

Diplôme de conservateur de bibliothèque

Mémoire d'étude / Mars 2026

Intelligence artificielle et signalement en bibliothèque : quelles évolutions pour les institutions, les agents et les chaînes de traitement ?

Octave BOCZKOWSKI

Sous la direction d'Emmanuel Jaslier

Directeur du département des Métadonnées, BnF

Remerciements

J'adresse mes remerciements les plus chaleureux à mon directeur de mémoire, Emmanuel Jaslier, qui m'a prodigué son temps et ses conseils avec largesse dès les premières étapes de ma réflexion, traçant avec sûreté les voies qu'elle a empruntées. Ce fut une chance et un plaisir de pouvoir discuter de ce travail à de nombreuses reprises avec le comité de suivi réuni par ses soins, dont je remercie les deux autres membres – Étienne Cavalié et Cécile Queffélec – avec tout autant de vivacité. Le présent mémoire leur doit à tous trois ce qu'il a de moins imparfait ; je garde la seule responsabilité du reste.

Je suis plein de reconnaissance envers toutes les personnes qui ont accepté de s'entretenir avec moi de l'automatisation du signalement, de Singapour jusqu'à Washington. Que ces discussions aient porté sur une institution individuelle ou sur l'ensemble des bibliothèques engagées dans une telle transformation, elles m'ont permis d'entrevoir, bien mieux qu'au travers de la littérature professionnelle ou des rapports d'expérimentation, les enjeux humains et organisationnels sous-jacents. J'espère que la richesse des idées partagées par ces professionnels pionniers se reflète quelque peu dans les pages qu'on va lire.

Merci encore à l'Enssib de m'avoir permis d'entreprendre la présente étude et de m'avoir offert le cadre intellectuel pour la mener à bien.

Que toute ma promotion trouve ici l'expression de ma gratitude pour sa camaraderie chaleureuse et stimulante. Merci à celles et ceux de mes collègues qui m'ont suggéré des références, et notamment Émilie et Sophie : la première pour ses réflexions sur l'avenir des métadonnées et la seconde pour son incitation à mener des lectures en économie et en sociologie.

Et sur un tout autre plan enfin, merci à Eleni pour sa présence radieuse et sa joie sans borne.

Résumé:

Face au poids croissant de la masse documentaire, les agences bibliographiques nationales se tournent de plus en plus vers l'IA pour accomplir leurs missions catalographiques. Le présent mémoire analyse ces modalités nouvelles de signalement, ce qu'une telle voie exige des institutions qui s'y engagent et le devenir du travail des agents concernés. Il s'appuie sur une revue de littérature et des entretiens avec des institutions pionnières pour proposer une typologie de chaînes allant de l'assistance pour la rédaction de notices une à une à des dispositifs d'indexation semi-automatisés, contrôlés par échantillonnage, qui traitent des centaines de milliers de titres par an. L'enquête montre que l'expertise catalographique demeure centrale, mais qu'elle se déplace dans les projets mûrs vers la relecture et la validation, ainsi que l'annotation de corpus à des fins de pilotage. Le mémoire insiste sur les conditions de réussite observées : ressources, temps, compétences internes, mobilisation des agents.

Descripteurs: Bibliothèques -- Informatique ; Indexation automatique ; Catalogage informatisé

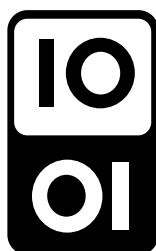
Abstract:

Given the growing scale of documentary production, national bibliographic agencies are increasingly turning to AI to carry out their cataloguing missions. This thesis analyses these new modes of library bibliographic control, what such a route demands from the institutions that pursue it, and how the work of the staff involved is likely to evolve. It draws on a literature review and interviews with pioneering institutions to offer a typology of workflows ranging from assistance in drafting records one at a time to semi-automated indexing systems—monitored through sampling—that process hundreds of thousands of titles yearly. The study shows that cataloguing expertise remains central, but in mature projects it shifts toward reviewing and validating output, as well as annotating corpora for steering and oversight. The thesis highlights the observed conditions for success: resources, time, in-house skills, and staff mobilization.

Keywords: Information services ; Artificial intelligence ; Cataloging

Droits d'auteurs

*Ce document est mis à disposition sous « Licence Ouverte / Open Licence » v2.0 (Etalab) (SPDX : etalab-2.0).
Texte de la licence : <https://www.data.gouv.fr/pages/legal/licences/etalab-2.0>*



SOMMAIRE

SOMMAIRE	5
SIGLES ET ABRÉVIATIONS	7
INTRODUCTION	9
ADOPTION INSTITUTIONNELLE DE L'IA : QUELS DÉFIS ?	17
QUELLES PERSPECTIVES POUR L'EMPLOI DES AGENTS CONCERNÉS ?	23
UNE PERMANENCE INDISPENSABLE DE L'EXPERTISE CATALOGRAPHIQUE	27
DU CATALOGUEUR PRODUCTEUR AU CATALOGUEUR RELECTEUR : UNE TRANSFORMATION AMBIVALENTE	35
FORMATION ET PROFILS TECHNIQUES	44
CONCLUSION	51
SOURCES	54
GLOSSAIRE	76
INDEX	77
TABLE DES MATIÈRES	81

SIGLES ET ABRÉVIATIONS

ADBU Association des directeurs et personnels de direction des bibliothèques universitaires et de la documentation (France)
AI4LAM *Artificial Intelligence for Libraries, Archives and Museums*
API *Application Programming Interface*
Abes Agence bibliographique de l'enseignement supérieur (France)
BERT *Bidirectional Encoder Representations from Transformers*
BnF Bibliothèque nationale de France
CENL *Conference of European National Librarians*
DNB *Deutsche Nationalbibliothek* (Bibliothèque nationale d'Allemagne)
ECD *Exploring Computational Description* (projet de la Bibliothèque du Congrès)
FAST *Faceted Application of Subject Terminology*
GLAM *Galleries, Libraries, Archives and Museums*
GND *Gemeinsame Normdatei* (fichier d'autorités national allemand)
GPEC Gestion prévisionnelle des emplois et des compétences
GPT *Generative Pre-trained Transformer*
IA Intelligence artificielle
IFLA *International Federation of Library Associations and Institutions*
KB *Koninklijke Bibliotheek* (Bibliothèque nationale des Pays-Bas)
KBR Bibliothèque royale de Belgique
LC *Library of Congress* (Bibliothèque du Congrès des États-Unis d'Amérique)
LC Labs Structure d'innovation de la Bibliothèque du Congrès
LCSH *Library of Congress Subject Headings*
LIBER *Association of European Research Libraries*
LLM *Large Language Model* (modèle de langage de grande taille)
MARC *Machine-Readable Cataloging* (famille de formats catalographiques)
MIT *Massachusetts Institute of Technology*
ML *Machine Learning* (apprentissage machine)
MMU *Manchester Metropolitan University*
MNHN Muséum national d'Histoire naturelle
NIST *National Institute of Standards and Technology* (États-Unis)
NLB *National Library Board* (Singapour)
OCLC *Online Computer Library Center*
OCR *Optical Character Recognition* (reconnaissance optique de caractères)
PCC *Program for Cooperative Cataloging*
RAG *Retrieval-Augmented Generation*
RAMEAU Répertoire d'autorité-matière encyclopédique et alphabétique unifié
RDA *Resource Description and Access*
RDF *Resource Description Framework*
RSC *RDA Steering Committee*
SCD Service commun de la documentation
SGCAT Outil de catalogage assisté par LLM (*National Library Board*, Singapour)
SIGB Système intégré de gestion de bibliothèque
SLM *Small Language Model* (LLM dont le nombre de paramètres est comparativement jugé peu élevé)
SRU *Search/Retrieve via URL* (protocole de recherche)
SUDOC Système universitaire de documentation, catalogue collectif des bibliothèques universitaires françaises
UNIMARC *Universal Machine-Readable Cataloging*
UX *User Experience* (expérience utilisateur)
ZBW *Leibniz Information Centre for Economics* (Allemagne)

INTRODUCTION

Les notices formant le catalogue d'une bibliothèque permettent d'identifier et de chercher les documents physiques ou digitaux qu'elle propose. On appellera *métadonnées* toutes les informations consignées dans ces notices, en distinguant les *métadonnées descriptives* qui « servent à décrire une ressource à des fins de compréhension, découverte ou identification du contenu » (comme le titre d'un ouvrage ou les noms de ses auteurs) des *métadonnées sémantiques*, celles qui sont « conformes à un vocabulaire standardisé ou une autorité de contrôle » tels qu'un vocabulaire contrôlé¹. Le terme de *descripteurs*² aurait aussi convenu. Nous ne reviendrons pas sur l'émergence, depuis le XIX^e siècle des normes régissant la production catalographique³.

La description bibliographique, l'indexation matière, la gestion des autorités ou la normalisation de données hétérogènes constituent des activités coûteuses, exigeantes, que l'accroissement continu de la masse documentaire rend difficiles à assurer. L'usage des différentes formes de l'intelligence artificielle (IA) pour créer ou compléter des métadonnées est aujourd'hui répandu : plusieurs établissements ont déjà instauré au sein de leurs chaînes de traitement des outils relevant de l'assistance ou de l'automatisation partielle. Ce mémoire étudie la façon des services innovants, du Chili à Singapour, mettent à profit l'IA pour le signalement, ce qui se joue dans ces déploiements, la manière dont ils transforment les métiers, et les stratégies de pilotage afférentes.

1.1. Accélération récente d'un phénomène ancien

1.1.1. Mobilisation des réseaux professionnels

L'idée d'automatiser tout ou partie du catalogage par des techniques relevant, au sens large, de l'IA, n'est pas nouvelle. Des projets existent depuis les années 1990, avec des vagues successives de méthodes, de corpus, et d'outillages, souvent centrées sur l'indexation automatique ou l'extraction d'informations⁴.

Mais depuis l'essor de l'apprentissage profond dans les années 2010, cette question a gagné en acuité, comme en témoigne l'attention accrue que lui accordent organisations nationales et internationales. La session organisée par la *Subject Analysis and Access Section* de l'IFLA portait sur l'indexation automatisée lors du congrès de 2018. En 2019, cette section s'est dotée d'un *Automated Indexing Working Group*. Signe de l'élargissement perçu du champ des possibles après ChatGPT, elle s'associait en 2023 à la *Cataloging section* pour animer une session au congrès de Rotterdam intitulée « *Utopia, threat or opportunity first? Artificial Intelligence and machine learning for cataloguing* ».

D'autres événements montrent le même essor : les congrès annuels *Fantastic Futures* organisés depuis 2018 par la fédération de bibliothèques, archives et musées AI4LAM, ou depuis 2019 les journées annuelles, en Allemagne, du *Netzwerks maschinelle Verfahren in der Erschließung*⁵.

