



HAL
open science

La numérisation

Aurélie Brest, Angélique Malec, Clarisse Siméant, Elisabeth Kneller, Pauline Riviere

► **To cite this version:**

Aurélie Brest, Angélique Malec, Clarisse Siméant, Elisabeth Kneller, Pauline Riviere. La numérisation : un atout pour valoriser nos archives scientifiques. 2023. hal-04254610

HAL Id: hal-04254610

<https://universite-paris-saclay.hal.science/hal-04254610>

Submitted on 25 Oct 2023

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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

La numérisation : un atout pour valoriser nos archives scientifiques
Journée d'étude du Réseau Archives scientifiques de l'Université Paris-Saclay – Orsay, le 13 juin 2023
Aurélie Brest
Bibliothécaire et archiviste à l'Institut des hautes études scientifiques (Bures-sur-Yvette)

Ce compte rendu a été rédigé en collaboration avec :

- Elisabeth Kneller, responsable de la bibliothèque mathématique Jacques-Hadamard (université Paris-Saclay) / CNRS ;
- Angélique Malec, responsable de la bibliothèque de recherche à la faculté Jean-Monnet (université Paris-Saclay) ;
- Pauline Rivière, cheffe de projet à la bibliothèque Sainte-Geneviève (Paris) ;
- Clarisse Siméant, chercheuse en histoire du droit à la faculté Jean-Monnet (université Paris-Saclay).

Le 13 juin 2023, le réseau Archives scientifiques Paris-Saclay a organisé une journée d'études intitulée : La numérisation : un atout pour valoriser nos archives scientifiques.

Cette rencontre entre professionnels de l'information et enseignants-chercheurs répondait au besoin d'échanger autour des pratiques de numérisation et de valorisation d'archives scientifiques. Elle a mis l'accent sur la collaboration entre chercheurs et professionnels de l'information. Elle proposait aussi des pistes de réflexions pour celles et ceux qui portent des projets de numérisation d'archives scientifiques. La journée a rassemblé une cinquantaine de participants et a été suivie par une soixantaine de personnes en visioconférence.

[Thierry Doré - vice-président "Recherche et valorisation" à l'Université Paris-Saclay](#) - a ouvert la journée. Selon lui, la valorisation des archives scientifiques est un enjeu majeur pour la recherche. Les archives en ligne constituent une source d'information pour les chercheurs car elles complètent les publications officielles et permettent d'enrichir la recherche scientifique. C'est pourquoi la collaboration entre enseignants-chercheurs et archivistes est essentielle. Cette journée permettra de mieux comprendre les enjeux de la préservation et de la numérisation des archives scientifiques. Les enseignants-chercheurs en ont besoin pour faire l'histoire des sciences.

[Introduction : les archives scientifiques par Cyprien Henry – conservateur des bibliothèques MENJS \(Ministère de l'Éducation nationale, de la Jeunesse et des Sports\), et MESRI \(Ministère de l'enseignement supérieur, de la recherche et de l'innovation\).](#)

Cyprien Henry a proposé une courte histoire de la notion d'archives scientifiques en guise d'introduction. Il a souligné qu'il n'existait pas de consensus sur une définition de la science et que la notion d'archives était un terme ambigu qui change en fonction de différentes professions, même si elles sont définies par [l'article L.211.1 du code du patrimoine](#).

Selon le statut du producteur (service de l'administration centrale, organisme de recherche, laboratoire, etc.), le versement et la conservation des archives se feront dans des services d'archives différents : Archives nationales ou archives départementales. Pour ce qui concerne les archives des enseignants-chercheurs, la situation est plus complexe. En tant que producteur d'archives, ils considèrent qu'elles leur appartiennent et peuvent décider de les donner à des services très différents comme une bibliothèque, un musée, etc. Mais ces chercheurs sont souvent attachés à une institution publique qui peut, à juste titre, revendiquer la propriété de ces archives. Le développement de nouveaux flux de diffusion et de valorisation, ainsi que la prise en

compte des données de la recherche, mettent en évidence la nécessité d'un dialogue entre les enseignants-chercheurs, les archivistes, les bibliothécaires et les documentalistes.

Deux réalisations de projets de numérisation d'archives scientifiques ont ensuite été présentées.

