différentes ressources. Le projet s'inscrit dans une démarche d'ouverture de la BnF au web de données et d'adoption des standards du web sémantique.

Explorer les collections

L'éditorialisation de Gallica s'inscrit dans une démarche de médiation et de valorisation. Il s'agit d'orienter des publics potentiels en organisant les collections numérisées dans des parcours structurés par thèmes ou par disciplines. L'objectif est de

¹ Nouvelle interface de Gallica : <http://gallica.bnf.fr/>

² Pour en savoir plus sur les nouvelles fonctionnalités de Gallica :

<http://blog.bnf.fr/gallica/index.php/2015/01/26/sur-gallica-labs-un-nouveau-visualiseur-de-documents/>

³ Informations reprises de la présentation en ligne. Pour en savoir plus : <http://data.bnf.fr/about>

faciliter la dissémination des ressources numériques et d'encourager la démocratisation de leurs usages. La difficulté réside dans l'impossibilité d'automatiser le requêtage, qui doit être fait manuellement et suppose un travail rétrospectif important car l'indexation n'est pas faite en fonction de l'éditorialisation. L'enjeu est d'éliminer le bruit par des requêtes complexes qui s'appuient sur le catalogue quand la cotation n'est pas suffisamment précise, ce qui est souvent le cas des collections antérieures à 1975. Une solution consiste à insérer dans les notices des codes projets que l'on peut ensuite requêter.

Les corpus ainsi éditorialisés concernent tous types de collections des différents départements de la BnF, par exemple : sciences naturelles, livrets xylographiques, archives de la Bastille, manuscrits carolingiens, cartes marines, histoire de France par l'image, partitions musicales par genre, etc. Le plus riche est le portail France-Japon ⁴, issu d'une coopération entre la Bibliothèque nationale de France et la Bibliothèque nationale de la Diète (Tokyo). Il présente plus de 2 000 œuvres concernant les relations franco-japonaises et la connaissance du Japon en France des origines à 1914. Le parcours thématique est centré sur les collections imprimées des années 1850-1914. Une exploration des collections précieuses liées au Japon vient le compléter : fonds anciens japonais, estampes japonisantes de la fin du XIX^e siècle, cartes anciennes, premières photographies et images des organes de presse. Ces ressources numériques sont articulées avec les modules réalisés dans le cadre des expositions virtuelles.

Éditorialisation sur les réseaux sociaux

La présence de Gallica sur les réseaux sociaux assure la promotion des contenus numériques et des services associés, en tenant compte de l'évolution des pratiques du web et en expérimentant de nouvelles formes d'interactivité avec les utilisateurs.

Chaque jour, la page Facebook de Gallica ⁵ met en valeur un document ou un ensemble documentaire numérisés pour leur dimension patrimoniale ou historique, ou leur caractère insolite et inattendu. Elle permet ainsi de faire connaître la richesse et la qualité des collections numériques de la BnF. Les réactions et les commentaires des fans sont nombreux témoignant d'un véritable intérêt pour l'offre documentaire de Gallica et pour ce mode de valorisation « sociale » qui permet de partager les documents. Le fil Twitter de Gallica ⁶ relaye l'information et signale, de manière plus originale, les « trouvailles » des internautes, leur réappropriation et leur propre mise en valeur des contenus. Il valorise ainsi non seulement Gallica, mais aussi ses utilisateurs et la manière dont ceux-ci mentionnent, citent, commentent les documents numérisés par la BnF. D'autres médias sont encore utilisés pour la valorisation de Gallica : blog.bnf.fr, Lettre de Gallica, Pinterest. Dernièrement, l'équipe Gallica était invitée par le *Huffington Post* ⁷ à commenter l'actualité à partir des collections numériques.

⁴ Corpus France-Japon en ligne : <http://gallica.bnf.fr/html/und/livres/france-japon>. Rééditorialisation dans le cadre des expositions virtuelles : <http://expositions.bnf.fr/france-japon/>

⁵ Créée en février 2010, la page Facebook de Gallica compte 87 027 « J'aime » au 4 juillet 2015 : <http://www.facebook.com/GallicaBnF>

⁶ Lancé en août 2010. Le fil Twitter de Gallica compte 1639 abonnés au 4 juillet 2015. <http://twitter.com/GallicaBnF>

⁷ En savoir plus : <http://blog.bnf.fr/gallica/index.php/2015/09/21/quand-le-huffington-post-invite-gallica/>

