

Transformations des métiers de l'information-documentation au prisme des humanités numériques : Entretien avec Jean-Philippe Moreux

Sokaina Al Morabit Emma Lavallée
Jean-Philippe Moreux

2025/04/23

Résumé

Entretien réalisé le 23 avril 2025 par les étudiants du Master MAVINUM
(AL MORABIT Sokaina et LAVALLÉE EMMA)

Mot-clés : Humanités numériques, DataLab, Médiation numérique, SHS,
Intelligence artificielle

Introduction

Ce billet prend la forme d'un entretien écrit sur les transformations des métiers de l'information-documentation au prisme des humanités numériques avec Jean-Philippe Moreux, en collaboration avec Emma Lavallée et Sokaina Al Morabit.

Jean-Philippe MOREUX, ingénieur diplômé de l'INSA Toulouse (informatique, 1990), est le chef de mission IA de la Bibliothèque nationale de France. Il a été auparavant l'expert scientifique de la bibliothèque numérique Gallica au département de la Coopération et l'expert OCR et formats éditoriaux du service de la Numérisation. Précédemment, il a exercé les métiers de chef de projet informatique, responsable éditorial et consultant (ingénierie éditoriale, édition scientifique). Il travaille sur l'application de l'intelligence artificielle dans le contexte du patrimoine, les programmes de valorisation du patrimoine numérique et participe à des projets de recherche sur ces sujets. Ses principaux sujets d'activité et de recherche sont la production et la valorisation des ressources patrimoniales numériques, les services aux humanités numériques, l'ingénierie de la donnée et du document (Source : BnF)

Emma Lavallée est diplômée de l'université de Limoges avec une Licence d'Histoire et une Licence Professionnelle Design de l'Information et Rédaction Technique. Elle poursuit actuellement le Master 2 MAVINUM à l'université de Montpellier Paul-Valéry en alternance. Elle réalise cette alternance à Safran Landing Systems, entreprise d'aéronautique, dans le service de documentation technique.

Sokaina Al Morabit est étudiante en Master 2 MAVINUM à l'Université Paul-Valéry Montpellier et documentaliste en alternance à la bibliothèque de TBS Education à Toulouse. Ses missions portent sur l'acculturation des étudiants aux compétences informationnelles, via la conception de contenus pédagogiques sur le LMS Moodle, ainsi que sur le développement de la science ouverte, à travers la création d'une rubrique dédiée sur le site de la bibliothèque et la participation à l'organisation d'ateliers pour les chercheurs. Sokaina Al Morabit s'est orientée vers le domaine de l'information-documentation à la suite d'un Master recherche Langues et Sociétés à l'Université Toulouse 2 Jean-Jaurès, où son travail de recherche l'a amenée à s'intéresser aux techniques documentaires et au monde des bibliothèques.

L'entretien avec Jean-Philippe MOREUX

Pour commencer, pourriez-vous présenter brièvement votre parcours professionnel et décrire vos principales missions au sein de la Bibliothèque nationale de France (BnF), notamment autour de Gallica et du BnF DataLab ?

Moi j'ai une formation initiale d'ingénieur en informatique. Après j'ai travaillé un petit peu dans le logiciel, et après je me suis tourné vers le secteur de l'édition, l'édition scientifique, donc à cheval entre l'édition scientifique, technique et puis un peu après l'édition numérique.

Et plus récemment, il y a dix ans, j'ai basculé du côté des bibliothèques. Donc c'est un parcours centré sur le document. À l'origine, mon premier poste dans l'informatique, c'était la gestion électronique de documents. Donc le fil, c'est le document.

Et aujourd'hui, à la BnF, après avoir longtemps travaillé auprès de la bibliothèque numérique Gallica, depuis plus récemment, deux ou trois années, je me suis occupé beaucoup d'innovation, de recherche-développement autour de l'intelligence artificielle dans le patrimoine.

Donc mon poste aujourd'hui, il est plutôt sur l'accompagnement à la transition vers l'innovation numérique, le lien pour le patrimoine. Et auparavant, dans le précédent poste côté Gallica, encore aujourd'hui, il y a beaucoup d'interactions bien sûr avec les humanités numériques, au sens où les chercheurs sont la principale catégorie d'utilisateurs de la BnF.