[LaCAS](#) par Sarah Cadorel et Emma Bahous de l'[INALCO](#)

Sarah Cadorel et Emma Bahous de l'INALCO (Institut des national des langues et civilisations orientales) sont revenues sur la création de [LaCAS](#) (Open Archive in Language and Cultural Area Studies). Initié en 2019, le programme de numérisation patrimonial des archives d'enseignants-chercheurs de l'INALCO s'appuie sur le projet de recherche [D-PaRSAS](#) pour le développement de la plateforme LaCAS qui diffuse des productions scientifiques dans le domaine des études aréales. Une collaboration entre bibliothécaires, archivistes dont une apprentie, informaticiens et chercheurs a permis la réalisation du projet. La numérisation externe d'archives de chercheurs s'est justifiée pour pallier un problème de conservation ou pour leur intérêt scientifique. Après la numérisation, l'outil [Mynakala](#) a permis l'import par lots des fichiers sur LaCAS afin de publier, partager et valoriser les contenus. Ces archives étant pour la majorité des documents figurés, un panel d'outils juridiques a été créé pour mettre en œuvre la diffusion des contenus en collaboration avec les différentes parties prenantes. Enfin, ce processus de diffusion d'archives numérisées s'est inscrit dans une politique d'archivage plus globale et dans la politique de science ouverte de l'établissement. En valorisant ces archives scientifiques le [SISAP](#) (Service de l'information scientifique, des archives et du patrimoine) de l'INALCO espère créer un cercle vertueux pour l'archivage de la production scientifique.

[Numistra](#) : Traitement et valorisation du fonds photographique Charles Sadron, par Lucile Schirr, Delphine Issenmann et Nicolas Di Méo de l'Université de Strasbourg.

Les intervenants ont présenté la méthodologie utilisée pour la valorisation d'archives audiovisuelles qui a permis de bâtir plusieurs missions impliquant une collaboration entre les services de l'Université de Strasbourg et le CNRS (Centre national de la recherche scientifique), ainsi qu'une montée en compétences au sein de cette démarche pilote. L'institut a été fondé par Charles Sadron vers 1947, le laboratoire s'appuie sur le concept d'usine à recherches pour des physiciens, des chimistes et des biologistes autour de la recherche sur les macromolécules et les polymères.

En 2019, un projet de collaboration autour de la question de la mémoire orale et de la transmission de témoignages entre l'[Institut Charles Sadron](#) et le [Jardin des Sciences](#) fait émerger une volonté de mettre en valeur des entretiens oraux et un fonds photographique. Le Jardin des Sciences contacte alors le service des archives au sujet du classement du fonds photographique et le département des collections du service des bibliothèques universitaires en vue de la numérisation du fonds. Les trois services identifient leur complémentarité pour le traitement de ce fonds dans son intégralité : du classement à l'étude des droits liés en vue de la numérisation. La valorisation est assurée par deux vidéos sur la dimension internationale du laboratoire et son statut "d'usine à recherches" et une BD numérique impliquant également d'autres services de l'université et du CNRS (les délégués à la protection des données et les services juridiques).

Financements et partenariats : le rôle de [CollEx Persée](#) par Clément Oury – Directeur adjoint de la bibliothèque du Muséum d'histoire naturelle

Le Groupement d'intérêt scientifique (GIS) CollEx-Persée, présenté par Clément Oury, est une infrastructure de recherche en information scientifique et technique créée en 2017. Son objectif est de financer des établissements et de les aider à gérer des collections d'excellence. Le GIS CollEx-Persée collabore aussi avec d'autres partenaires comme l'[ABES](#) (Agence nationale de l'enseignement supérieur), le [CTLes](#) (Centre technique du livre de l'enseignement supérieur) et la [BNF](#) (Bibliothèque nationale de France) pour bâtir des services innovants. Les projets financés sont axés sur la numérisation enrichie ou sur la création de services de recherche. Il y a deux critères fondamentaux : la logique partenariale et l'implication des chercheurs ou des laboratoires de recherche.

Le GIS CollEx-Persée réfléchit aujourd'hui au programme "Archives scientifiques". Trois groupes de travail ont été formés pour élaborer ce programme et des établissements volontaires sont déjà identifiés. Les axes retenus sont : "la gouvernance du programme", "les évolutions professionnelles et organisationnelles des archives dans les bibliothèques" et "le signalement, la diffusion et l'exploitation des archives déjà conservées".

Une présentation au Conseil des membres de CollEx-Persée est prévue à l'automne pour décider du devenir du projet. Et courant 2024, une journée d'étude réunissant des professionnels de l'IST (information scientifique et technique) et des chercheurs devrait lancer officiellement le programme "Archives scientifiques" et clarifier les questions de terminologie.

Réaliser un cahier des charges et choisir le prestataire : l'exemple de NumaHop par Oriane Boissel, chargée de projets numériques et amélioration continue à l'Université Paris-Saclay

Dans le cadre du projet [NumaHOP](#), Oriane Boissel est revenue sur la rédaction d'un cahier des charges pour un projet de numérisation réalisée par un prestataire extérieur.

