



# **GALLICA (1997 – 2016)**

# DE LA BIBLIOTHÈQUE DE « L'HONNÊTE HOMME » À CELLE DU GALLICANAUTE

SOPHIE BERTRAND ET ALINE GIRARD

Gallica propose aujourd'hui 3,7 millions de documents issus majoritairement des fonds de la Bibliothèque nationale de France (BnF), mais aussi de ceux de ses partenaires. Cette

Avec 3,7 millions de documents disponibles, la bibliothèque numérique Gallica proposée par la BnF connaît une évolution constante. L'accroissement de ses collections, l'optimisation de ses fonctionnalités, l'interaction dynamique avec de nombreux partenaires numériques, le renforcement des relations avec les Gallicanautes sont les principales caractéristiques de son développement, une vingtaine d'années après sa création.

bibliothèque numérique, devenue aujourd'hui collective, a une vingtaine d'années et connaît une évolution constante : accroissement de ses collections, fonctionnalités optimisées, interaction dynamique avec de nombreux partenaires numériques, relation de plus en plus étroite avec les Gallicanautes<sup>1</sup>. Gallica n'est pas simplement une interface de mise à disposition de documents patrimoniaux,

mais l'une des matérialisations complexes d'une politique de diffusion et de conservation du patrimoine menée par une institution nationale.

<sup>1</sup> Cf. les articles en ligne : <http://www.archimag.com/reseaux-sociaux/2016/03/01/gallica-cache-derriere-reseaux-sociaux> ; <http://www.archimag.com/bibliotheque-edition/2015/11/18/gallica-refait-beauté-grâce-utilisateurs> et <http://liber-libri.blogspot.fr/2014/05/la-bibliotheque-sur-le-web.html>



<sup>2</sup> *Fictions*, Jorge Luis Borges, Folio Gallimard, 1974, p. 76.

<sup>3</sup> *La grande conversion numérique*, Milad Doueïhi, Seuil, 2008.

<sup>4</sup> Pour faciliter la lecture, les données chiffrées sur la collection et la fréquentation de Gallica sont rassemblées dans l'encadré « Gallica en chiffres ». Comme le montre la répartition par types de documents, Gallica reflète la richesse et la diversité des collections de la BnF et des autres bibliothèques françaises, ce qui la distingue de Google Livres avec lequel on a longtemps cherché à la comparer. Cf. Rapport Tessier du 12 janvier 2010 : <http://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/104000016.pdf>

<sup>5</sup> Une deuxième version de la charte a été publiée en 2007.

<sup>6</sup> [http://www.bnf.fr/fr/collections\\_et\\_services/collections\\_departements.html](http://www.bnf.fr/fr/collections_et_services/collections_departements.html)

<sup>7</sup> [http://www.bnf.fr/fr/professionnels/anx\\_catalogage\\_indexation/a\\_cotation\\_bnf.html](http://www.bnf.fr/fr/professionnels/anx_catalogage_indexation/a_cotation_bnf.html)

Mettre en ligne est-il un acte suffisant pour répondre au devoir de diffusion du patrimoine sur le web ? Le sentiment que tout numériser était possible n'est pas sans faire écho aux premières réactions des personnages qui découvrent la bibliothèque de Babel : « *Quand on proclama que la bibliothèque comprenait tous les livres, la première réaction fut un bonheur extravagant*<sup>2</sup>. » Néanmoins, atteindre l'exhaustivité est à tel point un illusoire dessein qu'apparaît l'oxymore « d'exhaustivité relative » ou « exhaustivité raisonnée<sup>3</sup> » pour qualifier un des objectifs de l'entreprise d'enrichissement de Gallica<sup>4</sup> amorcée depuis deux décennies.

### DE QUELLE COLLECTION NUMÉRIQUE PARLE-T-ON ?

Lancée dans les années 1990 dans la perspective de créer une bibliothèque numérique de « l'honnête homme », encyclopédique et raisonnée, Gallica a ouvert en 1997 avec 20 000 titres et visait un objectif de 100 000 titres : éditions consacrées par la postérité dans toutes les disciplines ; textes présentant les contextes culturel et scientifique des grandes œuvres ; documents connexes pour l'étude et la recherche, comme les périodiques et les ouvrages de référence. Une nouvelle version est mise en ligne en 2000, avec les premiers parcours thématiques structurés au sein des collections numérisées. En 2004, la première charte documentaire précisait l'évolution des collections<sup>5</sup> : les quelque 100 000 documents imprimés, 80 000 images et 30 heures de son alors disponibles dans Gallica s'inscrivaient dans une dominante disciplinaire en Histoire, Littérature, Sciences et Techniques. Majoritairement francophones, ces

ressources libres de droits proposaient une large variété de supports et allaient de l'Antiquité à la première moitié du XX<sup>e</sup> siècle, avec une forte présence de documents publiés au XIX<sup>e</sup> siècle. Un premier programme ambitieux de numérisation était lancé en 2005, celui de la presse quotidienne. En 2007, Gallica offrait 350 000 titres.

La période 2007-2009 a vu la mise en œuvre de la numérisation de masse en réaction à l'initiative de Google et le choix de la reconnaissance optique de caractères (OCR) pour une recherche au cœur des textes. Gallica s'est d'abord rapidement étoffée par la numérisation des livres et revues (2007), puis des collections spécialisées (2009) de la BnF. Ayant à conduire presque du jour au lendemain une entreprise industrielle de numérisation avec plus de 100 000 imprimés par an à dématérialiser, la BnF a procédé dans un premier temps non par sélection des documents, mais par désélection : une fois les fonds choisis par les départements d'imprimés<sup>6</sup> à partir de la cotation Clément qui classe en 23 divisions systématiques les imprimés entrés à la BnF de la fin du XVII<sup>e</sup> siècle à 1996<sup>7</sup>, étaient retirés les documents sous droits, les documents étrangers et ceux trop fragiles pour supporter la numérisation de masse.