Plusieurs groupes de travail se structurent explicitement autour de l'IA dans les chaînes de traitement des métadonnées. Citons le groupe du CENL sur l'IA en bibliothèques institué en 2020⁶, qui organise des webinaires tout au long de l'année ; le groupe d'intérêt spécial de l'IFLA sur l'IA, dont le communiqué de formation en 2021 mentionne l'indexation automatique et la classification des documents⁷ ; le groupe de travail sur l'IA du conseil de gouvernance de RDA et de son comité directeur (RSC) en 2024⁸ ; le groupe de travail du PCC consacré à la planification stratégique sur l'IA et l'apprentissage machine, institué en décembre 2023, dont le rapport final préconisait la création d'un groupe sur l'IA et l'apprentissage machine pour le catalogage et les métadonnées, lequel a récemment rendu ses conclusions⁹ ; le groupe de travail de LIBER sur l'IA en 2025¹⁰, qui s'inscrit dans le sillage

¹ (MESGUICH, 2023, p. 77).

² (MOREUX, 2019, p. 114).

³ Voir par exemple (RONCAGLIA, 2026, p. 109-112).

⁴ (PHILIBERT, 2024, p. 83-84).

⁵ (POLEY, UHLMANN, BUSSE et al., 2025, p. 26).

⁶ (AI in Libraries Network Group – The Conference of European National Librarians [sans date]).

⁷ (Exploratory Meeting : Formation of an IFLA Special Interest Group on Artificial Intelligence 2021).

⁸ (COHEN, ALIVERTI, BARBUS et al., 2025).

⁹ (PCC TASK GROUP ON AI AND MACHINE LEARNING FOR CATALOGING AND METADATA, 2025b).

¹⁰ (LIBER Launches a Taskforce on Artificial Intelligence 2025).

de la stratégie quinquennale de LIBER, laquelle met l'accent, dans son volet services, sur la curation d'informations et des données¹¹ ; le groupe de travail mis en place par OCLC sur la gestion de l'IA dans les circuits de métadonnées¹². Cette liste, qu'on pourrait encore étoffer, donne une idée de l'ampleur de la mobilisation face aux risques et opportunités que présente l'IA dans le domaine qui nous intéresse.

Cet intérêt de plus en plus vif est concomitant de l'émergence d'Annif¹³, une trousse à outils libre pour l'indexation automatique. Développé depuis 2017 par la Bibliothèque nationale de Finlande, cet ensemble d'algorithmes et de modèles a été adopté par la DNB, la ZBW, la Bibliothèque nationale de Suède, l'Abes ou la Bibliothèque du Congrès, parmi d'autres. Si l'indexation automatique a déjà plusieurs décennies d'histoire¹⁴, il devient plus facile d'imaginer et de mettre en œuvre un projet d'IA pour le signalement dès lors que les briques de construction fondamentales assurant une précision élevée des résultats sont réunies.

1.1.2. Une cadence précipitée depuis 2022

Les LLM viennent augmenter encore la vitesse de cette transformation. Leur spécificité, en substance, est de pouvoir s'adapter à des situations inconnues : « *AI's capacity to depart from script, to improvise based on training and experience, enables it to engage in expert judgment — a capability that, until now, has fallen within the province of elite experts* »¹⁵. Cette nouvelle technologie peut enrichir des méthodes existantes ou en faire germer de nouvelles. L'équipe responsable d'Annif a ainsi obtenu d'excellents scores lors d'une récente compétition annuelle autour de l'indexation en combinant des LLM et ses méthodes traditionnelles d'apprentissage machine¹⁶. Une approche similaire, avec un module de recherches sur internet en sus pour récupérer des résumés de titres, est exposée par le directeur de la recherche en métadonnées du réseau des bibliothèques de l'Université de Pennsylvanie¹⁷. Utilisés par eux-mêmes, les LLM capables d'analyser des images remplissent avec une précision élevée les métadonnées descriptives à partir de la page de titre numérisée¹⁸. Enfin, insérés dans des chaînes complexes en tant qu'agents, ils peuvent prélever dans la documentation professionnelle les passages pertinents pour le traitement de chaque ouvrage soumis et chercher, par le recours aux API disponibles, les entités liées à insérer dans la notice. Ça et là commencent à être décrits des projets relevant de cette catégorie, qui exploitent les progrès les plus récents de l'IA¹⁹. Tous ces développements sembleraient paver la voie à une adoption plus vaste des LLM dans les chaînes de production, et l'IFLA, dans une brève introduction à l'IA en bibliothèque²⁰, centre l'un des trois scénarios sur un projet de ce type. L'arrivée de cette technologie a rendu plus évidente la possibilité avérée depuis une décennie d'automatiser, au prix certes d'importants efforts, une partie plus ou moins large des circuits de description des documents physiques ou digitaux.

Il devient délicat d'éluder la question du recours à cette technologie. Si elle ne fait pas l'objet d'un projet de recherche et développement au sein de l'établissement, elle risque de s'y introduire à court terme par le biais des SIGB ou d'autres systèmes dont dépendent les bibliothèques. Comme le souligne un récent rapport de l'ADBU, l'application de l'IA à l'amélioration de la qualité des métadonnées fait partie des priorités des prestataires en informatique documentaire²¹. Ex Libris, pour ne mentionner qu'une entreprise dominante, offre maintenant des outils d'IA pour traiter des notices bibliographiques (création, insertion des entités liées, dédoublonnage, indexation)²². Dans son outil WorldShare Record Manager, OCLC permet déjà l'indexation automatique²³, et il est raisonnable de penser qu'une organisation qui utilise l'IA pour le dédoublonnage de WorldCat depuis 2023²⁴ ne s'en tiendra pas là. Le seul fait que la possibilité existe et soit à portée de main impose une réflexion de

¹¹ (JOHNSON, 2023, p. 5).

¹² (PROFFITT 2025a).

¹³ Pour un exposé technique de son fonctionnement, voir (SUOMINEN, INKINEN, et LEHTINEN, 2022). Pour une présentation générale et une synthèse récentes, voir (LEHTINEN, 2025a) et (SCHUMANN et MERGEL, 2026).

¹⁴ (GOLUB, 2021, p. 707).

¹⁵ (AUTOR, 2024, p. 7).

¹⁶ (D'SOUZA, SADRUDDIN, ISRAEL et al., 2025, p. 4).

¹⁷ (HAHN, 2024).

¹⁸ Voir l'exemple de la KBR, décrit dans (LOWAGIE, 2024).

¹⁹ Voir le schéma très clair d'une telle chaîne de traitement dans (KORPET et REES, 2025, p. 35) et une présentation plus générale dans (MALMSTEN et LUNDBORG, 2025).

²⁰ (COX et DE BRASDEFER, 2025, p. 10).

²¹ (ADBU, COMMISSION SDSA 2025, p. 1).

²² (AI Metadata Enrichment for Libraries 2024, p. 1), (VELTZMAN et ELSTEIN, 2024).

²³ (WorldShare Record Manager Release Notes 2025).

²⁴ (Implementing AI to Further Scale and Accelerate WorldCat De-Duplication 2025).

fond. Comme le souligne une récente synthèse de l'ADBU, il en va des « évolutions de nos métiers et de nos compétences »²⁵.

Mais l'accélération n'a rien d'uniforme. Les bibliothèques ne disposent pas des mêmes ressources ni les mêmes objectifs. Dresser la liste des facteurs expliquant l'adoption de l'IA reviendrait à décrire bon nombre des caractéristiques constitutives d'une institution : contexte national, rôle dans un réseau, masse documentaire à traiter, propension culturelle à l'automatisation, expertise technique disponible, etc. C'est un défaut courant des analyses panoramiques qu'on trouve dans la littérature professionnelle que d'aplanir les forts reliefs du paysage, et donc d'échouer à distinguer les projets les plus ambitieux et aboutis des diverses tentatives menées de par le monde. La forte variation, d'un continent à l'autre, des progrès accomplis explique qu'on lise, dans la synthèse d'un questionnaire transmis sur des listes de diffusion catalographiques en Amérique du Nord, qu'aux yeux de certains l'IA en est à ses débuts dans le domaine du signalement ; il est vrai que les projets les plus mûrs sont européens et n'utilisent ni FAST ni LCSH²⁶. En omettant la DNB et la ZBW, une méta-étude récente peut conclure qu'il n'existe pas de consensus sur l'utilité des technologies actuelles d'IA *lato sensu* pour le catalogage²⁷, une autre au contraire qu'il faut les adopter et cultiver les compétences idoines²⁸. Or ces deux établissements sont d'une importance cardinale pour toute réflexion sur le sujet qui se voudrait ancrée dans les possibilités réellement ouvertes par l'apprentissage machine et les LLM. Nous ne nions pas que l'IA soit un objet de fantasmes, nimbé d'un imaginaire faisant parfois obstacle à l'évaluation objective de ses résultats²⁹, mais on ne peut révoquer en doute son potentiel pour les missions catalographiques : c'est un fait établi qu'en Allemagne des centaines de milliers de documents électroniques sont indexées sans supervision unitaire des notices. Une enquête menée il y a maintenant sept ans auprès de directeurs de bibliothèques et d'experts anglo-saxons révélait de fortes incertitudes sur l'importance de l'IA et ses conséquences³⁰, y compris dans le domaine de la génération de métadonnées. Gageons que les avis seraient différents aujourd'hui : AutoSE et le volet indexation d'EMa sont en production depuis respectivement 2021 et 2022.

1.2. Périmètre de l'étude

Nous nous concentrerons sur l'usage de l'IA pour le signalement des ressources principalement textuelles, entendu comme l'ensemble des opérations qui créent, enrichissent, ou normalisent les métadonnées d'un document, qu'il soit physique ou digital, tel qu'un livre, un article, un fragment de littérature grise, en vue de l'exploitation de celles-ci dans un catalogue, un portail, ou des systèmes connectés en aval (entrepôts, moteurs de recherche, etc.). Parmi les GLAM, nous nous en tenons aux bibliothèques. La longueur allouée à un mémoire d'étude imposait d'écarter le traitement des contenus non textuels (images, audio, partitions, etc.), quoique certains enjeux techniques et organisationnels soient comparables³¹.

Lorsque l'IA est appliquée aux documents patrimoniaux (manuscripts, livres anciens), le but est souvent d'en transcrire le contenu ou d'en extraire les images (miniatures, enluminures, vignettes, etc.) pour analyse³². Notre intérêt se porte plutôt sur les situations où des milliers de documents *contemporains* dont les métadonnées descriptives ou sémantiques manquent doivent être versés au catalogue, typiquement au sein d'un service aux flux importants (dépôt légal), ou bien à l'occasion d'un chantier de catalogage rétrospectif. Les motivations sont tout à fait différentes : la transcription d'un manuscrit permet à un plus large public de prendre connaissance de son contenu, de préserver l'original en réduisant le nombre de consultations physiques, puis de mener des recherches relevant des humanités numériques (fouille de texte massive, par exemple)³³.