4 Les « Essentiels de la littérature » dans Gallica

Peut-on aider à mieux voir, mieux lire, mieux connaître, en proposant une interface pour la bibliothèque numérique qui privilégie la découverte et favorise l'appropriation ? L'objectif des « Essentiels de la littérature » est d'amener vers Gallica des publics scolaires et des enseignants, en construisant des chemins d'accès, des parcours de connaissances, qui modélisent les usages et organisent les ressources – des ressources choisies, structurées, des contenus pertinents, qualifiés, travaillés. Offrir un accès simple et éditorialisé aux « essentiels » de notre patrimoine littéraire et faire découvrir les auteurs et les œuvres littéraires dans leur contexte historique et artistique. Le site comprend des repères sur les grands mouvements littéraires et artistiques, des modules présentant de manière attractive auteurs, œuvres et thèmes. Il importe de mettre en relation les imprimés avec les manuscrits, les illustrations, des corpus d'images, des enregistrements sonores et vidéos, des conférences en ligne... Le site permet enfin de tisser des liens intellectuels entre les œuvres et de créer des rebonds par des regards contemporains et des approches thématiques.

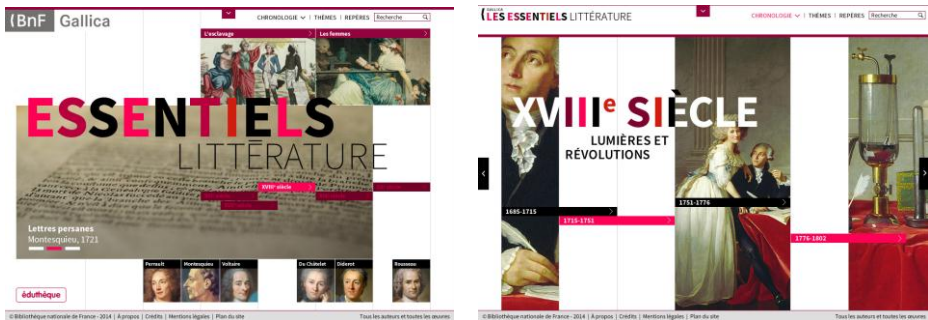


Fig. 2. Interface des Essentiels de Gallica : page d'accueil et entrée XVIII^e siècle

À terme, les « Essentiels » proposeront quelque 500 entrées dans l'histoire de la littérature. C'est un projet sur trois ans avec une déclinaison par siècle. Une première version du site, couvrant le XVIII^e siècle, a ouvert en octobre 2015 en même temps que la nouvelle interface de Gallica⁸. À côté de Voltaire, Diderot, Rousseau ou Montesquieu, le site fait place à des femmes comme Émilie du Châtelet, Manon Roland ou Olympe de Gouges et à des écrivains moins connus comme Louis-Sébastien Mercier, auteur du *Tableau de Paris*, chronique savoureuse de la vie parisienne à la veille de la Révolution.

C'est une triple approche anthologique qui a été privilégiée : sélection d'auteurs et d'œuvres accessibles selon trois niveaux ; sélection de ressources numériques commentées et visualisables au zoom ; sélection d'extraits qui renvoient vers le texte intégral. Des requêtes prédéfinies ouvrent vers des corpus plus larges de textes et d'images.

Le site s'articule autour de quatre entrées – périodes, auteurs, œuvres ou thèmes – qui sont mises en avant dès la page d'accueil. Au centre, une sélection d'œuvres s'affiche dans un *slider* qui fait le lien vers les modules concernés. Deux thèmes sont proposés ainsi qu'un choix d'auteurs. On accède à la barre de navigation générale de Gallica par une « tirette » dans le bandeau du haut.

⁸ Les Essentiels de la littérature : <http://gallica.bnf.fr/essentiels>

Des tableaux chronologiques

Auteurs et œuvres prennent place dans des tableaux chronologiques proposant des repères historiques sur les événements, les découvertes scientifiques, les temps forts artistiques ou le mouvement des idées dans lesquels s'inscrit l'histoire littéraire. Il importe ici de contextualiser la littérature par période. Le XVIII^e siècle est ainsi découpé en quatre grands tableaux historiques et culturels :

- 1685-1715 : Anciens contre Modernes
- 1715-1751 : L'âge des expériences
- 1751-1778 : Le triomphe des Lumières
- 1778-1800 : L'ère des révolutions



Fig. 3. *Les Essentiels* : tableaux chronologiques

Ces tableaux chronologiques se présentent comme des « cimaises virtuelles » avec un accrochage hiérarchisé des œuvres : au centre, un choix d'œuvres « majeures » accessibles par une grande image, de petites vignettes interactives signalent d'autres œuvres importantes de la période, tandis qu'un bouton « plus d'œuvres » affiche des titres complémentaires. Il ne s'agit pas ici d'un jugement de valeur, mais d'une modalité d'affichage qui tient compte du niveau d'éditorialisation des contenus proposés. Il importe de rester dans un volume de 20 à 30 œuvres par tableau. Le nom des auteurs phares renvoie vers leurs modules respectifs. Le contexte est documenté par des repères historiques, artistiques et scientifiques accessibles sous forme de fiches qui renvoient vers des requêtes dans Gallica.