Ont été néanmoins affirmés trois grands principes pour la politique de numérisation, qui a pris progressivement en compte une grande variété de supports :

- *l'intérêt patrimonial* : les documents auxquels la qualité esthétique, la valeur historique, la provenance, la rareté, confèrent un statut de chef-d'œuvre ou de trésor ;
- *l'intérêt documentaire* : sanctionné par les pratiques de la recherche, il caractérise les ensembles le plus demandés ;

# GALICA EN CHIFFRES

(au 31 décembre 2015)



## LA FRÉQUENTATION

Visites : 16 millions en 2015 <sup>1</sup>

Nombre moyen de pages vues par visite : 21,1

Temps moyen par visite : 13 min 46 s

## LA COLLECTION

### • Gallica

3 654 758 documents, soit 1 519 téraoctets de données.

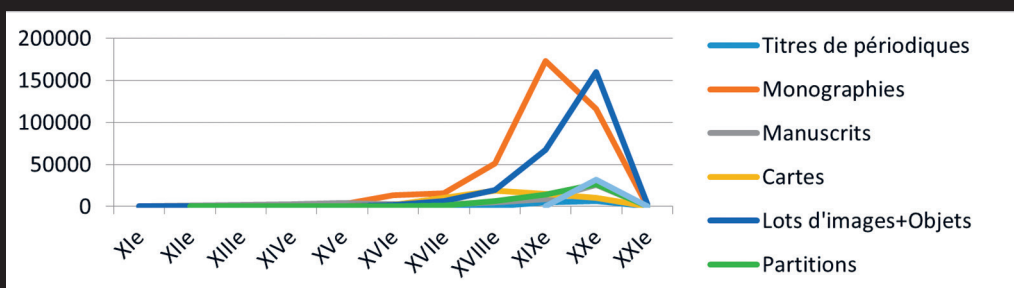
#### Dont, par provenance

- 3 437 465 documents issus des collections de la BnF et des bibliothèques partenaires et pleinement consultables dans Gallica
- 217 293 documents issus des collections des bibliothèques partenaires et référencés par interopérabilité des bibliothèques numériques

#### Dont, par type de documents (contenus consultables dans Gallica)

- 314 902 livres, dont 201 902 en mode texte
- 1 522 504 fascicules de presse et revues, dont 1 063 138 en mode texte
- 1 226 940 images
- 49 999 manuscrits
- 39 924 partitions
- 95 026 cartes
- 34 431 documents sonores
- 178 608 objets
- 13 vidéos

#### Dont, par siècle (à titre indicatif, pour les contenus consultables dans Gallica)



### • Gallica intra-muros

3 849 049 documents

<sup>1</sup> Gallica, Gallica embedded, Gallica intra muros, Gallica Labs et Numistral, d'une part, et site mobile et applications Gallica pour iPad et Android, d'autre part.



8 Documents acidifiés (en particulier la presse), grands formats, plaques de verre, objets avec des éléments mobiles, etc.

9 On consultera également l'article d'Arnaud Dherry «Gallica, construction et stratégie», in *La Pensée*, janvier-mars 2010, n° 361, pour un historique de Gallica. Disponible dans Gallica : <http://gallica.bnf.fr/ark:/12148/bpt6k9612074g.r>

10 [http://www.bnf.fr/fr/la\\_bnf/anx\\_connaitr/a\\_politique\\_documentaire.html](http://www.bnf.fr/fr/la_bnf/anx_connaitr/a_politique_documentaire.html)

11 <http://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/104000143.pdf>

12 [http://www.bnf.fr/fr/professionnels/cooperation\\_nationale/a\\_gallica\\_numerisation\\_partagee.html](http://www.bnf.fr/fr/professionnels/cooperation_nationale/a_gallica_numerisation_partagee.html)

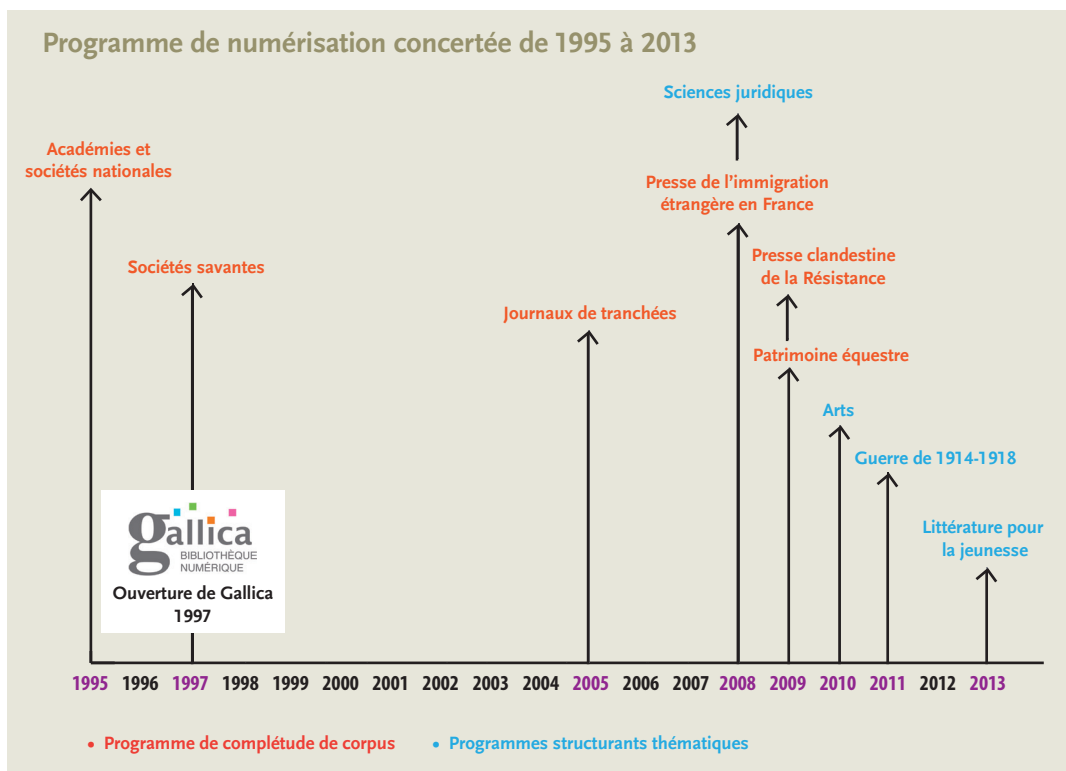
La BnF soutient la numérisation de documents du domaine public ou dont les droits ont été négociés, favorise la dématérialisation des collections d'imprimés hors presse (livres, revues, bulletins, actes, mémoires, thèses, etc.) et privilégie la numérisation de documents en français, dans les langues de France, en latin et en allemand (bibliothèques de l'ancienne Alsace-Moselle). La complémentarité optimale entre la numérisation des collections de la BnF et celles des partenaires est recherchée, afin de limiter la redondance documentaire, sans toutefois chercher à la supprimer totalement, la numérisation de plusieurs exemplaires d'un même titre étant parfois intéressante compte tenu des particularités d'exemplaire, d'envoi, de reliure, etc..