Ce document, structuré selon quatre grandes thématiques (contexte, objectifs de la numérisation, modalités logistiques, cadre réglementaire), permet de définir les attentes et les exigences d'un projet de numérisation. Après une description des éléments attendus, Oriane Boissel a expliqué la démarche dans le cadre du projet NumaHOP de l'Université Paris-Saclay.

NumaHop est un logiciel libre pour la gestion de la numérisation et l'archivage de documents. L'automatisation des contrôles et la conversion des données en formats interopérables, ainsi qu'une communauté d'utilisateurs active sont ses principaux atouts. Les limites du logiciel résident dans les constats d'état des documents et la description des fonds inadaptés aux spécificités des archives. La livraison directe des fichiers accélère le processus de contrôle qualité et facilite les échanges entre prestataire et commanditaire. Ces deux fonctionnalités ont permis d'alléger les prestations attendues dans le cahier des charges de Paris-Saclay.

Grâce à cette expérience, la direction des bibliothèques de l'Université Paris-Saclay propose aujourd'hui des services d'accompagnement de projets comme le soutien à la rédaction du cahier des charges, la formation et l'utilisation des outils de valorisation numérique comme NumaHop et [Omeka S](#).

5 exemples de valorisation de documents numérisés

[Numaclay](#) par Sylvie Sallé, chargée de projet au sein de la direction des bibliothèques de l'Université Paris-Saclay

Sylvie Sallé a présenté la bibliothèque numérique de l'Université Paris-Saclay : Numaclay. Mise en ligne en mars dernier, le principal objectif est de valoriser les patrimoines historiques, scientifiques et techniques de l'université. Numaclay est un site réalisé avec Omeka S. En plus de la mise à disposition d'outils tels que NumaHop et Omeka S, la direction des bibliothèques offre également des services de conseil, d'expertise et d'accompagnement, et de formation.

Numaclay propose plusieurs rubriques ayant chacune un intérêt différent. La rubrique "À la une" présente des actualités mensuelles thématiques pour mettre en valeur des documents que les utilisateurs ne consulteraient peut-être pas autrement. Un onglet "Collections" permet une découverte de documents spécifiques : archives scientifiques, ouvrages spécialisés, instruments scientifiques ou revues en histoire de l'enseignement, etc. La page d'accueil propose une sélection aléatoire de documents pour une exploration spontanée du site. Et une carte interactive est disponible pour explorer les différentes collections.

Une charte de politique documentaire est en cours de rédaction pour définir une ligne éditoriale encadrant la sélection des documents. Enfin, Numaclay prévoit de développer des expositions virtuelles. La première sera consacrée à Gustave Eiffel.

En conclusion, après une phase d'appropriation des différents outils disponibles, Numaclay se concentre désormais sur la construction de processus solides et de services adaptés pour accompagner les services et les laboratoires de l'université dans leur démarche de valorisation du patrimoine.

[Les archives de l'IHES](#) par Aurélie Brest, archiviste à l'IHES (Institut des hautes études scientifiques)

Aurélie Brest a présenté la bibliothèque numérique de l'IHES. Cet institut privé de recherche fondamentale en mathématiques et en physique théorique fait partie de l'Université Paris-Saclay. Le projet de bibliothèque numérique de l'IHES est le résultat de deux campagnes de numérisation de masse de ses archives, réalisées entre 2015 et 2018, et ayant pour objectifs de préserver le patrimoine immatériel de l'Institut, de mieux le connaître et de le faire connaître, ainsi que de participer à la diffusion de la culture scientifique.

Les archives de l'IHES se composent d'archives scientifiques, historiques et administratives, de fonds de chercheurs, d'une photothèque et de deux fonds privés. Elles ont été numérisées par un prestataire et une partie des documents est disponible sur la bibliothèque numérique réalisée avec Omeka Classic. La diffusion des contenus respecte les règles archivistiques et juridiques en vigueur.

Le projet a engendré d'autres idées. Une exposition s'est tenue à l'IHES en 2016 autour des documents de la première campagne de numérisation. Elle est aussi disponible sur le site des archives de l'IHES. Le service des archives a conseillé des chercheurs pour trier leurs propres archives ou réaliser des recherches historiques et scientifiques. L'Institut a reçu deux dons. Un livre sur les 60 ans de l'IHES. Et l'Institut a participé au projet AMOr CollEx-Persée2022 de la Bibliothèque Jacques Hadamard.