Modules auteurs et œuvres

Les entrées auteurs et œuvres se déclinent en modules plus ou moins élaborés, afin d'offrir une logique de découverte par niveaux. Les modules principaux



Fig. 4. *Les Essentiels* : module auteur

s'organisent en quatre rubriques qui structurent une approche multimédia : découvrir, rencontrer, explorer, approfondir. Plus ou moins complexes selon le statut de l'auteur ou de l'œuvre mais aussi des ressources et contenus disponibles, les modules proposent une approche allant du plus synthétique et visuel au plus textuel et approfondi.

- « Découvrir » introduit l'auteur ou l'œuvre en images pour en donner, sur le modèle de l'exposition virtuelle, les clés de compréhension.
- « Rencontrer » propose un entretien audiovisuel de quelques minutes avec un spécialiste qui fait partager sa passion et donne l'envie de lire.
- « Explorer » présente des collections d'estampes et de manuscrits organisées en albums thématiques selon la logique d'un propos qui se décline d'image en image.
- « Approfondir » rassemble des pages de dossier : À propos de l'auteur ou l'œuvre, articles plus détaillés et, le cas échéant, chronologie, bibliographie, repères, sites internet, etc. Une sélection d'extraits, soigneusement choisis, invite à découvrir le texte intégral dans une édition sélectionnée pour sa lisibilité.

Les modules « secondaires », moins éditorialisés, reprennent quelques-uns de ces éléments qui sont tous modulaires et peuvent s'enrichir progressivement. Depuis l'auteur, on accède aux œuvres sélectionnées. L'accès à data.bnf.fr permet d'ouvrir vers les ressources du catalogue. La page peut être partagée.

L'approche thématique

L'esclavage, l'image de l'autre, l'Orient, la place des femmes... : des thèmes transverses tissent des liens entre les œuvres dans un siècle donné ou à travers les époques. Ils permettent aussi d'éclairer le passé par des regards contemporains. Sont proposés quatre types d'approches thématiques, de la plus éditorialisée à la moins éditorialisée : parcours pédagogiques, genres littéraires, groupements documentaires, entretiens audiovisuels. Chaque parcours propose une piste pédagogique qui s'appuie sur une anthologie et un album d'images.

Des outils de recherche permettent de trouver aisément tout type de ressources dans le site. Réalisés en partenariat avec l'Éducation nationale, les Essentiels prendront place dans le portail « Éduthèque »⁹ grâce auquel les enseignants peuvent télécharger les images et les vidéos.

5 Conclusion

À l'image du web qui est une culture de l'hybridation, l'éditorialisation recourt à plusieurs modèles et requiert des compétences issues de différents métiers. De l'expert en sciences de l'information, on retient les techniques d'architecture de l'information, d'indexation et de gestion de contenu. Du savoir-faire de l'éditeur importent le choix des contenus, la légitimation de ces contenus et leur diffusion. Bien connaître son « lectorat », construire des narrations (*storytelling*) et mettre en place un processus de publication, le *workflow* éditorial. Du journaliste prévalent un principe d'écriture et une manière de construire des articles : accroche et intertitres qui relancent l'intérêt et facilitent la lecture indicielle, cette lecture par balayage très courante sur le web et axée sur la recherche d'information. Ainsi que la démarche de vérifier ses sources et de garantir l'information. Du conservateur, on retient l'idée de « prendre soin » (*curare*) de son contenu, de développer autour une expertise, et de le renseigner, le décrire avec des

⁹ Le portail Éduthèque : <http://www.edutheque.fr/accueil.html>

informations normalisées qui constitueront nos métadonnées. Du bibliothécaire, l'expérience de médiation autour des collections et son approche documentaire.

Document, corpus, canon

Toutes ces expertises dessinent une nouvelle manière de concevoir et présenter des contenus : construire un discours interactif et multimodal, en réseau avec des documents qui ouvrent sur les collections de la bibliothèque numérique, voire sur le web. En effet, « l'usage simultané et non séquentiel de différents contenus redessine la manière de construire et de transmettre les savoirs ». (Benhamou, 2014).