Le périmètre ainsi tracé est déjà fort vaste, comme on le verra. Aussi le mémoire laisse-t-il de côté tout recours à l'IA qui ne mènerait pas à la production ou à la modification de métadonnées réinjectées dans un système documentaire. Les agents conversationnels destinés aux services aux

²⁵ (ADBU, 2025, p. 2).

²⁶ (CHEN et AND LI, 2024, p. 325).

²⁷ (ENGEL, DO, SALEM et al., 2025, p. 8).

²⁸ (MWANTIMWA et MSOFFE, 2025, p. 164).

²⁹ Sur la construction de promesses autour de l'IA dans les organisations, voir (CHRISTENSEN, 2025).

³⁰ (COX, PINFIELD, et RUTTER, 2019, p. 16).

³¹ Sur les spécificités des ressources non textuelles dans l'automatisation, voir (ENGEL, DO, SALEM et al., 2025, p. 9).

³² Voir là-dessus l'excellent panorama fourni par un récent mémoire d'étude Enssib, (BOYER, 2025).

³³ (BOYER, 2025, p. 16).

publics sont ainsi exclus, par exemple, et en abordant la recherche sémantique dans un catalogue, nous n'étudierons que ses implications pour les métadonnées sous-jacentes.

1.3. Des chaînes de signalement hétérogènes : typologie et exemples

Nous distinguerons autant que possible les problèmes que l'on regroupe parfois sous un même terme. Automatiser l'indexation matière et automatiser la production du corps d'une notice ne mobilisent pas les mêmes méthodes ; la première tâche relève de la classification multilabel extrême³⁴, tandis que l'autre requiert de créer des zones MARC à partir d'une image, ou directement d'un bloc de texte lorsque le circuit de traitement évite l'analyse de fractions numérisées du document (métadonnées éditeur disponibles ou documents nativement digitaux).

Les opérations peuvent s'appuyer sur des techniques variées : apprentissage machine, traitement du langage naturel, modèles de vectorisation, LLM pour la création de métadonnées descriptives ou sémantiques. Loin d'être un objet unique, l'IA forme un faisceau de technologies et d'approches, qui se combinent différemment selon les cas d'usage. « *AI, it should be noted, is not a single technology, but rather a collection of multiple automation technologies that simulate human behaviour and cognition, and that can be applied in various ways across different sectors* », écrivent des membres de la bibliothèque de l'université de science et technique de Norvège³⁵. Ce constat est largement répandu, et on le trouve aussi sous la plume d'économistes : « *AI is no single thing but rather a variety of different AIs, in the plural, each with different characteristics, that do not necessarily replicate human intelligence.* »³⁶ Pour bien définir notre objet, nous proposons ci-dessous une typologie des chaînes de traitement rencontrées.

Une première famille regroupe les systèmes d'assistance au catalogage, où l'IA produit des suggestions et où la décision reste à l'humain : propositions de vedettes-matière, extraction de données à partir d'une page de titre, ou alignement d'entités. Le gain recherché est une accélération du travail, par réduction du temps de saisie et de recherche, tout en conservant un contrôle de qualité humain dans le flux. Dans ces configurations, parler d'humain dans la boucle peut être trompeur, avance-t-on parfois : il s'agit plutôt d'IA dans la boucle, l'outil restant subordonné à la décision et au raisonnement du professionnel au cours même du signalement. Cette dichotomie, nullement propre au monde des bibliothèques, est soulignée dans des études générales sur la place respective de l'IA et de l'humain dans des systèmes qui les mettent en synergie : « *In human-in-the-loop systems, AI systems drive the inference and decision-making process, but humans intervene to provide corrections and supervision. In AI-in-the-loop systems, humans make the ultimate decisions, while AI systems assist with perception, inference, and action* »³⁷. La KB et la KBR sont les établissements les plus représentatifs de cette tendance : dans leurs projets de rétrocatalogage, une notice est générée à partir de la numérisation de différentes parties de l'ouvrage (page de titre, table des matières, quatrième de couverture, colophon), qu'un catalogueur inspecte avant validation.

Une deuxième famille comprend des chaînes semi-automatisées où l'IA agit de façon plus autonome, mais où le contrôle humain est maintenu par des dispositifs de supervision, de relecture ciblée ou d'échantillonnage après coup. Le travail du catalogueur se déplace alors : il s'exerce moins sur chaque notice individuellement que sur des lots, pour déceler des erreurs typiques ou fournir de nouvelles données d'entraînement. Cela permet une forme d'industrialisation qui n'élude pas la responsabilité documentaire. Mais la notion de qualité du signalement s'en trouve ébranlée, et la soutenabilité à long terme de cette relecture ne va pas de soi, comme on le verra. Il s'agit souvent de corpus qui, sans automatisation, ne seraient pas traités. L'enjeu n'est pas seulement de remplacer un travail existant, mais d'étendre le périmètre de ce qui est signalé. Cette extension peut être présentée comme une amélioration de service, mais elle engage aussi une redéfinition de la qualité attendue : on accepte des métadonnées qui n'ont pas été vérifiées une à une si elles rendent possible une exploration auparavant impossible. Nous n'avons pas identifié d'autres cas que la ZBW et la DNB dans cette catégorie. Ces deux bibliothèques indexent automatiquement des centaines de milliers de publications digitales : articles

³⁴ (NICOLAS, 2025b, p. 4).

³⁵ (NOGUEIRA, MOLTUBAKK, FAGERVIK et al., 2025, p. 736).

³⁶ Sur cette pluralité, voir (AUTOR, MINDELL, et REYNOLDS, 2022, p. 44).

³⁷ (NATARAJAN, MATHUR, SIDHÉKH et al., 2024, p. 3). Le schéma illustrant cette distinction est très éclairant.

académiques d'économie pour la première, livres électroniques, publications universitaires, articles de presse, instantanés de sites internet, articles académiques de toute nature pour la seconde³⁸.

Ces deux grands groupes ne sont pas exclusives : passer d'un outil d'assistance à un outil automatisé ne requiert qu'une décision organisationnelle, les briques technologiques étant les mêmes. À la ZBW, les catalogueurs continuant de pratiquer l'indexation manuelle bénéficient des recommandations du modèle qui génère, par ailleurs, les vedettes-matière des dizaines de milliers de notices traitées automatiquement.

1.4. Quelle bascule organisationnelle ?

Notre question de départ était la suivante : que se passe-t-il lorsqu'une bibliothèque déploie des outils d'IA pour le signalement, et que ces outils cessent d'être des prototypes pour devenir des composants d'une chaîne de production ? Parmi les différents discours sur le sujet, on peut distinguer quelques antilogies structurantes :

- D'un côté, l'expertise humaine est considérée comme centrale : l'IA aurait besoin du bibliothécaire pour contrôler, corriger, et maintenir un niveau de qualité suffisant dans les notices traitées, de sorte que seul un système d'assistance est recevable. De l'autre, les objectifs d'échelle ne peuvent être atteints que si la décision humaine s'exerce lors d'un contrôle a posteriori, par échantillonnage ou par relectures ciblées, ce qui change la nature même du travail.
- D'un côté, l'automatisation est présentée comme un moyen de redéployer des compétences vers des tâches à plus forte valeur ajoutée, ou de maintenir un service malgré une contrainte de ressources humaines. De l'autre, elle pourrait produire un appauvrissement du travail pour les agents concernés.
- D'un côté, l'IA est perçue comme un moyen d'améliorer la qualité des métadonnées en réduisant certaines défaillances humaines. De l'autre, on craint qu'elle n'introduise des erreurs systématiques, des biais difficiles à détecter. Selon les pans de la littérature qu'on dépouille et les experts qu'on interroge, tantôt l'application des normes catalographiques et des vocabulaires contrôlés offrent un terrain où l'IA peut exceller, tantôt les difficultés d'interprétation de ces systèmes de gestion de la connaissance si quintessentiellement humains sont érigées en obstacles à la performance algorithmique, lesquels garantiront toujours la supériorité des spécialistes³⁹.

À partir de ces tensions, la question générale peut être ainsi formulée : dans quelles conditions l'IA appliquée au signalement constitue-t-elle un dispositif d'augmentation viable pour les bibliothèques, et dans quelles conditions risque-t-elle de déstabiliser des organisations, d'affaiblir des compétences ? Dans le périmètre circonscrit de la gestion des métadonnées, nous rejoignons ainsi les préoccupations de l'ADBU : « Comment faire de l'IA un soutien et une plus-value au développement de nos compétences, de nos métiers et de nos services en préservant les savoirs, savoir-faire et savoir être fondamentaux nécessaires au pilotage de nos missions et à leur qualité ? »⁴⁰. Notre problématique peut se décomposer en interrogations plus précises :

- Quelles finalités un établissement assigne-t-il à un projet d'IA appliqué au signalement : accélération du traitement rétrospectif, extension à des corpus non traités, amélioration de la qualité, réduction de coûts, maintien d'un service face à la contrainte de ressources humaines ? Comment ces finalités se traduisent-elles en choix techniques et organisationnels (répartition des tâches, formation, composition des équipes, recours à des prestataires) ?
- Quelles configurations se mettent en place concrètement entre l'humain et la machine : à quel moment la décision humaine intervient-elle, avec quels outils de contrôle, quels indicateurs, quelles responsabilités ? Dans quels cas observe-t-on une bascule vers un contrôle a posteriori, et comment cette bascule est-elle rendue acceptable au plan professionnel ?
- Quels effets observe-t-on sur les métiers : évolution des profils de postes, nouvelles compétences attendues, reconfiguration des collectifs de travail, et risques de perte ou de transformation de

³⁸ (POLEY, UHLMANN, BUSSE et al., 2025, p. 3).

³⁹ Sur les doutes exprimés par des professionnels à propos de l'indexation matière automatisée, voir par exemple (CHEN et AND LI, 2024, p. 325-326).

⁴⁰ (ADBU, 2025, p. 2).

l'expertise ? Dans quelle mesure l'IA sert-elle une logique d'augmentation ou une logique de substitution, et comment ces logiques coexistent-elles dans les discours et les décisions ?

- Quels effets observe-t-on sur les métadonnées ? Dans ce domaine, les principaux sujets sont le devenir des vocabulaires contrôlés, l'apparition de nouvelles données (vecteurs notamment), la question de l'étiquetage de métadonnées générées, et la transformation des critères de qualité.

1.5. Principes de sélection des cas

Nous avons privilégié des chaînes destinées à durer et à modifier les pratiques, traitant de vastes volumes de documents, plutôt que des prototypes sur un corpus étroit. Le mémoire portant sur les effets organisationnels, l'évolution des métiers et les métadonnées, nous donnerons des explications techniques quand elles sont nécessaires à l'intelligence des projets menés, sans produire un manuel sur l'exploitation des outils impliqués. Nous espérons ainsi dégager de l'état de la technologie à la date de remise (mars 2026) des régularités et des points de vigilance qui resteront au moins partiellement pertinents malgré les évolutions ultérieures.