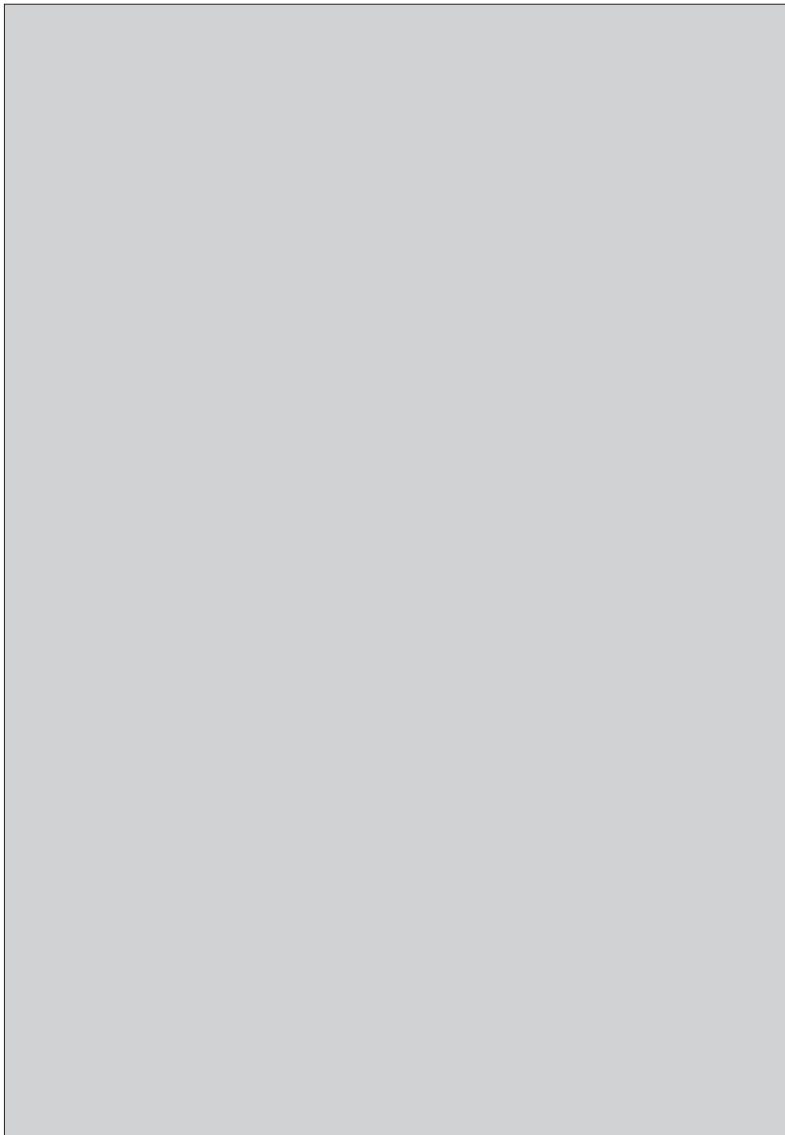
– *l'exigence de conservation* : elle motive la numérisation des documents fragilisés ou dont l'état rend la communication difficile ou risquée<sup>8</sup>. L'audiovisuel fait l'objet d'un traitement particulier (700 000 documents numérisés), les supports magnétiques, optiques et mécaniques étant fragiles et menacés de dégradation accélérée.

Depuis 2009, la dématérialisation des imprimés se poursuit, tout en s'efforçant d'équilibrer quantité et sélectivité, à travers notamment la numérisation par corpus, et de répondre aux demandes des usagers (numérisation à la pièce). Ce concept s'applique également aux collections spécialisées, à partir desquelles sont également constitués des ensembles documentaires à la fois cohérents et massifs, comme c'est le cas pour les cartes et plans, les partitions ou les monnaies (un quart des 700 000 monnaies de la BnF est numérisé)<sup>9</sup>. Mobilisée sur le front de la production et de la diffusion, la BnF a remis à plus tard la rédaction d'un document stratégique et programmatique pluriannuel. La Charte de la numérisation de la BnF est actuellement en cours d'élaboration couvrant la période 2015-2020; elle éclairera la

politique de numérisation de l'établissement, tout comme la Charte documentaire des enrichissements des collections (mars 2016) en éclaire la politique d'acquisitions<sup>10</sup>.