Prochainement, la bibliothèque numérique devrait migrer vers la base Omeka S de l'Université Paris-Saclay. L'import des archives sera géré avec NumaHop. Et une nouvelle version des "Archives de l'IHES" sera disponible sur Numaclay.

["D'or et de pixels](#) une collection de manuscrits médiévaux dans [Genovefa](#)" par Pauline Rivière, cheffe de projet numérisation et Antoine Boustany, chargé de la collection des manuscrits de la bibliothèque Sainte-Geneviève.

Antoine Boustany a expliqué que la bibliothèque numérique [Genovefa](#) avait été développée en interne sous Omeka S. L'équipe a aujourd'hui toutes les compétences pour développer et enrichir la bibliothèque numérique au gré des nouveaux projets. Techniquement, la bibliothèque repose sur le protocole [IIIF](#) ([International Image Interoperability Framework](#)) et elle est adossée à un serveur d'images IIIF [Cantaloupe](#). Le site d'or et de pixels est alimenté par import de métadonnées issues d'autres entrepôts de données ([Calames](#), [BVMM](#) etc.). Il a notamment pour objet de donner accès aux numérisations des 600 manuscrits médiévaux de la bibliothèque Sainte-Geneviève. La numérisation, menée en partenariat avec l'IRHT (Institut de recherche et d'histoire de textes, CNRS), est toujours en cours mais certains projets de valorisation sont déjà réalisés. L'exposition "D'or et de pixels", à la fois physique et virtuelle, en est un exemple.

Pauline Rivière a, quant à elle, montré, à travers 5 exemples, la diversité des usages permis aux publics de la bibliothèque à partir des numérisations et des services proposés par Genovefa.

[Bibliothèque numérique de Science Po](#) par Olesia Dubois, responsable du service de numérisation et d'archivage numérique à Sciences Po Paris.

Olesea Dubois a présenté les trois outils de diffusion et de valorisation de la bibliothèque numérique de l'institution. Ces outils ne se concentrent pas uniquement sur des documents classiques, mais plutôt sur des collections spécifiques, notamment les archives électorales. La bibliothèque numérique a été développée sous Omeka S, une plateforme de publication web dédiée aux collections du patrimoine culturel. En plus d'Omeka S, Sciences Po utilise Archimed Explorer pour regrouper et explorer les archives électorales, offrant une interface de recherche avancée. Les contenus sont également diffusés sur [Internet Archive](#) pour une meilleure visibilité mondiale. Des expositions virtuelles sont proposées, et des améliorations sont prévues, telles qu'un outil de transcription et une meilleure structuration des archives.

[SorbonNum](#) par Adeline Batailler, responsable adjointe du département des collections à Sorbonne université.

Adeline Batailler a évoqué le projet SorbonNum développé en collaboration avec le service des archives. Lancé en 2021, après un état des lieux sur une première bibliothèque numérique ancienne et en collaboration avec la bibliothèque et le service des archives, la bibliothèque numérique SorbonNum a été conçue comme le point d'entrée à toutes les collections scientifiques, documentaires et archivistiques. Les documents de SorbonNum sont archivés au [CINES](#) (Centre informatique national de l'enseignement supérieur) pour garantir un accès pérenne. Ils sont libres de droit et librement réutilisables.

Depuis, l'infrastructure a été repensée pour moderniser et mieux organiser les contenus. SorbonNum est interopérable avec NumaHOP. Grâce à cela, certaines collections ont été remaniées et cataloguées au format Dublin Core pour créer des facettes et des notices de collections détaillées.

A l'avenir, SorbonNum devrait être moissonnée par [France archives](#) et [Archives Portal Europe](#). Les collections seront scindées en 3 ensembles : collections documentaires, archives et collections scientifiques. Enfin, il est prévu d'intégrer les liens IIF dans les notices existantes. Des expositions virtuelles et la communication sur les réseaux sociaux seront proposées aux publics pour développer de nouveaux services de valorisation des contenus.

Les projets autour de SorbonNum montrent une volonté de développer une bibliothèque numérique utile au plus grand nombre.

Table ronde : la numérisation des archives scientifiques : un atout pour la recherche ?

Modérée par Pauline Lemaigre-Gaffier de l'UVSQ (Université Versailles Saint-Quentin) - Université Paris-Saclay

Autour de Laurent Rollet du laboratoire Archives Poincaré, Clarisse Siméant, Angélique Malec de l'Université Paris-Saclay et Elisabeth Kneller du CNRS.