Le choix des établissements a été dicté par le degré de maturité et l'originalité de leurs projets d'IA. Il s'agit principalement de bibliothèques nationales, pour des raisons exposées en première partie (pression de la masse documentaire, moyens techniques disponibles). On en trouvera la liste en parcourant la bibliographie thématique. La littérature bibliothéconomique a été complétée par des d'entretiens semi-directifs menés avec des acteurs directement impliqués dans des projets d'IA appliqués au signalement. Ces entretiens couvrent notamment la Bibliothèque nationale de France, l'Abes, la KB (Pays-Bas), la KBR (Belgique), la Bibliothèque du Congrès, la ZBW (Allemagne), la Bibliothèque nationale de Finlande, la Bibliothèque nationale d'Israël, le NLB (Singapour). Les bibliothèques en mesure de mener à bien de tels projets sont suffisamment rares pour qu'il soit raisonnable d'envisager de rencontrer une fraction substantielle des responsables concernés, et les spécificités qu'ils présentent dissuadent de recourir à une enquête générique par questionnaire.

Quoique les entretiens soient riches d'enseignements, une discussion d'une heure ne saurait permettre de reconstituer finement la myriade de décisions petites ou grandes ayant jalonné le développement d'un projet, ni l'évolution des attitudes des dizaines de personnes impliquées sur plusieurs années, leurs dynamiques collectives, la mue graduelle de leurs méthodes de travail. Il aurait fallu conduire une étude ethnographique, en immersion dans un service touché, comme des études pionnières dans d'autres domaines nous en donnent d'excellents exemples⁴¹. Certes le signalement dans les bibliothèques est une activité qui présente des caractéristiques très particulières, mais ces institutions ne sont pas si atypiques que les cadres d'analyse d'autres disciplines perdent toute validité. L'économie, la sociologie, le champ des études sur les interactions entre l'humain et la machine permettent de replacer le phénomène étudié ici, l'implantation de l'IA au sein d'un environnement de travail, dans un cadre plus large. Qu'on tire de ces études des micro-analyses de situation ou des enseignements généraux sur les effets de l'automatisation sur les professions, elles contiennent de nombreux aperçus applicables aux cas observés pendant notre enquête⁴².

1.6. Plan schématique

Nous examinerons d'abord les raisons poussant les institutions vers l'IA catalographique, les ressources qu'elles doivent mobiliser chemin faisant, et les enjeux juridiques ou éthiques indissociables de tels projets.

Le reste du mémoire est centré sur les deux pôles toujours présents dans l'accomplissement d'un projet d'IA, celui des spécialistes du catalogue, et celui des ingénieurs en IA et informaticiens. Quel que soit le pays et la nature de l'établissement considéré, la collaboration de ces deux forces est une condition *sine qua non* ; cette structure élémentaire stable rend justifiable la comparaison entre une agence bibliographique nationale européenne et une bibliothèque universitaire américaine, par exemple.

⁴¹ (GRØNSUND et AANESTAD, 2020) suit une entreprise de fret maritime pendant les deux années où le recours à l'apprentissage machine s'y répand progressivement. La répartition des tâches et des rôles entre les employés et l'algorithme est au cœur de l'enquête.

⁴² Sur l'intérêt de mobiliser d'autres cadres d'analyse pour comprendre l'IA dans les bibliothèques, voir (COX, 2022).

Nous embrassons d'abord le point de vue des bibliothécaires, au plus près du signalement. Cette partie traite des dangers perçus pour l'emploi, de l'adhésion des équipes et des opportunités de redéfinition de leurs missions que l'IA présente. Les deux suivantes portent elles aussi sur les détenteurs de l'expertise catalographique : la première montre pourquoi celle-ci reste indispensable au bon fonctionnement des circuits de traitement, et la seconde analyse les conséquences potentielles d'un passage du rôle de rédacteur de notices unitaires à celui de relecteur de lots.

Enfin nous nous penchons sur le problème de la formation. Cette dernière partie aborde les autres profils de poste nécessaires au succès d'un projet d'IA, et les différences fondamentales entre les établissements qui choisissent de cultiver leur expertise en ingénierie IA interne et ceux recourant à des prestataires.

ADOPTION INSTITUTIONNELLE DE L'IA : QUELS DÉFIS ?

FEUILLE DE ROUTE

Cette première partie se situe au niveau institutionnel. Elle explique pourquoi les grandes agences bibliographiques nationales offrent un terrain d'application idéal aux technologies d'IA : agissant dans un écosystème de données normalisé en profondeur, elles sont confrontées à une croissance continue de la masse documentaire à traiter, relevant de disciplines infiniment variées, dans des langues innombrables. Pour tirer parti de cette technologie, elles doivent mobiliser de très larges ressources tant pour la phase d'idéation et de prototypage que pour celle, si délicate, du passage en production. La profession est consciente des enjeux éthiques que l'IA soulève, et les institutions ne sauraient les éluder.

Liens des sections :

[↗ Un domaine propice à l'automatisation contrôlée](#)[↗ Un problème d'échelle : volumes et arriérés](#)[↗ Un substitut à la spécialisation linguistique et disciplinaire ?](#)[↗ Mobiliser les ressources nécessaires](#)[↗ Interopérabilité, passage à l'échelle et mise en production](#)[↗ Enjeux éthiques](#)

2.1. Un domaine propice à l'automatisation contrôlée

Quoique certaines propriétés des bibliothèques facilitant le recours à l'IA soient évidentes, elles méritent d'être soulignées. Le signalement est informatisé et normalisé à l'échelle internationale : codes, formats, vocabulaires, autorités font l'objet d'efforts d'interopérabilité depuis les années 1960 afin de faciliter la mutualisation des efforts⁴³. L'automatisation s'en trouve facilitée, par rapport à des domaines où les pratiques sont hétérogènes et peu standardisées, comme par exemple les fonctions supports administratives en milieu médical⁴⁴.

L'uniformisation qui en découle n'a pas les connotations négatives qu'elle peut prendre ailleurs⁴⁵, parce que la finalité du signalement est précisément la production de données standardisées. Néfaste aux activités créatives, la standardisation a ici l'effet bénéfique d'améliorer la cohérence des catalogues.

L'IA forme souvent le rouage d'une chaîne plus étendue⁴⁶ : dans celle imaginée par la KBR, après détection de l'ISBN sur la page de titre, celui-ci est recherché dans le catalogue de la bibliothèque⁴⁷, dans WorldCat ou d'autres services. L'interrogation de ces bases de données locales ou distantes est rendue possible par une toute une tuyauterie (protocoles SRU, API) mise en place depuis longtemps. Sans cela, un LLM, aussi performant soit-il, ne pourrait correctement identifier les autorités liées d'une notice ou ses doublons, par exemple.

Enfin, notons que pour les projets ne se contentant pas d'un LLM dans son état de base, la masse disponible de données à peu près homogènes dans les catalogues est une condition clef de leur réussite. Riches de millions de notices, ils offrent un réservoir de choix pour mener l'entraînement des modèles d'IA, qu'il s'agisse d'apprentissage machine traditionnel ou de spécialisation (*fine-tuning*) d'un LLM. Comme l'écrivait Jean-Philippe Moreux en 2019 à propos des métadonnées sémantiques, « l'indexation automatique trouve dans les bibliothèques un terreau favorable à sa pleine expression, du fait de l'expertise de ces dernières en sciences de l'information et de la disponibilité de corpus structurés et annotés indispensables à l'apprentissage machine »⁴⁸.

⁴³ (WENZ, 2009, p. 91).

⁴⁴ (AUTOR, MINDELL, et REYNOLDS, 2022, p. 52).

⁴⁵ (TAVARES et JARAVEL, 2025, p. 28).

⁴⁶ (LOWAGIE, 2025b).

⁴⁷ (LOWAGIE, 2024, p. 7).

⁴⁸ (MOREUX, 2019, p. 133).

2.2. Un problème d'échelle : volumes et arriérés

Les bibliothèques font face à des masses documentaires qui excèdent aujourd'hui la capacité de traitement manuel, à cause de l'accroissement continu des flux (dépôt légal, ressources numériques) ou de l'accumulation d'arriérés (rétrocatalogage, mise à niveau de métadonnées anciennes) tandis que les services concernés sont loin de grandir dans les mêmes proportions, quand ils ne perdent pas d'agents⁴⁹. L'exploration du recours à l'automatisation du signalement est souvent attribué à ces tendances inverses, quel que soit le pays⁵⁰. La description bibliographique représente dans ce contexte un défi considérable, et l'indexation plus encore, alors que des études prouvent la grande utilité des recherches par vedette-matière⁵¹. L'IA est donc souvent présentée comme un moyen non de surpasser ou supplanter l'humain, mais simplement de rendre consultables des ressources qui, autrement, demeureraient invisibles. L'argument est universel, et cette formulation par la Bibliothèque du Congrès en offre une synthèse claire :

*LAM institutions share common challenges that arise from large, and often ever increasing, collections which are often undescribed, or under described; which are inaccessible to users because they have not been digitised or exist only in languages or scripts that are not widely understood; and which cannot be found because good quality metadata that supports search does not exist. Solving these problems has historically required expert labour which is a challenge when staff time and budgets are spread thinly. AI and ML methods are potentially a solution to some of these problems by supplementing expert staff time with computational methods.*⁵²

Certes, les bibliothèques universitaires décrivent des arriérés importants et un manque de personnel spécialisé pour y faire face, de l'Université Paris-Saclay à l'Université autonome du Chili⁵³, mais ce sont surtout les bibliothèques nationales qui se trouvent exposées à cette contrainte d'échelle. À l'augmentation générale des volumes de publication s'ajoute l'extension continue de leurs obligations réglementaires de collecte. Les nouveaux domaines couverts, qu'ils soient digitaux⁵⁴ ou imprimés (auto-édition à la BnF)⁵⁵, rendent la gageure intenable à masse salariale constante. Le recours à l'IA est dans ces conditions présenté parfois comme un impératif : « *emph[It is not possible to manually index or annotate every single one of the plethora of documents available in multiple formats anymore which leads to the necessity of exploring the potential use of automated means on every level.]* »⁵⁶

Cet enjeu d'échelle conduit les établissements à interroger leurs exigences de qualité : lorsqu'il faut choisir entre traitement manuel exhaustif et absence de signalement, quel parti prendre ? Ce n'est pas la première fois qu'une technologie met face à pareil dilemme ; la transcription automatique, lors des débuts de Gallica, avait causé des controverses entre deux camps, les « pragmatiques » d'une part, les [orthodoxes de la qualité] de l'autre pour résumer à grands traits les positions en présence⁵⁷. Les projets d'IA bénéficient d'une réflexion déjà en cours sur la question. Ainsi, la BnF prévoit dans sa politique de qualité des données le « recours plus systématique, dès que le contexte le rend pertinent, aux traitements automatisés et semi-automatisés »⁵⁸. Notons aussi le cas de l'Abes, qui dans son projet d'établissement 2024-2028, établit une « politique de la qualité différenciée », visant à concentrer l'effort expert là où il est jugé le plus nécessaire⁵⁹. Cette ligne directrice, tracée plutôt pour les versements et l'importation de données externes, pourrait valoir aussi pour les compromis à consentir lors d'une adoption éventuelle de l'IA.

