



HAL
open science

De l'accès aux documents mathématiques de Gallica et des coordinations avec d'autres programmes de numérisation

Laurent Guillopé

► **To cite this version:**

Laurent Guillopé. De l'accès aux documents mathématiques de Gallica et des coordinations avec d'autres programmes de numérisation. [Rapport Technique] Mathdoc. 2001. hal-03793075

HAL Id: hal-03793075

<https://hal.univ-grenoble-alpes.fr/hal-03793075>

Submitted on 22 Dec 2022

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.

De l'accès aux documents mathématiques de Gallica et des coordinations avec d'autres programmes de numérisation

Laurent Guillopé
laurent.guillopé@math.univ-nantes.fr

Cette note analyse quelques directions de mise en valeur du fonds mathématique de Gallica. Il est souhaitable que ce travail se développe en parallèle à la sélection de nouvelles collections à numériser (dans la suite de Gallica ou par des programmes extérieurs à la BNF).

La première direction est celle d'un catalogage plus fin des ressources actuellement présentes dans Gallica afin que le chercheur puisse localiser rapidement un document au niveau de l'unité logique (l'article d'une revue ou d'œuvres choisies, le chapitre d'une monographie) et le récupérer par un appel simple (évitant un feuilletage pour déterminer la portion du document à télécharger). Une granularité du catalogage est à l'œuvre dans des projets de numérisation de périodiques, par exemple JSTOR (dont l'accès requiert un abonnement payant) ou à la bibliothèque de l'Université de Göttingen et de l'État de Basse-Saxe (d'accès libre, mais qui n'offre pas actuellement de format d'impression de qualité minimum, PDF p. ex.).

Un second objectif est d'insérer la signalisation des fonds mathématiques de Gallica aux cotés d'autres ressources de ce type, issues de projets de numérisation nationaux (comme le programme NUMDAM piloté par la Cellule MathDoc pour le compte du CNRS) ou étrangers. Ce montage sera à la base de coopérations, évitant des duplications et renforçant la visibilité de Gallica dans le secteur de la recherche en mathématique et histoire des mathématiques.

Ces travaux seraient placés sous l'égide d'une structure de type *pôle associé pour les mathématiques* qui rassemblerait les bibliothèques des laboratoires de mathématiques fédérées au sein du RNBM, certaines bibliothèques ayant un fonds mathématique important, les sociétés savantes, le Comité national des mathématiciens, la section *Mathématiques* de l'Académie, la Cellule MathDoc et serait dirigée par un Conseil scientifique. Cette structure devra être spécialement attentive à impulser diverses collaborations : au niveau international (Göttingen, Linz, Cornell, Ann Arbor,...) , mais aussi au niveau national (l'unité mixte de service UJF/CNRS Cellule MathDoc, l'unité propre du CNRS *Archives de la création mathématique*, le Laboratoire de Philosophie et d'Histoire des Sciences - Archives H. Poincaré de l'Université Nancy II,...)

Le corpus Gallica en ligne actuellement propose différents types de documents imprimés. La nature de la mise en valeur envisagée varie suivant le type de document, mais les impacts tant scientifique qu'au niveau de la visibilité internationale sont similaires. Dans la suite, on développe quelques remarques sur les

- périodiques,
- œuvres complètes,
- monographies,

en terminant par l'évocation du *Répertoire bibliographique des sciences mathématiques*.

I. Périodiques

Dans Gallica, les périodiques sont signalés par année et il n'est pas possible d'interroger au niveau des articles : l'interrogation sur les sommaires est possible, mais l'orientation reste très difficile (un nom propre peut apparaître comme le nom d'un auteur ou comme élément d'un titre) et requiert un examen en ligne du document (préciser sa nature, déterminer sa longueur avant téléchargement).

Dans un premier paragraphe, on décrit un travail aujourd'hui affiché sur le serveur de la Cellule MathDoc pour un titre disponible dans Gallica. Puis, on évoque un travail analogue sur autre titre classique, travail dont l'intérêt scientifique est internationalement reconnu, mais qui demande des ressources qu'il conviendrait d'estimer plus précisément...