Depuis la fin des années 2000, la BnF a fait de la numérisation l'axe central de sa politique de coopération en France comme avec l'étranger. Au plan national, elle s'attache à favoriser la numérisation du patrimoine écrit des bibliothèques françaises et à créer de manière collaborative de vastes ressources patrimoniales numérisées destinées à enrichir Gallica, mais aussi les autres bibliothèques numériques. Pour faire suite aux recommandations du Schéma numérique des bibliothèques (mars 2010)<sup>11</sup> – mais confirmant ainsi une démarche engagée par la BnF depuis la fin des années 1990 avec les premiers programmes coopératifs de numérisation (publications des sociétés savantes et des académies nationales) –, la numérisation partenariale est organisée préférentiellement autour de programmes structurants – disciplinaires et d'intérêt régional – et de programmes de corpus<sup>12</sup>. La complétude documentaire est recherchée pour que les ensembles constitués soient cohérents et fassent sens. Ainsi les pro-





grammes de numérisation en Sciences juridiques, Art et Littérature pour la jeunesse ambitionnent-ils la mise en ligne à moyen/long terme d'ensembles homogènes scientifiquement pertinents (par exemple, pour les programmes évoqués, toutes les sources du droit coutumier, tous les catalogues de vente français, tous les abécédaires). Les projets de reconstitution virtuelle de corpus plus ciblés ont les mêmes objectifs (l'intégralité des portulans français)<sup>13</sup>. Cette option n'exclut cependant pas des programmes bilatéraux spécifiques, également légitimes sur le plan documentaire et producteurs de ressources numériques de grande

qualité<sup>14</sup>. Les partenariats nationaux de recherche sont eux aussi apporteurs de contenus numérisés, comme dans le cas de l'Equipex Biblissima<sup>15</sup> ou du Labex Obvil<sup>16</sup>. La coopération internationale et le mécénat font de leur côté entrer dans Gallica des contenus très diversifiés, dans le cadre de programmes européens (Biodiversity Heritage Library for Europe<sup>17</sup>, Europeana Regia<sup>18</sup>, Europeana Collections 1914-1918<sup>19</sup>), bilatéraux (France-Chine<sup>20</sup>, France-Brésil<sup>21</sup>, France-Japon<sup>22</sup>), multilatéraux (International Dunhuang Project<sup>23</sup>) ou sur projets (Roman de la rose<sup>24</sup>, presse ottomane en français<sup>25</sup>, numérisation de manuscrits arabes,

<sup>13</sup> [http://www.bnf.fr/documents/num\\_concertee\\_imprimes.pdf](http://www.bnf.fr/documents/num_concertee_imprimes.pdf)

<sup>14</sup> Entre 2009 et 2014, la numérisation partenariale a permis la numérisation de 17380000 pages. 176 partenaires nationaux ont été soutenus par la BnF, à hauteur de 6 109 000 €.

<sup>15</sup> <http://www.biblissima-condorcet.fr/fr/ressources/numerisations-a-bnf>

<sup>16</sup> <http://obvil.paris-sorbonne.fr/partenaires-institutionnels>

<sup>17</sup> <http://www.bhle.eu/fr/presentation>

<sup>18</sup> <http://www.europeanaregia.eu/fr>

<sup>19</sup> <http://www.europeana-collections-1914-1918.eu/>

<sup>20</sup> <http://gallica.bnf.fr/html/und/asia/france-chine>

<sup>21</sup> <http://bndigital.bn.br/francebr/frances/presentation.htm>

<sup>22</sup> <http://gallica.bnf.fr/html/und/asia/france-japon>

<sup>23</sup> <http://idp.bnf.fr/idp.a4d>

<sup>24</sup> <http://romandelarose.org/>

<sup>25</sup> [https://www.archives.saltresearch.org/R/PIDT44SGLTKUGSHCMRJA86PAJ85MN4SKQQ5AEE2BCLVHVKPS41-03858?func=collection\\_s&collection\\_id=2244&pds\\_handle=GUEST](https://www.archives.saltresearch.org/R/PIDT44SGLTKUGSHCMRJA86PAJ85MN4SKQQ5AEE2BCLVHVKPS41-03858?func=collection_s&collection_id=2244&pds_handle=GUEST)



<sup>26</sup> Pour soutenir l'effort de numérisation, de multiples sources de financement sont mises en œuvre, se cumulant pour donner à la BnF une capacité d'action déterminante dans le domaine de la numérisation du patrimoine : fonds propres de l'établissement; crédits du Centre national du livre pour les imprimés; mécénat d'entreprise et individuel; financements par projets au niveau national ou international, sur fonds privés ou européens; mobilisation des « Investissements d'avenir »; prestations de reproduction.

<sup>27</sup> Une série de douze vidéos, *L'aventure de la numérisation*, présente les apports des partenaires de la BnF dans le marché de numérisation des imprimés, financé sur crédits du CNL et ouvert pour un tiers de sa volumétrie globale aux collections des autres bibliothèques françaises [http://www.bnf.fr/fr/collections\\_et\\_services/aux\\_bib\\_num/a.gallica\\_videos.html](http://www.bnf.fr/fr/collections_et_services/aux_bib_num/a.gallica_videos.html)

<sup>28</sup> [http://www.bnf.fr/documents/partenaires\\_gallica.pdf](http://www.bnf.fr/documents/partenaires_gallica.pdf)

<sup>29</sup> Le Schéma numérique de la BnF (mars 2016) [http://www.bnf.fr/documents/bnf\\_schema\\_numerique.pdf](http://www.bnf.fr/documents/bnf_schema_numerique.pdf) permet une approche panoramique de la stratégie numérique de l'établissement et inscrit Gallica dans la politique numérique globale de la Bibliothèque nationale de France.

turcs et persans, etc.), souvent ouverts aux autres bibliothèques françaises ainsi associées à des projets scientifiques à large visibilité<sup>26</sup>. Aujourd'hui, Gallica s'enrichit grâce à trois modes d'entrée de contenus patrimoniaux externes dans la bibliothèque numérique : l'intégration par numérisation de documents physiques des partenaires dans les marchés de dématérialisation et ateliers de la BnF<sup>27</sup>, l'intégration de fichiers numériques et le référencement des bibliothèques numériques des

partenaires par moissonnage (interopérabilité OAI-PMH). Gallica donne accès aux collections numérisées de 300 partenaires, avec un apport de plus de 400 000 documents en provenance d'autres institutions<sup>28</sup>.

L'ancienneté de l'entreprise, l'éventail des axes de numérisation de la BnF, la diversité des programmes coopératifs et la variété des canaux qui alimentent Gallica obligent à s'interroger sur la nature de cette collection numérique hors norme de près de 4 millions de documents<sup>29</sup>.