⁴⁹ (LOWAGIE, 2025a, p. 16).

⁵⁰ (CAVALIE, 2019a, p. 111). (ANDERSON et TIAN, 2025p. 363).

⁵¹ (GOLUB, 2021, p. 702-703) et (GOLUB et PESTANA, 2023, p. 205-207).

⁵² (POTTER et MCGRATTAN, 2025, p. 1).

⁵³ (DIAZ et VERGARA, 2025, p. 6). (BELLIER et MINQUILAN, 2025).

⁵⁴ Voir (MERGEL, KÜHLER, SCHÜMANN et al., 2026, p. 28) pour la DNB par exemple.

⁵⁵ (PHILIBERT, 2024, p. 33).

⁵⁶ (KASPRZIK, 2024b, p. 200).

⁵⁷ (MOREUX, 2025a).

⁵⁸ (PHILIBERT, 2024, p. 36).

⁵⁹ (PHILIBERT, 2024, p. 51).

2.3. Un substitut à la spécialisation linguistique et disciplinaire ?

La variété des écritures, des langues et des spécialités disciplinaires offre un point de convergence entre les besoins des bibliothèques et les possibilités offertes par l'IA. La profession fait grand cas de cette facette de la technologie : dans un sondage distribué sur des listes de diffusion catalographiques aux États-Unis, 38% des réponses (sur un total de 201) faisaient de la traduction la tâche où l'IA apporte le plus, 30% privilégiaient l'aptitude à générer un résumé⁶⁰. Dans un établissement comme la Bibliothèque du Congrès, traitant une vaste étendue de documents relevant de domaines très différents, les entrées sont réparties, logiquement, selon les compétences de chaque catalogueur⁶¹ : l'IA favorisera-t-elle une plus grande polyvalence et une répartition plus homogène des entrées ? On peut se demander quelle sera la trajectoire d'emploi pour les profils spécialisés : les bibliothèques nationales ou de recherche dont les acquisitions couvrent une mosaïque de champs auront-elles les moyens et la volonté politique de conserver en leur sein les experts pour vérifier, à l'unité ou par échantillonnage, la validité des métadonnées descriptives et sémantiques dans les notices de traités bouddhiques tibétains ou de monographies ethnographiques en norvégien ? Pour l'instant les LLM états-uniens et propriétaires les plus connus (ChatGPT, Claude, Gemini) ne manient qu'imparfaitement ces deux langues, mais les communautés de locuteurs se mobilisent pour entraîner des modèles compétents. Lorsque le taux de précision constaté pour la traduction avoisinera les 99%, qu'en sera-t-il ? Et sans spéculer sur cette perspective, qu'en est-il de langues étrangères où tout LLM est déjà bien versé ? Cette question n'est pas à prendre isolément : dès lors qu'il faut des acquéreurs spécialisés pour déterminer et appliquer la politique documentaire, le sort du signalement des collections dont ils ont la charge ne détermine pas nécessairement la pérennité des postes qu'ils occupent.

À l'inverse, la technologie peut aussi aider les spécialistes à participer à l'effort de signalement en renforçant les compétences proprement catalographiques dont ils manqueraient : l'outil RetroTool de la KB a donné à un catalogueur maîtrisant l'alphabet cyrillique l'opportunité d'utiliser MARC 21, et à l'avenir l'institution espère que son outillage d'IA servira à développer les compétences catalographiques des agents⁶². L'IA offrirait une aide décisive à l'insertion, dans les circuits de traitement, d'une expertise externe : dans un scénario proposé par les membres du groupe *Primary cataloging workflows* au sein du *Managing AI in Metadata Workflows Working Group* formé par l'OCLC Research Library Partnership, l'établissement solliciterait des chercheurs ou étudiants possédant une langue dont aucun bibliothécaire n'a connaissance pour vérifier la fiabilité des métadonnées générées par l'IA⁶³.

2.4. Mobiliser les ressources nécessaires

Pour tirer profit des gains potentiels que l'IA fait espérer, l'institution doit être prête à consentir des investissements de grande ampleur, comme le soulignent de nombreuses études de cas et analyses panoramiques. Les coûts en recherche et développement (entraînement des modèles, comparaisons statistiques pour trouver la recette la plus appropriée à la tâche), éventuellement en acquisition de matériel, sont considérables⁶⁴. Dans le cas des LLM, l'inférence de modèles hébergés par des prestataires peut être coûteuse⁶⁵, tout comme le matériel nécessaire à un circuit purement local⁶⁶. La synthèse de dizaines d'études sur la question conclut que l'aspect financier (coût d'entrée et de maintenance) est le thème de discussion le plus fréquent⁶⁷. Le temps humain nécessaire (évaluation des options, calibrage de la solution, contrôles) est tel qu'il ne faut pas s'attendre à dégager des économies ou réduire la masse salariale, disent clairement Avaha Cohen et ses pairs⁶⁸. La DNB, qui porte pourtant l'un des projets les plus avancés, aurait voulu lui accorder une enveloppe budgétaire plus large, notamment pour dans le personnel et le matériel⁶⁹. La KBR insiste sur la nécessité de veiller à la dépense⁷⁰.

⁶⁰ (CHEN et AND LI, 2024, p. 323).

⁶¹ (POTTER, 2025).

⁶² (BUESINK, 2025b).

⁶³ (PROFFITT 2025b).

⁶⁴ Cela concerne les projets d'IA relevant aussi bien du signalement que d'autres domaines bibliothéconomiques, voir (MANNHEIMER, BOND, YOUNG et al., 2024, p. 21).

⁶⁵ (DIAZ et VERGARA, 2025, p. 21).

⁶⁶ (POLEY, 2025), (KASPRZIK, 2023p. 11).

⁶⁷ (PINAR et COX, 2025, p. 578).

⁶⁸ (COHEN, ALIVERTI, BARBUS et al., 2025).

⁶⁹ (MERGEL, KÜHLER, SCHUMANN et al., 2026, p. 36).

La question des économies potentielles est liée à celle de la productivité des agents. Certes, il y aura sans doute quelques gains d'efficacité à la Bibliothèque du Congrès grâce au projet *Exploring Computational Description*, mais il faudra investir beaucoup d'argent⁷¹. Certains catalogueurs du réseau SUDOC constatent des gains de temps lors de l'utilisation de l'outil d'assistance à l'indexation du Labo, d'autres non, qui mettent plutôt l'accent sur l'amélioration de la qualité⁷². L'équipe d'Annif ne le présente pas comme un outil d'optimisation du temps de traitement (si on l'insère dans un système d'assistance) : « *If you are a professional indexer, Annif doesn't make the process go any faster. You still have to check the work and you still have to think about it. But it may help you by coming up with some suggestions that perhaps you didn't think about in the first place because the vocabularies are vast.* »⁷³ Un projet d'automatisation sur un corpus jusqu'alors non traité prendra un temps considérable aux catalogueurs avant de mener à des résultats fiables et publiables.

Les sommes à engager sont élevées car greffer une IA catalographique dans une chaîne de traitement prend du temps. Le système requérant un contrôle régulier après sa mise en place, on ne peut traiter pareille initiative comme un projet ponctuel. À la ZBW, AutoSE a vu le jour après que la direction eut fait du développement d'un outil d'indexation automatique une mission permanente⁷⁴. Les changements technologiques sont rapides et nécessitent une expertise interne qui prend des années à cristalliser : ce n'est pas un sprint mais un marathon, témoigne un membre de la DNB⁷⁵. Il s'agit d'un processus continu⁷⁶, dont on ne peut déterminer à l'avance combien de temps il prendra, sur quels obstacles temporaires il achoppera⁷⁷ : des financements de trop courte durée risquent de conduire à un arrêt des opérations avant l'obtention de résultats et au départ des acteurs impliqués, comme on l'observe dans le domaine spécifiquement patrimonial⁷⁸. Le projet *Exploring Computational Description* de la Bibliothèque du Congrès est en cours depuis 2022 et n'est pas encore entré en production ; Retrotool à la KB est terminé depuis 2024 mais le développement de sa version pérenne, destinée à traiter un plus large éventail de collections, n'a commencé que début 2026⁷⁹. Quoique la complexité inhérente à une IA catalographique appelle des compétences spécifiques, son développement devrait idéalement s'insérer dans une stratégie d'ensemble de l'établissement sur les applications et leur cycle de vie, telle que Gildas Illien l'appelaient de ses vœux il y a déjà dix ans : « traiter le problème au cas par cas et à la petite semaine »⁸⁰, alors que les ressources à mobiliser sont si considérables, ne saurait mener au succès.

La KB recommande de consacrer à la préparation d'un projet d'IA un temps de réflexion suffisant⁸¹. Dans un projet purement interne, de nombreux problèmes techniques requièrent une méticuleuse enquête préalable : par quels circuits les données circuleront-elles, comment garantir la disponibilité continue des serveurs et créer un environnement de développement continu pour continuer d'améliorer le système après sa mise en production⁸² ? La question du matériel, de la puissance de calcul nécessaire, devient critique lorsqu'on choisit de recourir à la vectorisation, à des modèles transformateurs ou à des LLM comme le montre un tableau comparatif établi par la DNB⁸³. Des méthodes ou cadres de réflexion existent pour formaliser les aspects principaux du projet lors de cette phase initiale, comme le *AI Risk Management Framework* du *National Institute for Standards and Technology* aux États-Unis, portant sur la fiabilité, la validité, la sécurité, la capacité de résistance, la responsabilité, la transparence, l'explicabilité et l'explicabilité et l'impartialité des systèmes construits⁸⁴. La division *LC Labs* de la Bibliothèque du Congrès, avant de s'engager dans *ECD*, s'est appuyée sur les préconisations du NIST et sur un schéma de planification conçu par ses soins, reposant sur la division de tout projet en trois phases (compréhension, prototypage, mise en production) et en trois composantes (données, personnes, infrastructure technique)⁸⁵.

⁷⁰ (SCHUMANN et MERGEL, 2026, p. 89).

⁷¹ (POTTER, 2025).

⁷² (NICOLAS, 2025b, p. 21).

⁷³ (LEHTINEN, 2025b).

⁷⁴ (KASPRZIK, 2025a), (KASPRZIK, 2024cp. 33).

⁷⁵ (MERGEL, KÜHLER, SCHUMANN et al., 2026, p. 35).

⁷⁶ (SCHMIDT, KÜHLER, SCHUMANN et al., 2026, p. 186).

⁷⁷ (SCHMIDT, KÜHLER, SCHUMANN et al., 2026, p. 187).

⁷⁸ (BOYER, 2025, p. 42).