Donc effectivement, moi je n'ai pas un parcours classique en info-doc. Je n'ai pas de formation initiale dans les métiers de bibliothèque. Mais je travaille ici depuis douze ans.

Comment définiriez-vous les humanités numériques, et comment cette notion s'est-elle transformée ou enrichie, selon vous, au cours de ces dernières années ?

Écoutez, les humanités numériques, c'est une manière d'envisager la recherche en sciences humaines et sociales en utilisant un outillage... enfin, un outillage numérique qui peut être très divers.

C'est à la fois, je pense, une communauté, une méthodologie de travail, qui s'appuie aussi, bien sûr, sur des bibliothèques d'outils ou des approches techniques, qui relèvent de l'informatique au sens large. Sous ce terme assez générique, on peut mettre beaucoup de choses. Nous, à la BnF, on se positionne en tant que partenaire d'équipes de recherche, ou de chercheurs, qui utilisent ces approches, qui s'inscrivent dans cette communauté, dans cette dynamique.

On fait de la recherche, mais la BnF, c'est une bibliothèque de recherche. Mais quand on parle d'humanités numériques ici, on parle vraiment de services qu'on peut rendre ou de l'accompagnement à des chercheurs, ou des laboratoires de recherche, ou des équipes de recherche, qui ont ce type de pratique. Ce qui a beaucoup changé, je dirais depuis une dizaine d'années, c'est la démocratisation de l'outillage technique, la montée en puissance, en performance, des approches techniques permises, y compris par l'IA.

Donc c'est cette démocratisation qui a permis de constater qu'on avait plus de demandes d'étudiants, de chercheurs, qui se sont engagés pour leur propre activité, pour leurs propres objectifs scientifiques. Ce qu'on fait avec eux aujourd'hui ou ce qu'on faisait il y a cinq ou dix ans, les objectifs restent les mêmes : accompagner, servir des objectifs scientifiques.

Les méthodes sont peu ou prou restées les mêmes, par contre l'outillage a beaucoup changé. Les chercheurs sont plus autonomes aujourd'hui qu'ils ne l'étaient il y a trois ou cinq ans, parce que la technologie est plus performante, parce que les outils se sont démocratisés, et parce qu'il y a aussi plus de pratiquants. Il y a une sorte de communauté qui s'est construite, en France et ailleurs.

Comment les outils et approches des humanités numériques transforment-ils la façon dont vous valorisez et diffusez le patrimoine culturel auprès du public et des chercheurs ?

On s'inscrit dans cette dynamique parce qu'il y avait une demande de la part de nos usagers, une demande renforcée d'accès aux collections numérisées. Pour qu'ils soient en capacité, eux, de travailler avec ces collections, nous, on a été amenés

à ouvrir, par exemple, des API. À s'engager dans une dynamique d'ouverture des données, qui est aussi promue par l'État au sens large.

C'est donc cette interaction, cet échange, en termes de demande adressée à la bibliothèque, qui nous a amenés à faire évoluer à la fois la manière dont on envisage les collections lorsqu'elles sont numériques ou numérisées, et puis les portails d'information, les outils qui s'appuient sur ces collections. C'est aussi une réponse de la BnF à des demandes exprimées par ses usagers.

Ce n'est pas qu'une question de demande, c'est aussi l'intérêt de travailler de manière transdisciplinaire avec des usagers ou avec des partenaires, sur des projets concrets qui s'appuient sur les collections. Ça permet de valoriser les collections, pas uniquement en les décrivant dans des catalogues ou en les diffusant sur des portails comme Gallica, mais ça permet d'élargir les usages qui sont faits de ces collections, puisqu'elles redeviennent vivantes dans la manière dont les chercheurs les appréhendent, les réutilisent pour leurs propres objectifs.

C'est une nouvelle manière de faire vivre, de communiquer, de faire connaître ces collections. C'est intéressant pour toutes les parties prenantes.

Comment décririez-vous votre posture professionnelle face à l'essor des humanités numériques ? Qu'est-ce qui caractérise, selon vous, le « nouveau » rôle du professionnel de l'information-documentation dans ce contexte ?