## GALLICA SE DÉFINIT-IL SEULEMENT PAR LE VOLUME CROISSANT DE SON CONTENU PATRIMONIAL ?

La Bibliothèque nationale de France<sup>30</sup> et les bibliothèques municipales<sup>31</sup> de l'Hexagone détiennent quelque 25 millions de livres ou de manuscrits antérieurs au XX<sup>e</sup> siècle. Il faut y ajouter les fonds patrimoniaux des bibliothèques universitaires et spécialisées, encore difficilement quantifiables à ce jour malgré les remarquables avancées de la description des fonds et de la rétroconversion des catalogues, que reflète le Catalogue collectif de France<sup>32</sup>. Au regard de ces seuls chiffres, il apparaît donc que Gallica ne donne à voir qu'un faible échantillon des fonds disponibles dans les bibliothèques françaises. En 2012, Denis Bruckmann, à raison, pose la question de manière directe : «*Que représentent les 300 000 ouvrages numérisés par rapport aux millions conservés*<sup>33</sup> ? » Nombre de bibliothèques possèdent des fonds qui s'apparentent à un « bois dormant » avec des livres

qui ne sont jamais empruntés ou feuilletés. Pierre-Marc de Biasi rappelle qu'«*un calcul effectué en 1995 a montré qu'en moyenne chaque livre conservé à la Bibliothèque nationale de France avait une chance d'être communiqué à un lecteur tous les treize ans*<sup>34</sup> ». Mais ce résultat doit être pondéré puisque certains titres suscitent plus de communications que d'autres. Numériser semble une entreprise presque dérisoire, il reste tant à faire.

S'ajoute à ce constat le fait que, grâce au dépôt légal, la collection de livres et revues s'accroît à raison de 6 km de papier par an, sans parler des autres documents imprimés (presse, partitions, cartes, etc.). La révolution de l'édition numérique est en cours, mais pas au point de réduire drastiquement le volume de papier, d'autant plus que la production éditoriale est exponentielle au regard du nombre de titres publiés par an (près de 76 000 titres par an en moyenne sur les quatre dernières années). Le patrimoine n'est pas un ensemble aux limites clairement définies et doit s'appréhender comme une matière en perpétuelles extension

<sup>30</sup> Les collections de la seule BnF s'élevaient à environ 14 millions de livres et 300 000 volumes de manuscrits, auxquels s'ajoutent 360 000 titres de périodiques, 12 millions d'estampes et photographies, plus de 800 000 cartes et plans, 2 millions de partitions, 700 000 monnaies et médailles, 1,1 million de documents sonores et 250 000 vidéos.

<sup>31</sup> Le ministère de la Culture et de la communication indique en 2005 : «*Environ 500 bibliothèques détiennent des fonds rares et précieux, soit 21,7 % des bibliothèques publiques. Les bibliothèques publiques conservent plus de 9,5 M de volumes dans leurs fonds patrimoniaux, dont 6 M d'imprimés antérieurs au XIX<sup>e</sup> siècle, 30 000 manuscrits médiévaux et 150 000 manuscrits modernes* » (enquête précédant la mise en œuvre du Plan d'action pour le patrimoine écrit).

<sup>32</sup> <http://ccfr.bnf.fr/portailccfr/>

<sup>33</sup> Denis Bruckmann et Nathalie Thouny, « La numérisation à la Bibliothèque nationale de France et les investissements d'avenir », *Bulletin des bibliothèques de France (BBF)*, n° 4, 2012, p. 49-53. En ligne : <http://bbf.enssib.fr/consulter/bbf-2012-04-0049-010>

<sup>34</sup> Pierre-Marc de Biasi, « Le patrimoine écrit », *Item*. <http://www.item.ens.fr/index.php?id=13597>





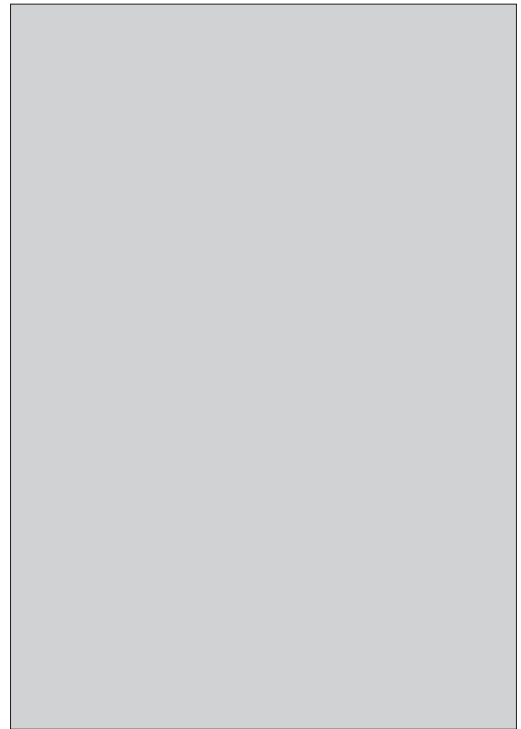
et reconstitution<sup>35</sup> du fait de la patrimonialisation progressive des collections. Le travail de catalogage *a posteriori* réalisé par le service de l'Inventaire rétrospectif de la BnF illustre à quel point le signalement des collections à travers les siècles peut s'avérer lacunaire. Tous les jours, par des dons et des legs de collectionneurs privés, le patrimoine s'enrichit et s'agrandit en parallèle de la procédure instituée du dépôt légal.

Recenser le patrimoine écrit est donc un travail de navigateur qui se doit de tenir le cap face aux marées de documents «endormis», «oubliés» ou inédits qui sont stockés dans les bibliothèques. Aussi, à la question «peut-on tout numériser?», la réponse est clairement «non». Le principal enjeu n'est donc pas là, même si Gallica acquiert une visibilité et une notoriété internationale grâce à sa nature d'important «réservoir» documentaire<sup>36</sup>.