⁷⁹ (*Automatic Metadata Generation at the KB, National Library of the Netherlands* [sans date]).

⁸⁰ (ILLIEN, 2015, p. 105).

⁸¹ (WILLEMS, ANDUCAS, et SILVESTRINI, 2026, p. 216).

⁸² (KASPRZIK, 2020, p. 7).

⁸³ (KAHLER, 2025, p. 25).

⁸⁴ (SACCUCCI et POTTER, 2024, p. 231).

⁸⁵ (MANCHESTER 2023).

2.5. Interopérabilité, passage à l'échelle et mise en production

Fabriquer à l'écart, dans un cadre expérimental vierge, un prototype capable de transformer des pages de titre en notices ou d'émettre l'indexation matière d'articles numériques ne suffit pas à passer en production. L'intégration au système informatique de l'établissement constitue un défi à part entière⁸⁶. À la BnF, de plus en plus de projets de recherche donnent lieu à un prototype prometteur, notamment dans le cadre de la cellule IA, mais l'intégration au SI de l'établissement, riche déjà d'un portefeuille d'un millier d'applications, et le passage à l'échelle constituent de grands défis⁸⁷. C'est l'un des enseignements que formule une méta-étude récente : « *The integration of new AI-based services with legacy systems seems a key dimension of AI capability. This is less a weakness of the capability model and more a reflection that AI capability is just part of the organization's overall operations.* »⁸⁸ Elle requiert souvent la transmission du projet à d'autres départements, selon des moyens, des calendriers, et une méthode d'accompagnement à définir, surtout si des actualisations régulières sont nécessaires pour maintenir le modèle à jour et donc la fiabilité des résultats, comme dans le cas du prototype RAMEAU de l'Abes⁸⁹.

Le cadre technique où vient se greffer un projet d'IA détermine parfois, par les contraintes qu'il pose, le champ des possibles. La KB a ainsi restreint l'automatisation à des documents physiques absents du catalogue, non pas à cause de problèmes d'efficacité ou précision de l'outil Retrotool, mais parce que l'infrastructure ne permet pas de réinjecter facilement des métadonnées dans des notices déjà présentes⁹⁰. L'exploitation des résultats de l'IA est souvent dépendante d'une architecture de données et de décisions de conception antérieures à son application.

La difficulté s'accroît lorsque l'ambition est d'étendre l'outil à plusieurs chaînes de traitement, comme le souhaite la Bibliothèque du Congrès, où le signalement se ramifie en circuits variés selon les types de documents, les supports et les flux d'acquisition. Concevoir un service d'indexation ou d'enrichissement réutilisable dans des contextes divers implique d'anticiper cette diversité dès la conception : interfaces, formats d'entrée, critères de qualité, modalités de supervision, et compatibilité avec des systèmes variés en aval⁹¹.

Enfin, la temporalité des grands chantiers SI peut entraver toute injection de solutions IA à court terme. Une migration majeure comme le passage à Noemi à la BnF peut dissuader de s'engager dans des projets d'IA catalographique non parce qu'ils seraient jugés inutiles, mais parce que la priorité est autre : que les catalogueurs disposent d'un outil correspondant à leurs attentes, qui ne pourront être finement exprimées qu'après une certaine pratique du nouveau code et du nouvel outil⁹². Sur le plan de l'accompagnement au changement, après de lourdes transformations d'infrastructure qui ont vu la KB changer de SIGB et de format (passage à Alma et MARC 21), renouveler l'expérience de l'adoption d'une nouvelle technologie a pu nourrir le scepticisme⁹³. Du point de vue d'un établissement tête de réseau, les changements en cours à l'Abes (abandon de WinIBW et remaniement subséquent de l'offre de services) posent la question de l'opportunité d'investir dans le déploiement de l'assistant d'indexation Rameau⁹⁴.

2.6. Enjeux éthiques

Sur le plan éthique, de toutes les formes d'IA ce sont les LLM qui soulèvent dans les bibliothèques le plus d'interrogations. Pour le dire sans ambages : « *How do we feel, consciously, in partnering with a company such as OpenAI that outsources its data labeling of toxic content to a company that exploits Kenyan workers so the end users benefit ?* »⁹⁵. Un établissement public, en choisissant de recourir au produit d'une entreprise, peut-elle ne pas se soucier des chaînes d'étiquetage et de filtrage déléguées à une main-d'œuvre précaire⁹⁶ ?

⁸⁶ (MERGEL, SCHMIDT, SCHUMANN et al., 2024, p. 19), (HENRY, 2026, p. 116).

⁸⁷ (MOREUX, 2025b).

⁸⁸ (PINAR et COX, 2025, p. 587).

⁸⁹ (NICOLAS, 2025a).

⁹⁰ Voir (BUESINK, 2025b).

⁹¹ Voir (POTTER, 2025).

⁹² (KOSKAS, 2025).

⁹³ Voir (BUESINK, 2025b).

⁹⁴ (NICOLAS, 2025a).

⁹⁵ (MURPHY, 2025, p. 624).

⁹⁶ (KASPRZIK, 2025b, p. 7).

L’empreinte environnementale de l’IA est débattue dans le discours public et naturellement les publications bibliothéconomiques s’en émeuvent⁹⁷, parfois énergiquement⁹⁸. Les gains d’efficacité espérés compensent-ils la pollution générée⁹⁹ ? Pour répondre, il faudrait mesurer l’ensemble des externalités négatives d’un projet d’IA, ce qui est faisable en cas de circuit local, mais à peu près impossible si l’on recourt à des LLM hébergés par un prestataire.

Sur le plan juridique, les questions de droit d’auteur et de réutilisation sont centrales dès que les chaînes traitent des contenus protégés : peut-on utiliser l’apprentissage machine pour analyser des textes aux droits réservés, et que devient le contenu après l’entraînement ? Ces interrogations sont explicitement formulées par des institutions engagées dans des projets d’IA¹⁰⁰. Elles rejoignent un autre enjeu : celui de la confidentialité des données traitées (documents non encore publiés, collections sensibles, données personnelles éventuelles) et de la souveraineté sur les corpus internes.

L’externalisation déplace une part de la responsabilité vers des acteurs dont les intérêts ou les conceptions de l’éthique ne coïncident pas nécessairement avec ceux des bibliothèques, et qui opèrent sous une autre juridiction. « *We take great care to make sure that our metadata enrichment not only conforms to legal requirements, but also to ethical considerations* »¹⁰¹, soutient tel fournisseur de SIGB qui s’appuie sur un modèle d’OpenAI. De quelle éthique est-il question dans cette déclaration ? Les exigences légales dont Ex Libris fait mention recouvrent sans doute le droit d’auteur ; or cette notion appliquée aux millions de documents hors du domaine public ingérés par l’IA fait l’objet d’interprétations divergentes en Europe et aux États-Unis. Quelles clauses contractuelles, quelles garanties techniques une institution publique devrait-elle imposer à l’entreprise lui offrant un service grâce à des modèles propriétaires dont les données et méthodes d’entraînement sont inconnues ?

Plutôt que d’accepter pareille dépendance à des services externes et à des modèles surdimensionnés par rapport aux tâches à accomplir, la voie de la souveraineté et de la sobriété bien entendue est promue par de nombreux praticiens. Elle requiert de la part des bibliothèques un haut degré de compétence technologique (les enjeux éthiques et réglementaires sont liés aux outils choisis). Argie Kasprzik, responsable de l’IA de signalement à la ZBW, énonce des principes clairs : utiliser le plus petit modèle dont la performance suffit à la tâche envisagée, ne pas utiliser l’IA commerciale, évaluer et créer soi-même des jeux de données et adopter une démarche transparente¹⁰². La position de Géraldine Geoffroy constitue une feuille de route pour l’application technique de ces idées :

Notre conviction : l’avenir de l’IA dans les métiers documentaires repose sur des Small Language Models (SLM) ouverts – des modèles open, légers, CPU-friendly, déployables localement, qui garantissent la sécurité des données, permettent un fine-tuning sur mesure au plus près des besoins réels, et favorisent l’acculturation des équipes documentaires à l’usage raisonné de l’intelligence artificielle.¹⁰³

Le degré de maturité collective de la réflexion sur l’IA au sein de la profession rend les bibliothèques capables d’exercer la responsabilité publique que certains lui attribuent, celle de donner l’exemple à la société d’une utilisation réfléchie de l’IA :

*Publicly funded GLAM institutions, who work under resource constraints and face considerable pressure to demonstrate value to taxpayers, need to demonstrate strategic and proactive engagement with AI – both as a tool to be enrolled in the day-to-day work of GLAM practitioners, and as a phenomenon that GLAM institutions can help the broader public understand and interact with.*¹⁰⁴

⁹⁷ (COX, 2024, p. 162).

⁹⁸ Voir (KASPRZIK, 2025b, p. 3) et (SLATER, 2025, p. 601).

⁹⁹ (ANDERSON et TIAN, 2025, p. 365).

¹⁰⁰ Voir (MERGEL, SCHMIDT, SCHUMANN et al., 2024, p. 20) et, pour un cadrage plus général, (GRINBAUM, CHATILA, DEVILLERS et al., 2023, p. 21-22).

¹⁰¹ (AI Metadata Enrichment for Libraries 2024).

¹⁰² (KASPRZIK, 2025b, p. 12).

¹⁰³ (GEOFFROY [sans date]).

¹⁰⁴ (BERMAN, SMITHIES, TAIURU et al., 2025).

QUELLES PERSPECTIVES POUR L'EMPLOI DES AGENTS CONCERNÉS ?

FEUILLE DE ROUTE

Les pages qui suivent tentent d'exposer le point de vue des catalogueurs. Ils expriment leurs craintes dans de nombreux établissements, mais manifestent adhésion et enthousiasme s'ils peuvent contribuer à mettre en place une IA améliorant la qualité du travail fourni. Il en résulte pour eux la possibilité de concentrer leurs efforts sur les tâches à plus forte valeur ajoutée.

Liens des sections :

[↗ Craintes pour l'emploi](#)

[↗ Adhésion des agents](#)

[↗ Redéployer le travail](#)

3.1. Des craintes pour l'emploi

Dans les discours économiques sur l'automatisation, la crainte la plus immédiatement perceptible est qu'elle ne serve à supprimer ou dégrader des emplois plutôt qu'à transformer positivement le travail¹⁰⁵. L'IFLA voit là l'un des quatorze risques associés à l'IA : « *Threats of job displacement or exploitative employment relations.* »¹⁰⁶ Des méta-études sur l'IA en bibliothèque l'identifient elles aussi¹⁰⁷. Dans plusieurs enquêtes, des bibliothécaires expriment sans fard la crainte que l'IA soit comprise par les instances de pilotage comme un moyen de réduire les effectifs, alors même que les agents contribuent eux-mêmes à entraîner ou à évaluer des systèmes destinés à automatiser une partie de leurs fonctions¹⁰⁸. Cela vaut par exemple pour les bibliothèques nationales de d'Allemagne¹⁰⁹, de Belgique¹¹⁰, d'Estonie¹¹¹, des États-Unis¹¹², de Finlande¹¹³, de France¹¹⁴.