1. Le *journal de mathématiques pures et appliquées* (1836-1869, 1871-1880) [JMPA]

Les tables de matières ont été récupérées, analysées (éclatement en champs auteurs, titre, série, pages), mises au format (traduction TeX de toutes les formules, rétablissement des formules non saisies car trop complexes), avec ajout d'une URL pour chaque article (permettant le téléchargement à partir de Gallica de tout l'article, de la première à la dernière page). Une base de données a été réalisée, qui est interrogeable suivant de multiples critères et les réponses donnent un accès immédiat aux documents originaux. En outre, pour les articles parus entre 1868 et 1880 et signalés dans le *Jahrbuch über die Fortschritte des Mathematik* (JFM), un lien vers la notice dans la version numérisée du JFM a été introduit ; inversement, un lien dans la notice du JFM numérisé pointe vers l'article du JMPA correspondant lorsqu'il est présent dans Gallica. C'est le Mémoire sur l'emploi de l'épaisseur dans la théorie des surfaces élastiques de Sophie Germain.

Chaque article a une notice du type suivant

<ARTICLE>

<ID_NUMDAM>JMPA_1846_1__197_0</ID_NUMDAM>

<JOURNAL>Journal de mathématiques pures et appliquées</JOURNAL>

<ISSN>0021-7824</ISSN>

<ANNEE>1846</ANNEE>

<SERIE>1</SERIE>

<PAGES>197-200</PAGES>

<AUTEUR>Svanberg, A.-F.</AUTEUR>

<TITRE>Sur les intégrales définies $\int_0^\infty \frac{e^{-\beta x} x^{m-1} dx}{1+x^2}$, $\int_0^\infty \frac{\cos\{\beta x\} x^{m-1} dx}{1+x^2}$, $\int_0^\infty \frac{\sin\{\beta x\} x^{m-1} dx}{1+x^2}$.</TITRE>

<URL_NUM>http://gallica.bnf.fr/scripts/get_page.exe?O=16390&E=00000205&N=4&CD=0&F=PDF</

URL_NUM>

</ARTICLE>

Poursuivre la numérisation de ce journal (qui est encore vivant) serait évidemment d'un grand intérêt.

2. Comptes rendus de l'Académie des Sciences [CRAS]

Il s'agit d'une revue interdisciplinaire (mathématiques, physique, biologie, astronomie, médecine,...), numérisée à partir de 1835 ; la dernière année accessible actuellement sur Gallica est 1931 (respectant le délai des 70 ans de droit d'auteurs). Les articles sont courts, donc nombreux (70 000 ?). Un catalogage précis est souhaité (demande d'un collègue américain) et l'exemple des publications de H. Lebesgue (dont une vingtaine sont déjà numérisées, sur Gallica ou ailleurs) montre la pertinence d'un tel travail : en témoignent tant l'existence de travaux de référence (la note de 1901 faisant date par l'annonce de ce qui est désormais connu, dans le monde entier, sous le nom d'intégrale de Lebesgue a été l'objet récemment d'une conférence à l'ÉNS Lyon) que le bruit important de l'accès actuel sur Lebesgue (34 réponses en interrogeant sur le mot Lebesgue dans le champ *sommaires* pour 17 notes seulement avec comme auteur Lebesgue). En outre, l'utilisateur de Gallica apprécierait de pouvoir télécharger par un appel simple les pages contenant le texte d'une note (actuellement, il est contraint à faire quelques essais).

Effectuer pour les CRAS ce qui a été réalisé pour le JMPA est un travail considérable : il pourrait faire l'objet d'une collaboration entre la BNF (fourniture des sommaires, avec identifiant physique) et un groupe de mathématiciens. Avant démarrage, le temps de réalisation (et le coût) doit être estimé.

3. Autres revues

D'autres journaux devraient être l'objet de numérisation, p. ex. la *Correspondance de l'École Polytechnique*, le *Journal de l'École Polytechnique*, le *Journal de Gergonne*, les *Nouvelles annales de mathématiques*, le *Bulletin de la société philomatique de Paris*, les *Annales de la Faculté des sciences de l'Université de Toulouse pour les sciences mathématiques et les sciences physiques*.

D'autres revues sont aussi, de manière encore plus partielle, concernées : *Revue générale des sciences pures et appliquées*, le *Journal des savants*, les *Procès-verbaux et mémoires de l'Académie Royale des Sciences*. Un état des lieux est à faire.