### GALLICA N'EST-IL PAS AUSSI UN MIROIR RÉVÉLATEUR DES ENJEUX NUMÉRIQUES DE LA BNF ?

Alors même qu'il s'agit d'une entreprise sans fin au regard de la définition dynamique du patrimoine écrit, un des paradoxes de la numérisation patrimoniale est de se fonder pour grande partie sur des processus industriels de traitement encadrés par des marchés publics de numérisation. Ce dispositif a l'avantage d'assurer une productivité certaine, puisqu'à peu près 1500 nouveaux documents de tous types sont mis en ligne sur Gallica chaque semaine. Afin d'intensifier la production de fichiers numériques, la BnF a initié des partenariats publics/privés permettant, selon Bruno Racine, «une démultiplication phénoménale de l'accès à (une) collection quasiment inconnue et pratiquement inaccessible en dehors de quelques spécialistes<sup>37</sup>».

Une des premières raisons qui légitime la numérisation patrimoniale est donc la mise à disposition de ressources documentaires pour chaque citoyen. Le document numérisé devient alors un véritable bien commun dès lors qu'il est diffusé librement et gratuitement sur la Toile puisqu'il est non rival et non exclusif. Gallica représente tout à la fois un objectif (chaque ci-



toyen doit avoir accès à un bien commun culturel) et un moyen (Gallica est un outil de diffusion). Cette dimension politique – dans le sens où cela participe à l'organisation de l'accès au savoir – a été perceptible à travers le positionnement de Jean-Noël Jeanneney<sup>38</sup> face au défi lancé par Google en décembre 2004. Lorsque l'entreprise californienne annonce qu'elle souhaite numériser sur six ans 15 millions d'ouvrages, le risque pressenti par le président de la BnF de l'époque est de voir, pour certains livres, se réduire leur capacité à demeurer un bien commun. Google pourrait décider un jour ou l'autre de restreindre les accès à ces documents patrimoniaux. Cette position a été largement soutenue par les pouvoirs publics qui ont donné à la BnF les moyens financiers d'une ambition numérique nationale à travers la numérisation de masse de ses seuls fonds dans un premier temps, des fonds d'autres bibliothèques dans un second temps. Mis à part cet événement, Gallica continue de se construire en «réaction» à de nouveaux enjeux proposés à la BnF. Il ne s'agit plus seulement de se positionner par rapport à certains acteurs du GAFA<sup>39</sup>. Gallica révèle aussi les ressources numériques nouvellement créées par les différents professionnels du secteur du livre

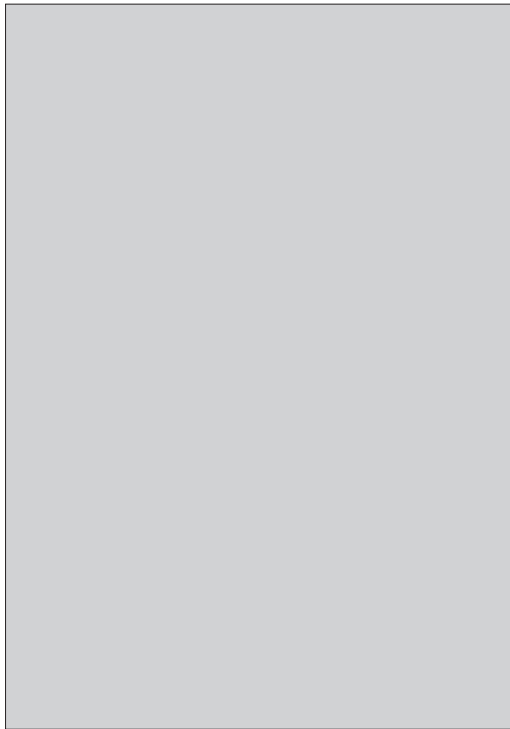
<sup>35</sup> L'extension et la reconstitution du patrimoine prennent aussi en compte le patrimoine numérique constitué depuis les années 2000 par l'archivage du web (26 milliards d'URL et 668 To de données fin 2015). Ce dernier dépassera bientôt celui accumulé par des siècles de conservation de documents physiques. De plus, le dépôt légal des livres numériques fera évoluer les rapports entre matérialité et dématérialité du patrimoine.

<sup>36</sup> Gallica est une des plus grandes bibliothèques numériques au monde avec celles de la British Library et de la Library of Congress.

<sup>37</sup> «La Bibliothèque de France au défi de la numérisation», Bruno Racine, <http://www.lemonde.fr> en date du 1<sup>er</sup> février 2013.

<sup>38</sup> Jean-Noël Jeanneney, *Quand Google défie l'Europe*, éd. Mille et une nuits, 2005, p. 18.

<sup>39</sup> GAFA pour Google, Apple, Facebook, Amazon.



au travers de chantiers nationaux inédits, comme le projet de numérisation des livres indisponibles du XX<sup>e</sup> siècle<sup>40</sup> ou bien encore les partenariats publics-privés dans le cadre des Investissements d'avenir<sup>41</sup>, sans oublier, bien sûr, le dépôt légal des livres numériques. Gallica n'est plus seulement une plateforme d'accès au contenu relevant du domaine public<sup>42</sup>, mais devient peu à peu un outil de signalement. Ainsi, dans les listes de résultats de Gallica, sont aujourd'hui référencés des documents exclusivement consultables dans leur intégralité sur Gallica intra-muros<sup>43</sup> que ce soit des ressources nées numériques ou des avatars de documents physiques.