De fait, les gains de temps de traitement après adoption de l'IA sont parfois mis en exergue par les institutions. Sous la rubrique « *Streamline bibliographic metadata creation* », une diapositive de présentation de l'outil singapourien SGCAT souligne qu'il réduit de presque deux tiers le temps nécessaire à la description d'une notice : alors qu'il fallait 30 minutes, auparavant, 12 suffisent désormais¹¹⁵. Dans les bibliothèques de l'Université autonome du Chili, pendant les périodes de pic de travail, le nombre d'ouvrages catalogués après adoption d'un système reposant sur ChatGPT est passé de 15 à 45, entraînant une augmentation de « 300% » (sic) de la productivité¹¹⁶. Même lorsque l'IA est conçue comme mécanisme d'assistance et non d'automatisation totale, on s'inquiète d'un glissement : ce qui commence par le traitement d'un fonds que le manque de ressources humaines empêchait de traiter, peut, pour des raisons budgétaires, offrir une raison de ne plus recruter, et de réduire progressivement la part de l'expertise catalographique humaine. Ce risque est mis en avant par les détracteurs les plus fermes de l'IA en bibliothèque¹¹⁷, qui fustigent un attachement trop fort à la productivité des systèmes, au détriment de l'emploi humain : « *The deployment of AI technologies in information systems is often driven by a demand for efficiency and cost reduction, aligning with capitalist goals of maximizing profit and reducing human labor.* »¹¹⁸

En termes schématiques, on redoute donc que l'IA, en augmentant le rendement des agents, ne réduise la taille des services : cette inquiétude est d'autant plus aiguë qu'on fait souvent du signalement le cœur de l'identité professionnelle, en particulier pour une partie des personnels de catégorie B dans le

¹⁰⁵ (CASILLI, 2024, p. 4-10).

¹⁰⁶ (COX et DE BRASDEFER, 2025, p. 5).

¹⁰⁷ (PINAR et COX, 2025, p. 570).

¹⁰⁸ (SCHMIDT, KÜHLER, SCHUMANN et al., 2026, p. 188).

¹⁰⁹ (MERGEL, SCHMIDT, SCHUMANN et al., 2024, p. 18).

¹¹⁰ (MERGEL, SCHMIDT, SCHUMANN et al., 2024, p. 18).

¹¹¹ (SCHMIDT, 2026, p. 128).

¹¹² (MERGEL, SCHMIDT, SCHUMANN et al., 2024, p. 27).

¹¹³ (SCHUMANN et MERGEL, 2026, p. 50).

¹¹⁴ (BERMES, 2021a, p. 10) et (MERGEL, SCHMIDT, SCHUMANN et al., 2024, p. 17).

¹¹⁵ (LING et GOH, 2025, p. 9).

¹¹⁶ (DIAZ et VERGARA, 2025, p. 20).

¹¹⁷ (SLATER, 2025, p. 596).

¹¹⁸ (FREEMAN, 2025, p. 228).

contexte français¹¹⁹. Les éditeurs du numéro le plus récent de la revue *IFLA Journal* notent qu'il n'en va pas seulement de la gestion prévisionnelle des compétences, mais aussi de la nature même des emplois affectés : « *What are the implications of AI for librarians' work, knowledge, skills and other competencies, and for professional identities?* »¹²⁰ Comme le souligne Jean-Philippe Moreux, le catalogage est parfois perçu comme le « cœur d'identité cognitive de la bibliothèque » – toute opération qui toucherait à son statut a donc chance d'être particulièrement douloureuse pour les agents concernés¹²¹. Cette sensibilité particulière est bien identifiée : « *Most profoundly it [AI] is a threat of automation for a core professional form of knowledge/skill.* »¹²²

L'IA pourrait devenir le moyen de défaire un compromis déjà fragile autour de la valeur de ces activités et de la place qu'elles méritent. Le catalogage présente un caractère indispensable, mais souffre d'une dévalorisation relative au sein des établissements, dont les priorités se déportent vers d'autres domaines (services au public notamment)¹²³. La mutualisation du signalement par dérivation ou son externalisation par recours à des prestataires servent à réduire au mieux le temps qu'il prend aux équipes. À une époque où « tout est mis en œuvre pour diminuer la charge de travail représentée par le catalogage local, tant au niveau des agences bibliographiques (qui diffusent leurs métadonnées), des fournisseurs de notices (qui vendent les leurs) ou des établissements eux-mêmes (qui veulent consacrer une plus grande part de leurs forces vives à la médiation, notamment) »¹²⁴, les économies de temps que l'IA fait espérer mineraient un peu plus une position symbolique menacée. À quoi bon continuer d'accorder des ressources à une activité dont la machine s'acquitte fort bien ? Or pour l'instant l'implication de l'humain dans les projets réussis d'IA catalographique est forte : leur expertise permet de construire des systèmes fiables, et leur relecture, unitaire ou par échantillonnage, est indispensable. La croyance hâtive en une automatisation possible et souhaitable des chaînes de traitement risquerait de nuire à la qualité des métadonnées produites : « *However, given the rise of AI and the perception that it is able to solve specialized problems in cataloging easily, with the click of a button, if only the right prompt is created, is problematic to perpetuate – at least for now.* »¹²⁵.

3.2. Adhésion des agents : curiosité, prudence, co-construction

Dans certaines institutions étudiées, les catalogueurs se disent curieux des outils d'IA et ouverts à leur utilisation, parfois avec enthousiasme, s'ils s'insèrent harmonieusement dans leur pratique. Une certaine impatience est même rapportée à la KB¹²⁶, à la ZBW¹²⁷, à la DNB¹²⁸, ou à la BnF¹²⁹, plusieurs années pouvant s'écouler entre l'annonce d'un projet et sa mise en production. Cette ouverture n'est pas en contradiction avec des inquiétudes sur la permanence de l'emploi ou de sa valeur, mentionnées plus haut ; elle indique plutôt une condition d'acceptabilité. L'outil doit être conçu comme ancillaire, laissant au catalogueur un pouvoir effectif de décision et la perspective, naturellement, de conserver des missions intéressantes. Cela s'observe par exemple à la Bibliothèque du Congrès¹³⁰, dans le réseau Sudoc¹³¹, à la Bibliothèque nationale de Finlande¹³² et celle des Pays-Bas¹³³. Une telle attitude concorde avec des observations menées sur l'ensemble de l'économie étasunienne : la forme de relation avec l'IA privilégiée par les travailleurs est celle d'un partenariat collaboratif¹³⁴.

Cette acceptabilité est renforcée lorsque les agents sont consultés en amont et associés au développement de l'outil. Un questionnaire sur des dispositifs d'indexation semi-automatisée aux États-Unis, administré en 2022 à des experts de la communauté LD4 (données liées), souligne précisément l'importance de cette collaboration : elle améliore la compréhension et réduit la suspicion¹³⁵. Il en va de même dans l'expérience de l'Abes¹³⁶, à la Bibliothèque du Congrès¹³⁷, et c'est l'une des

¹¹⁹ (PHILIBERT, 2024, p. 62).

¹²⁰ (COX et WANG, 2025, p. 568).

¹²¹ (MOREUX, 2025a).

¹²² (COX et HAJIBAYOVA, 2025, p. 404).

¹²³ (PHILIBERT, 2024, p. 29 et 61-76).

¹²⁴ (CAVALIÉ, 2019b, p. 12).

¹²⁵ (MOULATON-SANDY et AND COBLE, 2024, p. 382).

¹²⁶ (BUESINK, 2025b).

¹²⁷ (KASPRZIK, 2025a).

¹²⁸ (MERGEL, KÜHLER, SCHUMANN et al., 2026, p. 35).

¹²⁹ (TULEU, 2025).

¹³⁰ (BRADOR, 2024, p. 5).

¹³¹ (NICOLAS, 2025b, p. 22).

¹³² (LEHTINEN, 2025b).

¹³³ (BUESINK, 2025b).

¹³⁴ (SHAO, ZHOU, JIANG et al., 2025, p. 12).

¹³⁵ (HAHN, 2024, p. 3-5).

¹³⁶ (NICOLAS, 2025a).

recommandations issues du projet LibrarIN : « *National Libraries should engage staff members who will use AI tools in the future to make sure the developed solution is trustworthy and efficient, that it meets the needs of its employees and enhances the library's operations. Engaging staff throughout the process will foster a collaborative environment and increase acceptance of the new technology.* »¹³⁸ Ateliers, votes, entretiens avec les catalogueurs sont autant de moyens de leur donner voix au chapitre. Ce serait là appliquer, à l'échelle des bibliothèques, les recommandations sociales générales de l'école d'économie du MIT sur la place des travailleurs dans la grande transformation en cours : « *The direction of AI will have profound consequences for all workers. Creating an institutional framework in which workers also have a voice would be a constructive step [...].* »¹³⁹ Intégrer les agents tôt dans le processus d'adoption de ces nouvelles technologies en favorise l'acceptation¹⁴⁰ ; la perception que les agents en ont dépend en partie de l'attitude de l'encadrement lui-même¹⁴¹.

Un autre motif d'adhésion est enfin l'apprentissage que les spécialistes peuvent tirer du recours à l'IA : un système de suggestions bien calibré peut permettre aux catalogueurs d'explorer des zones moins familières d'un vocabulaire contrôlé. Cette dynamique, parfois observée, renverse l'idée d'une technologie substitutive. Conçue avec finesse, elle peut devenir un instrument de formation et de consolidation de l'expertise. Certains bibliothécaires de la Bibliothèque nationale de Finlande se servent d'Annif quand le sujet d'un ouvrage les intéresse particulièrement ou au contraire qu'ils sont confrontés à un cas difficile¹⁴². Tel catalogueur a pu, grâce au prototype d'indexation Rameau de l'Abes, se détromper sur la signification d'un titre allemand¹⁴³, et les participants à l'expérimentation ont pour certains apprécié un outil qui « nourrit leur réflexion pour orienter leur recherche dans la base RAMEAU »¹⁴⁴. Un débutant qui s'en servirait en étant accompagné par son équipe pourrait découvrir plus facilement RAMEAU et la pratique de l'indexation¹⁴⁵. Plus généralement, cette situation bénéfique où l'humain apprend de la machine est un phénomène bien repéré¹⁴⁶.

Le pacte tacite d'acceptation n'est cependant pas infrangible. L'adhésion tient au fait que les outils actuels constituent des aides à la décision. Si le périmètre s'étend vers une automatisation plus complète, on imaginerait sans peine des résistances plus fortes se manifester¹⁴⁷.