Je pense qu'il y a quelques années encore, c'était encore relativement nouveau, c'est-à-dire qu'on découvrait cette posture. Aujourd'hui, elle s'est un peu imposée, c'est-à-dire que l'accompagnement de chercheurs dans des pratiques techniques, spécifiques, nouvelles, qui se renouvellent très vite, c'est devenu un peu une mission classique dans une bibliothèque comme la BnF.

Le rôle du professionnel de l'info-doc a un peu évolué, dans le sens où auparavant on orientait les chercheurs dans les catalogues, vers des outils de recherche documentaire.

Aujourd'hui, il faut qu'on soit en capacité de comprendre ce que veulent faire les chercheurs avec les collections, et de leur fournir un accompagnement, qui est à la fois technique, ou bien qui consiste à identifier des sources, ou à les orienter vers des outils, vers des plateformes. Ça fait apparaître aussi des nouveaux profils dans les bibliothèques, c'est-à-dire des gens qui sont un peu à la croisée de plusieurs mondes : de l'informatique, du patrimoine, de l'enseignement supérieur.

On recrute aujourd'hui dans des bibliothèques, nationales ou universitaires, des ingénieurs d'étude ou des ingénieurs de recherche, qui sont capables de dialoguer avec des chercheurs. C'est donc un peu une recomposition des équipes dans les bibliothèques, pour certaines d'entre elles en tout cas.

Sur quels types de projets d'innovation travaillez-vous actuellement, ou prévoyez-vous de travailler, et comment s'inscrivent-ils dans l'évolution plus globale du champ des humanités numériques ?

Il y a un ensemble de projets qui relèvent des missions classiques de la bibliothèque : numériser, cataloguer, diffuser des collections, par exemple à travers Gallica. Et il y a un autre ensemble de projets qui s'appuient plutôt sur une logique de collaboration avec le monde académique. Depuis cinq ou six ans, on participe à des projets transdisciplinaires qui associent à la fois des laboratoires en sciences humaines, des chercheurs en informatique, et puis une bibliothèque patrimoniale comme la BnF.

Ces projets peuvent porter sur des collections spécifiques, par exemple la presse ancienne, les documents iconographiques, les estampes, les manuscrits, etc.

L'intérêt, c'est que chacun y trouve un bénéfice. Le laboratoire en SHS peut faire avancer ses hypothèses ou sa méthode. Le laboratoire en informatique a un corpus à tester. Et la bibliothèque, elle, peut créer des outils, des prototypes qui lui seront utiles ensuite pour ses propres missions. Donc ce sont des projets qui enrichissent toutes les parties.

C'est aussi un moyen pour nous de tester des technologies, d'identifier des besoins, d'expérimenter avec des chercheurs, tout en gardant une maîtrise des usages.

Ce sont des projets qui ne sont pas déconnectés de la bibliothèque : les corpus sont bien les nôtres, les chercheurs sont bien ceux que l'on accueille ou que l'on accompagne, et les outils développés sont pensés pour être utiles à terme.

Quelles sont, selon vous, les transformations majeures qui ont eu lieu ces dernières années sous l'impulsion des humanités numériques, et en quoi ont-elles modifié vos méthodes de travail ?

Je pense que ce qui a le plus changé, c'est le niveau d'exigence des chercheurs. Il y a cinq ou dix ans, les chercheurs qui s'intéressaient aux humanités numériques étaient plutôt des pionniers, des gens très au fait, qui venaient nous voir avec des demandes assez exploratoires.

Aujourd'hui, ce n'est plus le cas. Ce sont des chercheurs qui savent ce qu'ils veulent faire, qui ont souvent les compétences techniques, ou bien qui ont dans leur équipe des personnes capables de développer.

Ils veulent donc utiliser des corpus, des API, des formats de données bien structurés, des jeux de données prêts à l'emploi. Nous, en tant que bibliothèque, ça nous oblige à faire évoluer nos services, à revoir nos systèmes d'information.

Ce n'est pas uniquement le DataLab qui doit répondre à ces demandes : c'est toute la bibliothèque, tout son écosystème technique, qui doit évoluer. On ne peut

pas avoir d'un côté un laboratoire innovant, très performant, et de l'autre des systèmes qui datent des années 90 ou 2000. Donc on essaie d'articuler innovation et fonctionnement quotidien. Il faut que l'ensemble suive.

Ce sont des transformations profondes, qui concernent autant les outils que les méthodes de travail, et même les métiers dans la bibliothèque.