Car, si la numérisation change la nature du document, elle bouleverse autant la nature du corpus, désormais multimédia, multipliant les accès aux œuvres et interconnectant les documents. Le canon de la littérature évolue avec la numérisation : grâce à la mise en réseau, une attention plus soutenue peut être apportée aux mémoires et aux correspondances qui éclairent une sociologie de la littérature et remettent en contexte les « grandes œuvres ». Ainsi par exemple de l'éditorialisation des lettres de Voltaire dans le projet *Electronic Enlightenment (EE)*¹⁰. Cette base de données ne se limite pas aux seules lettres de Voltaire. Elle ouvre sur des correspondances en série, avec des possibilités de recherche sémantique, biographique et thématique. Hébergé par la Bodleian Library (bibliothèque universitaire d'Oxford), *Electronic Enlightenment* comprend actuellement près de 60 000 lettres et documents. Provenant de plus de 7 000 personnes – des penseurs et savants, politiciens et diplomates, mais aussi des gens ordinaires, fonctionnaires, commerçants, écrivant depuis l'Europe, l'Amérique ou l'Asie – ces correspondances permettent de saisir dans toute sa complexité la naissance de la société moderne au XVIII^e siècle.

De telles entreprises d'éditorialisation scientifique changent notre regard sur la littérature. C'est là un des enjeux des humanités numériques : grâce aux nouvelles technologies, changer l'approche du document, la nature du corpus, le canon même de la discipline, construire ainsi les conditions d'une épistémè numérique.

¹⁰ Informations reprises d'après le site web Electronic Enlightenment : <http://www.e-enlightenment.com/about/>

Références bibliographiques

- Bachimont, B., (2007). « Nouvelles tendances applicatives. De l'indexation à l'éditorialisation », Gros, P., (dir.), *L'indexation multimédia : description et recherche automatiques*. Paris, Lavoisier, Hermès sciences.
- Bachimont, B., (2011). « Enjeux et technologies : des données au sens », *Documentaliste-Sciences de l'Information*, vol. 48, n° 4.
- Benhamou, F., (2014). *Le livre à l'heure numérique : Papier, écrans, vers un nouveau vagabondage*. Paris, Le Seuil.
- Bon, F., (2011). *Après le livre*. Paris, Le Seuil.
- Dacos, M., Mounier, P., (2010). *L'édition électronique*. Paris, La Découverte.
- Di Pietro C., dir. (2014). *Produire des contenus documentaires en ligne : quelles stratégies pour les bibliothèques ?* Villeurbanne, Presses de l'enssib.
- Doueihi, M., (2008). *La Grande conversion numérique*. Paris, Le Seuil.
- Doueihi, M., (2011). *Pour un humanisme numérique*. Paris, Le Seuil.
- Dujol, L., (2011). « La bibliothèque, un plus pour le web social », La Bibliothèque approuvée, 23 février 2011. Disponible à : <http://labibapprouvee.wordpress.com/2011/02/23/la-bibliotheque-un-plus-pour-le-web-social/>
- Grafton, A., (2012). *La page de l'Antiquité à l'ère du numérique : histoire, usages, esthétiques*. Paris, Hazan.
- Jeanneney J.-N. (2007). *Quand Google défie l'Europe : Plaidoyer pour un sursaut*. Paris, Mille et une nuits.
- Laborderie, A., (2015). « Le livre numérique enrichi : enjeux et pratiques de remédiatisation ». Journées doctorales de la SFSIC. Lille, les 21 et 22 mai.
- Moss, A., (2002). *Les recueils de lieux communs. Apprendre à penser à la Renaissance*. Genève, Droz.
- Paquenseguy, F., Bossier, S., (2014). *Le livre numérique en questions*. Lille, Études de communication 2, n° 43.
- Pédaque, R. T. (2006). *Le document à la lumière du numérique*. Caen, C&F éditions.
- Schiffrin, A., (1999). *L'édition sans éditeurs*. Paris, La Fabrique éditions.
- Sinatra, M. E., Vitali-Rosati, M. (2014). *Pratiques de l'édition numérique*. Montréal, Presses de l'Université.
- Stiegler, B., (2012). « Pharmacologie de l'épistémè numérique », conférence prononcée aux Entretiens du Nouveau Monde Industriel. Paris, Centre Georges Pompidou, 17 décembre 2012.
- Vandendorpe, C., (1999). *Du papyrus à l'hypertexte. Essai sur les mutations du texte et de la lecture*. Paris, La Découverte.