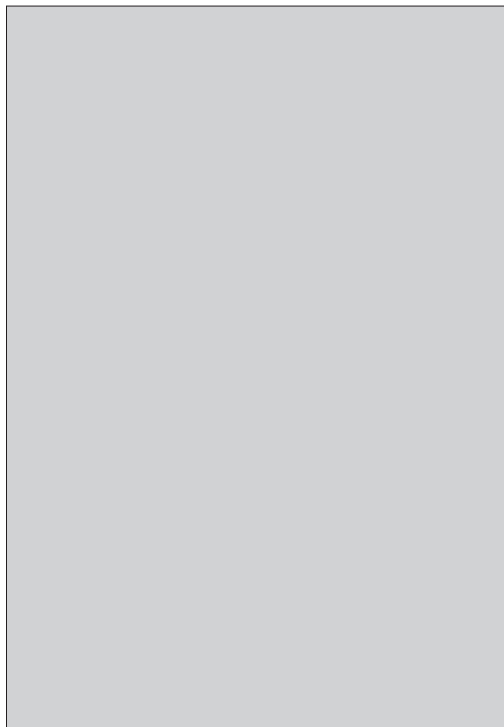
Gallica mue régulièrement et le développement de ses fonctionnalités constituent autant de marquages explicites de ces évolutions<sup>44</sup>; elle réinvente, au fil des nouveaux enjeux numériques de la BnF, l'imbrication pertinente qu'elle doit avoir avec les missions de l'institution. Gallica était à l'origine l'outil informatique de mise à disposition des documents à la BnF (sur des PLAO<sup>45</sup>) avant de devenir un service « hors les murs » avec l'arrivée du web. Aujourd'hui, Gallica réaffirme sa complémentarité avec les salles de lecture de la BnF mais d'une manière différente. L'exemple des archives du Web, dis-

ponibles dans les emprises de la BnF, illustre le fait que le Gallicanaute, pour embrasser toutes les formes du patrimoine numérique, doit aussi exploiter des services en intra-muros.

C'est un des paradoxes assumés de l'histoire de Gallica, qui tout en étant le symbole de la numérisation des fonds patrimoniaux, recrée avec régularité une logique d'ancrage dans les sites physiques de la BnF pour assurer un lien entre patrimoine dématérialisé et patrimoine né numérique.

### LES MUES DE GALICA EN FONT-ELLES EN SOI UN OBJET PATRIMONIAL ?

Gallica répond à une définition plurielle à plus d'un titre. D'abord, la bibliothèque numérique propose de plonger dans un grand réservoir documentaire aussi riche que varié et suggère aussi des documents en les signalant sans rien dévoiler de leur contenu ou en n'en montrant que des extraits. Elle construit de plus en plus dans l'acte de recherche un lien entre le domaine public et les contenus sous droits en créant une étroite articulation avec Gallica intra-



<sup>40</sup> Projet ReLire : <https://relire.bnf.fr/accueil>

<sup>41</sup> BnF Partenariat, filiale de la Bibliothèque nationale de France : [http://www.bnf.fr/fr/la\\_bnf/anx\\_actu\\_bib/a.partenariats\\_numerisation\\_bnf.html](http://www.bnf.fr/fr/la_bnf/anx_actu_bib/a.partenariats_numerisation_bnf.html)

<sup>42</sup> Les documents sont consultables directement sur Gallica ou indirectement sur d'autres bibliothèques numériques par le biais de l'interopérabilité.

<sup>43</sup> Site consultable dans les seules emprises de la BnF. Il permet d'accéder à l'ensemble des ressources présentes dans Gallica, à des contenus soumis au droit de la propriété intellectuelle ou relevant de règles spécifiques de diffusion (Partenariat Believe/Memnom, Proquest, projet ReLire).

<sup>44</sup> Par exemple, les modes de visualisation par défaut qui optimisent la consultation du document, la puissance du zoom qui exploite les excellentes résolutions de numérisation de certains documents, le mode feuilletage pour les indisponibles, sans parler des fonctionnalités incitant à la dissémination des contenus.

<sup>45</sup> Les premiers documents numérisés étaient disponibles sur des PLAO (postes de lecture assistée par ordinateur) dans les nouvelles salles de la bibliothèque François-Mitterrand à Tolbiac. Pour les origines de Gallica, on consultera « La bibliothèque numérique de la BnF : conception, méthodes, résultats » par Daniel Renault, in *Document numérique*, 1998.



muros. Ensuite, Gallica offre une forêt touffue de documents mais, depuis 2013, trace avec régularité des chemins et des carrefours balisés vers des collections par le biais de la médiation numérique. Enfin, Gallica se laisse de plus en plus apprivoiser par ses usagers et tend à se définir à travers eux.

Ces différents éclairages traduisent que Gallica n'est pas seulement un contenant dématérialisé mais est *en soi* un objet patrimonial révélant l'aventure de la numérisation de la BnF. La bibliothèque numérique traduit l'évolution méthodologique et technique de la numérisation patrimoniale réalisée par une institution nationale. La BnF, pour garantir la mémoire des objets qu'elle conserve, doit s'assurer que les différentes déclinaisons d'un contenu sont préservées. Gallica illustre non seulement les strates technologiques de la numérisation réalisée depuis vingt ans<sup>46</sup> mais exprime aussi le continuum qui existe dans les supports qui se succèdent, par exemple la numérisation en continu des microfilms. Ainsi, la conservation des documents patrimoniaux impliquerait une mise en abyme sans fin.

Gallica est donc le résultat d'une double exigence : numériser toujours mieux pour améliorer le confort de lecture de l'utilisateur, numériser toujours plus pour réduire le risque de perte irréversible d'un document. Bref, comme le dit joliment l'architecte Massimiliano Fuksas : « Réunir des besoins différents, souvent opposés, est une manière d'appréhender et de vivre les mille contradictions de notre temps. D'où le chaos. Un chaos sublime<sup>47</sup>. »

#### LE GALLICANAUTE EST-IL L'ANIMA<sup>48</sup> DE GALLICA ?

Ces contingences techniques mêlées à la productivité numérique exceptionnelle de ces quinze dernières années auraient-elles éloigné Gallica de sa définition initiale de « bibliothèque de l'honnête homme » du dix-septième siècle qui « sait quelque chose de tout » pour reprendre la formule pascalienne ? Gallica est-elle à considérer comme un avatar de fonds exis-

tants ou au contraire, reproduit-elle au niveau documentaire et, dans une certaine mesure, « un chaos sublime » comme l'évoque Fuksas ? Auquel cas, la collection numérique que présente Gallica serait à appréhender comme inédite. En effet, Gallica ne reproduit pas l'histoire de la constitution des fonds de la BnF et de celles des bibliothèques partenaires. Elle propose à l'utilisateur une pluralité d'ensembles documentaires.