Les projets en humanités numériques sont souvent interdisciplinaires. Pourriez-vous nous donner des exemples concrets de collaborations (avec des chercheurs, des développeurs, des conservateurs, etc.) et expliquer comment votre rôle s'articule dans ces dynamiques de travail collaboratif? Avez-vous déjà rencontré des difficultés ou fait face à des défis à ce niveau ?

Alors oui, chaque année par exemple, le DataLab, qui est un dispositif d'accueil à la BnF pour les chercheurs en humanités numériques, lance un appel à projets.

On sélectionne chaque année entre quatre et six projets, qui sont portés par des chercheurs ou des équipes de recherche. Et ce sont des projets très variés, qui peuvent aller de la littérature à l'histoire de la publicité, des arts du spectacle aux sciences de l'information.

À chaque fois, ces projets impliquent des corpus issus des collections de la BnF, et une équipe de chercheurs qui veut les explorer avec des méthodes numériques.

De notre côté, on constitue une équipe projet en interne, avec des conservateurs, des informaticiens, des responsables de données, etc. On met en place des réunions régulières, on accompagne le projet pendant un an. Ce sont donc des collaborations concrètes, où chacun apporte son expertise.

Les difficultés, on les rencontre surtout sur la question des temporalités. Un chercheur ou un postdoctorant travaille sur un temps court, un an, deux ans, trois ans au maximum.

La BnF, elle, est sur des cycles beaucoup plus longs, cinq, dix ans parfois. Donc il faut synchroniser ces temporalités.

L'autre difficulté, c'est la pérennisation. Comment on fait vivre un outil, un prototype, une méthode développée dans le cadre du projet, une fois que le projet est fini ?

C'est un vrai défi, parce que la BnF n'est pas une entreprise, ce n'est pas une start-up. Intégrer de l'innovation dans une structure publique patrimoniale, ça demande du temps, de l'organisation, des moyens. Mais c'est aussi ce qui rend le travail intéressant.

Selon vous, les professionnel·les de l'information-documentation (archivistes, documentalistes, etc.) sont-ils aujourd'hui suffisamment formé·es aux humanités numériques, ou constatez-vous encore des lacunes dans ce domaine ? Quelles compétences clés (techniques, scientifiques, communicationnelles, etc.) devraient être développées pour répondre aux nouveaux besoins induits par le numérique ?

Selon vous, les professionnel·les de l'information-documentation (archivistes, documentalistes, etc.) sont-ils aujourd'hui suffisamment formé·es aux humanités numériques, ou constatez-vous encore des lacunes dans ce domaine ? Quelles compétences clés (techniques, scientifiques, communicationnelles, etc.) devraient être développées pour répondre aux nouveaux besoins induits par le numérique ?

Ça dépend vraiment beaucoup des filières de formation. Je connais un peu les masters en sciences de l'information, l'ENSSIB, ce genre de formations.

Aujourd'hui, dans n'importe quelle bibliothèque, même une bibliothèque de lecture publique, vous êtes confronté·e au numérique. Il faut savoir que cela dépasse très largement la question des catalogues ou même de Gallica. Donc, ces formations intègrent désormais le numérique.

Mais sur la question plus spécifique des humanités numériques, c'est-à-dire la manière dont les chercheurs vont mobiliser les collections patrimoniales avec des outils d'analyse, d'extraction, de fouille, etc., il y a encore beaucoup à faire, notamment du côté de l'intelligence artificielle.

Aujourd'hui, on reçoit des projets au DataLab : 90 % des projets qui nous sont proposés utilisent des outils d'IA, que ce soit pour faire de la reconnaissance d'entités nommées, de la classification, de l'analyse d'images ou du traitement du langage.

Donc, si on veut pouvoir répondre à ces projets, il faut qu'on comprenne ce que les chercheurs veulent faire. Et pour cela, il faut que les professionnels soient formés, pas forcément pour développer eux-mêmes, mais pour comprendre, pour dialoguer, pour accompagner.

Donc oui, il y a encore un effort de formation à faire, notamment sur ces outils récents et sur les enjeux qu'ils posent.