Pour explorer l'intérêt de numériser des archives scientifiques pour les mettre au service de la recherche actuelle, trois projets ont été évoqués. Ils sont le fruit de rencontres souvent interindividuelles entre plusieurs acteurs : des professionnels de l'information, managers de projet, et des chercheurs pour les ressorts scientifiques.

*Les « [Archives scientifiques de Jean Gaudemet](#) » sont constituées de fiches de travail qui formaient le prélude aux articles scientifiques de ce chercheur. Ces documents se présentent sous la forme de feuillets A5 manuscrits ordonnés dans un plan de classement thématique et chronologique utilisé tout au long de sa carrière. Ce projet de numérisation dont l'objectif est de révéler la méthodologie d'un chercheur du XXe siècle, à travers un échantillon de 500 feuillets, a été porté par le laboratoire Droit et Sociétés religieuses en partenariat avec la DiBISO.

*Le deuxième projet est consacré aux [procès-verbaux des réunions du Bureau des longitudes](#).

Né en 1795, le Bureau des longitudes est une académie scientifique. Sœur de l'Académie des sciences, elle a vu ses missions s'enrichir au fil du temps et a accueilli de grands savants et des personnalités politiques de premier plan. La numérisation des procès-verbaux des réunions hebdomadaires tenues entre 1795 et 1932 met en valeur la vie de l'institution. Le projet réalisé par le laboratoire Archives Poincaré en partenariat avec la

bibliothèque de l'Observatoire de Paris et l'Université PSL a permis la mise en ligne d'une vingtaine de registres de réunions, soit 22 000 fichiers images.

* Le [projet ColLEx-Persée AMOr](#) des [Archives Mathématiques d'Orsay](#), mené par la Bibliothèque mathématique Jacques Hadamard en collaboration avec plusieurs partenaires, a pour ambition de mettre à disposition de la communauté scientifique des archives de la recherche mathématique provenant du Laboratoire de Mathématiques d'Orsay. Il s'agit, d'un côté, de créer un inventaire de ces archives pour les rendre plus facilement utilisables pour la recherche historique et mathématique. De l'autre côté, il s'agissait de les numériser et éditorialiser pour les rendre exploitables sous forme numérique. Une équipe projet a développé l'édition de formules mathématiques dans un éditeur XML/TEI. En s'appuyant sur des documents numérisés de l'IHES, une autre équipe a travaillé sur des modèles d'apprentissage (OCR et HTR) pour la transcription d'objets mathématiques en code LaTeX.

Ces projets visent à mettre en valeur la méthodologie d'un chercheur de sciences humaines, l'histoire d'une académie scientifique et la recherche de l'institut mathématique d'Orsay. Cette numérisation peut être le matériau pour de nouvelles recherches, comme en témoigne le projet AMOr qui a fourni ainsi de nouvelles données pour les outils liés à l'intelligence artificielle.

La réalisation de ces projets a nécessité une connaissance fine des environnements institutionnels locaux, ainsi qu'une recherche de collaborations interinstitutionnelles. En outre, il a été indispensable au préalable de formaliser les licences les plus ouvertes possibles avec les auteurs ou les ayants-droits. Ces étapes nécessitent un temps long souvent en décalage avec des postes non pérennes liés à ces tâches.

La visée scientifique de ces projets permet de distinguer cette opération de numérisation d'une mise en ligne d'archives par des services d'archives nationales ou départementales évitant ainsi une éventuelle concurrence entre établissements. La diffusion numérique de ces archives scientifiques ne fait pas l'économie d'un questionnement sur la pertinence de l'archivage pérenne. Au final, outre de potentielles nouvelles pistes de recherche, ces collections numérisées participent à la notoriété des laboratoires de recherche.

[Conclusion de la journée par Julien Sempéré, directeur de la DIBISO de l'Université Paris-Saclay](#)

En conclusion, Julien Sempéré a souligné l'intérêt croissant porté aux archives au sein de l'université, notamment en termes de diffusion et de valorisation. L'archivage scientifique s'inscrit dans le domaine de la science ouverte. C'est grâce à la diffusion des archives scientifiques que les chercheurs pourront les utiliser pour faire avancer leurs domaines de recherche.

Sur le plan juridique, il existe des enjeux mais ils ne doivent pas freiner le développement de services utiles aux chercheurs ni la création de communautés de pratiques susceptibles de faire évoluer certains champs disciplinaires.

La collaboration interdisciplinaire fait avancer les projets et permet de construire des outils communs. S'il est vrai qu'il n'existe pas de service d'archives à l'Université Paris-Saclay, il y a des compétences chez les bibliothécaires et une sensibilisation croissante des chercheurs sur la nature archivistique de leur production scientifique en tant que patrimoine vivant.