Ce qui importe, ce n'est pas le nombre de documents ainsi rassemblés, mais l'écho signifiant qui naît de leur mise en interaction. La structuration de Gallica s'articule alors selon plusieurs modes : réticulaire (les documents entretiennent une relation « horizontale » où n'existe aucune priorisation) et arborescente (les documents entretiennent des relations « verticales », ils sont subordonnés les uns aux autres et présentent entre eux une hiérarchie génétique, structurelle, qualitative). Cette organisation à plusieurs facettes est riche : un même document peut appartenir à plusieurs corpus en révélant dans chacun d'eux une particularité différente. Le système est donc « plus » que la somme de ses parties. Il démultiplie le potentiel d'une unité documentaire.

Cependant, la collection ne prend vie que si le Gallicanaute s'en empare. Cette prise de conscience est revendiquée explicitement sur la page d'accueil de Gallica depuis 2015 : « Une bibliothèque numérique n'est pas seulement faite de contenus numérisés et de services et fonctionnalités associés à ces contenus. Les usages et pratiques qui s'y greffent contribuent à en structurer l'identité. Aussi les Gallicanautes – internautes qui utilisent Gallica et participent activement à la diffusion de ses documents sur le web – ont leur place au sein de la bibliothèque numérique<sup>49</sup>. » Pour s'en convaincre, il ne suffit pas de se contenter des 50 000 visites par jour sur le site. « Tomber sur Gallica » au détour d'une requête effectuée sur un des plus gros moteurs de recherche n'est pas « tomber dans Gallica » comme plusieurs usagers en témoignent à travers les réseaux sociaux. Par exemple, la communauté des généalogistes, qu'ils soient amateurs ou professionnels, par

<sup>46</sup> Comme le mode image, le mode texte, la numérisation en noir et blanc, puis en niveaux de gris et en couleurs avec des résolutions de plus en plus optimisées, création d'e-pubs, de fichiers accessibles, etc. Voir : [http://www.bnf.fr/fr/professionnels/anx\\_numerisation/a\\_numerisation\\_historique.html](http://www.bnf.fr/fr/professionnels/anx_numerisation/a_numerisation_historique.html)

<sup>47</sup> Massimiliano Fuksas, *Chaos sublime*, éd. Arléa, 2010, p. 15.

<sup>48</sup> *Anima, ae*, n. f. : le souffle de vie.

<sup>49</sup> <http://gallica.bnf.fr/html/und/du-cote-des-Gallicanautes>

LES  
PARTENAIRES  
DE GALICA

Les partenaires de Gallica par type de partenaires (toutes filières après dédoublement) – 1997-2015	
➤ Partenaires des territoires	➤ 204
➤ Partenaires de l'Enseignement supérieur et de la Recherche	➤ 45
➤ Autres partenaires	➤ 56
<b>Total des partenaires de Gallica</b>	<b>➤ 305</b>
Les partenaires de Gallica par mode d'entrée des documents numériques (sans dédoublement : un même partenaire participant à plusieurs filières est compté pour chacune des filières) – 1997-2015	
➤ Filière intégration des documents	➤ 334
• Intégration de documents physiques dans les chaînes de numérisation BnF – imprimés	270
• Intégration de documents physiques dans les chaînes de numérisation BnF – documents à haute valeur patrimoniale	40
• Intégration de fichiers numériques	24
➤ Filière référencement des documents (moissonnage)	➤ 68
<b>Total des partenaires de Gallica</b>	<b>➤ 402</b>
Nombre de documents des partenaires accessibles dans Gallica par filières – Fin 2015	
➤ Intégration par numérisation des documents dans les marchés et ateliers de la BnF	➤ 157 861
➤ Intégration des fichiers numériques	➤ 33 771
➤ Référencement par moissonnage des bibliothèques numériques partenaires	➤ 217 293
<b>Total des documents des partenaires</b>	<b>➤ 408 925</b>

leurs fouilles approfondies dans Gallica, montrent comment la pertinence des ressources mises en ligne est sans cesse réinterprétée par eux. Sophie Boudarel commente en ces termes sur son blog le travail de médiation numérique sur les « Ressources généalogiques<sup>50</sup> » co-construites avec des Gallicanautes : « *Nous sommes tous conscients que, si tout est potentiellement intéressant pour le généalogiste dans Gallica, tout ne peut pas être regroupé dans cette page<sup>51</sup>.* » La compilation de ressources, l'organisation des accès dans Gallica est nécessaire mais insuffisante. Aussi, les résultats de l'étude en trois volets<sup>52</sup> des usages de Gallica prévue en 2016 fourniront des éléments déterminants pour que la BnF s'achemine vers une bibliothèque numérique plus contributive.

La numérisation patrimoniale de masse, dès lors qu'elle est mise en ligne, donne à voir autrement le patrimoine écrit. Le véritable enjeu d'une bibliothèque numérique qui rassemble en son sein un seuil volumétrique conséquent de documents est la manière dont elle va « multiplier les livres sans fin<sup>53</sup> » en proposant une diversité d'expériences de lectures. C'est là que se situe en partie l'ambition des humanités numériques. Lire ou relire autrement. Structurer, classer, définir sans restreindre les champs d'investigation possibles ajoutent à la puissance d'un corpus numérisé. Cependant, pour que le tissu dont résulte ce complexe maillage fasse sens, il lui faut un couturier. Gallica ne prend vie que si le Gallicanaute se l'approprie et invente de nouveaux usages. Le projet est donc ambitieux, collectif et passionnant.

**B:F**

<sup>50</sup> <http://gallica.bnf.fr/html/und/histoire/ressources-genealogiques>

<sup>51</sup> <http://la-gazette-des-ancetres.fr/toujours-plus-de-genealogie-dans-gallica/>

<sup>52</sup> Analyse des logs de connexion à Gallica, puis enquête en ligne, et enfin phase d'ethnographie numérique.

<sup>53</sup> Cf. citation de l'Ecclésiaste (12,12) par Anthony Grafton dans *La page, de l'antiquité à l'ère du numérique*, éd. du Louvre, 2012, p. 28.