Quels sont les principaux défis techniques et méthodologiques que vous rencontrez actuellement dans l'exploitation des données ou des corpus pour la recherche en humanités numériques ? La pérennité et l'interopérabilité des données, ainsi que le développement rapide des technologies, font-ils partie de ces défis ?

Oui, absolument. Ce sont des défis majeurs aujourd'hui. Les chercheurs, comme je le disais, sont de plus en plus exigeants, avec des outils de plus en plus puissants, des attentes très précises. Donc, ça veut dire que nous, de notre côté, on doit faire évoluer notre offre de services, notre environnement technique et nos systèmes d'information. Et ce ne sont pas que des questions d'innovation dans un coin de la bibliothèque, comme au DataLab : ce sont des questions qui doivent concerner toute la structure.

Il faut faire en sorte que ce ne soit pas seulement quelques ingénieurs ou quelques bibliothécaires innovants qui soient en capacité de répondre à ces besoins, mais que toute l'institution monte en compétence. Sur les données, les défis sont clairs : qualité, structuration, interopérabilité.

Il faut que les données soient utilisables, bien documentées, compatibles avec les outils des chercheurs. Et puis, il y a aussi la question de la pérennité. Beaucoup d'outils ou de plateformes sont développés dans le cadre de projets de recherche, mais une fois le financement terminé : qui maintient ? Qui héberge ? Qui réutilise ?

Là aussi, les bibliothèques ont un rôle à jouer, à condition d'avoir les ressources et l'appui institutionnel pour le faire.

Pensez-vous que les outils technologiques sont pleinement exploités et optimisés au profit des humanités à l'heure actuelle ? Ou bien leur utilisation reste-t-elle encore limitée ou inégalement généralisée ?

Pensez-vous que les outils technologiques sont pleinement exploités et optimisés au profit des humanités à l'heure actuelle ? Ou bien leur utilisation reste-t-elle encore limitée ou inégalement généralisée ?

Non, je pense que l'utilisation reste encore inégalement généralisée.

Il y a des équipes très structurées, avec des moyens, qui sont capables d'utiliser des outils très puissants. Et puis, à côté de ça, il y a des chercheurs isolés ou des laboratoires qui manquent de moyens, qui redéveloppent parfois des outils qui existent déjà ailleurs, parce qu'ils n'en ont pas connaissance ou parce qu'ils n'y ont pas accès.

Il y a donc un vrai problème de mutualisation, de capitalisation. La communauté des humanités numériques est très dynamique, mais elle est encore jeune, elle fonctionne beaucoup en silos. Il manque parfois des infrastructures pérennes.

Huma-Num, par exemple, joue un rôle important pour mutualiser les outils, mais ses moyens sont limités ; ce n'est pas suffisant pour structurer l'ensemble du champ.

Donc oui, les outils technologiques ne sont pas encore pleinement exploités. Il y a un enjeu à mieux les faire connaître, à mieux les partager, à mieux les maintenir aussi dans la durée.

Pour conclure, quels conseils donneriez-vous à un · e futur · e professionnel · le de l'information-documentation souhaitant s'impliquer dans les humanités numériques ? Et, d'après votre expérience, quels sont les principaux impacts que les humanités numériques ont sur la profession aujourd'hui et à l'avenir ?

Je dirais que si c'est un domaine qui l'intéresse, il faut d'abord trouver un établissement qui permet de faire ça.

Ce n'est pas le cas partout. Dans certaines bibliothèques de lecture publique, vous n'aurez pas de projets en humanités numériques.

En revanche, dans les bibliothèques universitaires, les bibliothèques de recherche, les bibliothèques patrimoniales, là, oui, il y a une vraie demande.

Ensuite, il faut se former. Il existe des formations hybrides, des masters croisés entre info-doc, numérique, sciences humaines.

Et puis il faut s'impliquer dans des projets concrets, en stage ou en début de carrière, pour comprendre comment ça fonctionne, pour monter en compétence. Il n'y a pas de profil unique. Certains professionnels viennent de l'informatique, d'autres des sciences humaines, d'autres encore du patrimoine. Ce qui compte, c'est la capacité à comprendre les enjeux du numérique, à dialoguer avec les chercheurs, à construire des services.

Pas besoin d'être développeur, mais il faut pouvoir se repérer dans un environnement technique complexe, et avoir envie de travailler avec des chercheurs sur des projets qui font bouger les lignes.