II. Œuvres complètes

L'exemple typique est celui des *Éléments* d'Euclide. Cette oeuvre peut être considérée comme le paradigme de l'oeuvre mathématique, que ce soit par sa place centrale dans l'enseignement des mathématiques au cours des siècles passés, par la nature de l'exposition qui en fait le modèle de l'exposé mathématique ou ses multiples éditions. L'édition présente dans Gallica compte près de 700 pages, une présentation par chapitres (les 15 éléments, les données) rend abordable l'ensemble. Un autre exemple est celui des oeuvres de Lobachetsky.

Il y a des multiples éditions d'oeuvres complètes, dans des volumes (Abel, Beltrami, Brioschi, Desargues, Fourier, Halphen, Hermite, Jacobi, Laguerre, Laplace, Mobius,...) ou dans des revues (celles de Galois dans le JMPA) dont l'accès par un catalogage au niveau des articles faciliterait énormément la lisibilité. Quelques unes de ces oeuvres sont volumineuses (Cauchy, Euler, Poincaré,...). Beaucoup sont d'une grande pertinence pour les mathématiques d'aujourd'hui.

Le type de catalogage souhaité est analogue à celui d'un périodique et de ses articles. Ainsi, l'utilisateur pourrait accéder directement à un travail précis d'un mathématicien, indépendamment du fait que cet article provienne d'oeuvres complètes ou du périodique de première parution.

III. Ouvrages

1. Monographies

Un important programme coordonné entre Göttingen, Cornell et Ann Arbor et avec les soutiens d'agences nationales de la recherche (NSF et DFG) a été initié récemment, comme le décrit un collègue américain en ces termes :

Currently Cornell has a joint project with Goettingen & Michigan funded by the DFG & NSF to develop a system to distribute digitized & OCR'd math books. The digitization is funded separately. Cornell will contribute 500+ books, Goettingen something similar, & Michigan about 1000. The main items new in this collaboration is the development of the software for distribution, OCR of the books already scanned at Cornell, something similar in Goettingen I think, and the scanning & OCR of the 1000 books at Michigan. They have in fact just completed the first 400 (not yet available online). The development of the software has just started.

Les ressources numérisées en mathématiques dans Gallica sont estimées à quelques centaines d'ouvrages : elles concernent des mathématiciens français (certains tombés dans l'obscurité), mais aussi des mathématiciens étrangers (en traduction ou dans la langue originale) : Airy, Bachmann, Baltzer, Beltrami, Bianchi, Boole, Braunmühl, Brioschi, Cantor, Cartan, Clebsch, Cremona, Darboux, Descartes, Gauss, Hadamard, Legendre, Poisson, ...

Certains ouvrages sont numérisés en doublon comme p. ex. la monographie *Analyse infinitésimale des courbes planes* de B. Aoust, dans Gallica et à la bibliothèque de Cornell.

Il est donc important de donner de la visibilité, en coordination avec ces projets internationaux, aux monographies déjà numérisées d'une part, de compléter ces collections numérisées par d'autres opérations de numérisation suivant les standards techniques les plus fins (l'exigence des 600 dpi p. ex.) d'autre part.

2. Ouvrages d'un type particulier

Les éditions d'œuvres complètes ont été mentionnées ci-dessus

Des volumes contenant les actes de congrès internationaux de mathématiciens (Zürich 1897, Paris 1900, Heidelberg 1904, Rome 1908, Cambridge 1912) ont été numérisés. Un traitement spécifique est souhaitable.

IV. Le Répertoire bibliographique des sciences mathématiques

En 1885, la Société mathématique de France prend l'initiative de créer le répertoire Bibliographique des Sciences Mathématiques pour répertorier la littérature mathématique, projet acquérant rapidement une envergure internationale : pendant environ 27 ans, une cinquantaine de mathématiciens disséminés dans 16 pays différents dépouillèrent plus de 300 revues. Au final, entre 1894 à 1912, plus de 20 000 travaux mathématiques furent identifiés, répertoriés et présentés suivant un classement systématique et méthodique. D'autres initiatives bibliographiques ont marqué la fin du XIXe siècle, parmi lesquelles le *Jahrbuch über die Fortschritte der Mathematik* (1868-1940), qui est accessible aujourd'hui via l'interrogation d'une base de données.

Numériser ce répertoire en constituant une base de données donnerait une contribution notable aux projets de numérisation en France et dans le monde, un complément utile à la base de données issue du *Jahrbuch*.

Dernière modification : 4 juin 2001