3.3. Redéployer le travail humain

L'argument le plus fréquent en faveur de l'IA dans les chaînes de métadonnées est qu'elle présente une opportunité de redéployer du travail humain vers des tâches requérant une plus haute expertise¹⁴⁸. Le même raisonnement est employé à propos de l'impact de ces technologies sur l'économie en général : « *In the workplace, AI can free people from routine duties and allow them to take on more complex tasks that use their high-level skills.* »¹⁴⁹. Une partie des agents exprime une lassitude face au caractère répétitif de certaines activités de catalogage¹⁵⁰, qualifiées ici et là d'ennuyeuses¹⁵¹. Cela explique sans doute partiellement l'adhésion aux projets d'IA. Dans leur enquête auprès d'un vaste échantillon de travailleurs américains, des économistes de Stanford constatent que le souhait d'automatisation est d'autant plus vif que les tâches affectées sont répétitives ou ennuyeuses ; leur délégation à la machine est vue comme l'occasion de donner plus de temps à des activités à plus forte valeur ajoutée¹⁵². Il y aurait donc là l'opportunité de rendre le travail plus intéressant et agréable. Comme le dit ailleurs l'un des chercheurs de cette équipe : « *If AI is a substitute for the most routine and formulaic kinds of tasks, then by taking tedious routine work off human hands, AI may complement genuinely creative and interesting tasks, improving the basic psychological experience of work, as well as the quality of output.* »¹⁵³ Dans une classification des tâches selon le désir des agents de la voir automa-

¹³⁷ (POTTER, 2025).

¹³⁸ (SCHMIDT, KÜHLER, SCHUMANN et al., 2026, p. 187).

¹³⁹ (ACEMOGLU, AUTOR, et JOHNSON, 2023, p. 8).

¹⁴⁰ (AUTOR, MINDELL, et REYNOLDS, 2022, p. 40).

¹⁴¹ (ARMSTRONG, CHEN, CUELLAR et al., 2024, p. 3).

¹⁴² (LEHTINEN, 2025b).

¹⁴³ (NICOLAS, 2025a).

¹⁴⁴ (NICOLAS, 2025b, p. 15).

¹⁴⁵ (NICOLAS, 2025b, p. 22).

¹⁴⁶ (TE'ENI, YAHAV, ZAGALSKY et al., 2023, p. 2).

¹⁴⁷ (RUSSEIL-SALVÁN, 2024, p. 14).

¹⁴⁸ (COX, PINFIELD, et RUTTER, 2019, p. 4, 12).

¹⁴⁹ (TAVARES et JARAVEL, 2025, p. 28).

¹⁵⁰ (PHILIBERT, 2024, p. 29) citant un rapport de Catherine Aurérin sur l'évolution des tâches numériques des catégories B et C, mené en 2021 à la BnF.

¹⁵¹ (DE LA ROSA, 2025, p. 63), (SUSSMEIER et HENRY, 2025, p. 3).

¹⁵² (SHAO, ZOPE, JIANG et al., 2025, p. 8).

¹⁵³ (BRYNJOLFSSON et UNGER, 2023b, p. 24).

tiser et la possibilité technique d'y parvenir, une certaine fraction de celles constituant le signalement se trouverait généralement dans la zone « feu vert »¹⁵⁴, où l'humain *veut* déléguer et la machine *peut* accomplir les missions confiées.

Dans un scénario optimiste, la part de temps consacré à des gestes routiniers est reportée sur des opérations plus délicates de signalement où l'IA trébuche encore (choix d'indexation dans les cas ambigus, contrôle d'autorités, résolution de doublons épineuse), ou sur d'autres missions des bibliothèques¹⁵⁵. Cette opinion est exprimée par des organismes internationaux¹⁵⁶ et par les responsables des services concernés. Ainsi, Avaha Cohen, responsable de la section de catalogage à la Bibliothèque nationale d'Israël, voit dans l'automatisation l'opportunité de réduire le temps passé par des catalogueurs hautement qualifiés à saisir à la main des informations¹⁵⁷. L'avantage de Retrotool, l'outil conçu par la KB, est qu'il s'acquitte des tâches répétitives, « *allowing the cataloguers to focus more on complex cataloguing aspects* »¹⁵⁸. L'opinion est la même en Belgique¹⁵⁹, en Suède¹⁶⁰, à la Bibliothèque du Congrès¹⁶¹, à Singapour¹⁶². L'argument est aussi mis en avant par les vendeurs de logiciels, comme Ex Libris¹⁶³, ou WorldCat vantant les mérites du dédoublonnage automatisé¹⁶⁴ : « *This approach to de-duplication is not about reducing the role of people—it's about refocusing their expertise where it matters most. Catalogers can focus on high-value work that connects them to their communities instead of spending hours resolving duplicate records.* » Cela suppose de montrer aux équipes les missions qu'elles pourraient exercer si le catalogage recule, de souligner la continuité du métier en profondeur et même les avantages que la transformation peut apporter : ainsi, à l'Université Paris-Saclay, si l'IA était adoptée, il est envisagé que le centre de gravité se déplace du signalement traditionnel vers la science ouverte et la valorisation des données de la recherche, avec une relation plus directe au chercheur très valorisante¹⁶⁵. L'enjeu foncier pour le professionnel resterait bien, dans un tel cas, la facilitation de l'accès à la production scientifique.

Lorsque les projets atteignent une maturité suffisante pour passer d'une assistance à une automatisation partielle, on remarque que des rééquilibrages ont lieu. Ainsi à la ZBW, des profils techniques de *data librarian* ont remplacé quelques spécialistes de l'indexation matière après la mise en production d'AutoSE en 2021 – néanmoins quelque 20 000 notices font encore l'objet d'une indexation intellectuelle chaque année, et nous verrons plus bas que l'institution doit conserver en son sein des experts pour vérifier le modèle et le faire évoluer. Le rééquilibrage des forces peut aussi se produire en amont, afin de constituer l'équipe qui s'attellera au projet d'IA. À la DNB, pour trouver les ressources humaines nécessaires à la création d'un système d'indexation automatique des documents numériques, il a fallu graduellement réduire celles allouées à l'indexation des imprimés, destinés à être eux aussi traités sans intervention humaine quand la technique le permettrait¹⁶⁶. Comme l'explique Benoît Tuleu, l'automatisation ne supprime pas nécessairement des emplois, car il s'en crée d'autres pour rendre des services nouveaux¹⁶⁷. Cette équation suscite plusieurs questions, auxquelles les parties suivantes apportent quelques éléments de réponse. Quelles compétences seraient nécessaires pour contribuer utilement, en tant que bibliothécaire, à une chaîne de métadonnées partiellement automatisée ? Comment une institution peut-elle organiser au mieux la collaboration entre experts en IA et spécialistes du catalogue ? Si la part du signalement diminue parmi les postes futurs, comment se déroulera la transition pour les catalogueurs aujourd'hui en poste ?

¹⁵⁴ (SHAO, ZOPE, JIANG et al., 2025, p. 10).

¹⁵⁵ (COX, 2022, p. 1) et (COX, 2024, p. 160).

¹⁵⁶ (COHEN, ALIVERTI, BARBUS et al., 2025, p. 1), (PCC TASK GROUP ON AI AND MACHINE LEARNING FOR CATALOGING AND METADATA, 2025p. 2).

¹⁵⁷ (COHEN, 2025).

¹⁵⁸ (MERGEL, SCHMIDT, SCHUMANN et al., 2024, p. 18).

¹⁵⁹ (MERGEL, SCHMIDT, SCHUMANN et al., 2024, p. 21).

¹⁶⁰ (MERGEL, SCHMIDT, SCHUMANN et al., 2024, p. 26).

¹⁶¹ (MERGEL, SCHMIDT, SCHUMANN et al., 2024, p. 27).

¹⁶² (LING et GOH, 2025, p. 3) et (JAILANI, 2025).

¹⁶³ (*AI in Academic Libraries : Trusted Solutions by Ex Libris* 2024).

¹⁶⁴ (RAJAPATIRANA 2025).

¹⁶⁵ (BELLIER et MINQUILAN, 2025).

¹⁶⁶ (JUNGER et SCHÖLZE, 2021, p. 60).

¹⁶⁷ (TULEU, 2025).

UNE PERMANENCE INDISPENSABLE DE L'EXPERTISE CATALOGRAPHIQUE

FEUILLE DE ROUTE

Cette troisième partie reste concentrée sur le catalogueur ; elle montre que, même après la mise en production d'une IA, son expertise demeure cruciale au cours du signalement et pendant tout le cycle de vie du système, notamment grâce à ses compétences de data librarian. Les normes catalographiques ne semblent pas devoir disparaître, tant s'en faut, et les plus futuristes des chaînes de traitement envisagées, chaînes agentiques où l'IA jouit d'une large autonomie, nécessiteront un travail d'explicitation méticuleux de ce corpus de conventions réglant la pratique du signalement. Le cas de l'indexation matière laisse croire à la permanence des méthodes en vigueur jusqu'à présent pour la description des ressources, et donc à l'importance de l'expertise accumulée.

Liens des sections :

↗ Implication de l'expert dans le signalement

↗ L'expertise au cœur du cycle de vie des systèmes

↗ Un précédent favorable : le data librarian

↗ Vers une nouvelle application de l'expertise

↗ Avenir de l'indexation

4.1. Implication de l'expert dans le signalement

Dans les projets observés, la figure de l'expert est loin de s'effacer. Le taux d'erreurs encore élevé des systèmes d'IA, leur manque de précision et de régularité, sont autant d'obstacles fondamentaux à l'automatisation pleine et entière selon Abigail Potter¹⁶⁸, de la Bibliothèque du Congrès. Le *National Library Board* de Singapour, qui recourt à l'IA depuis 2024 pour accélérer le catalogage des acquisitions absentes de WorldCat ou pour l'indexation des archives nationales et des archives du web, souligne la nécessité de maintenir un humain dans la boucle¹⁶⁹. Plus précisément, les concepteurs de SGCAT, l'outil de signalement à partir des métadonnées fournies par les éditeurs, mettent en garde : « *AI is not perfect. Human participation is still required.* »¹⁷⁰ Mona Lehtinen, de l'équipe Annif, doute que les LLM se hissent jamais au niveau humain pour l'indexation :

*AI keeps getting better, of course, but I think there will be a threshold of what it can actually do. And I'm somehow skeptical of it ever reaching the level of a professional human indexer. The task of subject indexing is almost an art form. It takes an experience and a certain knowledge of the world, it takes a certain knowledge of the field, it takes a certain knowledge of the vocabulary, and I don't think LLMs are fully ever going to be capable of doing that without the emergence of general AI.*¹⁷¹

Ce primat donné au catalogueur tient au constat pragmatique des insuffisances perçues ou constatées de la machine, mais surtout à une axiologie fortement ancrée dans la profession, qui accorde un rang élevé au travail humain. Les catalogueurs sont consultés dès la conception des systèmes. « *We wanted to center the catalogers in this experiment, include their expertise from the beginning, and allow their decisions and real-life workflows to set our direction.* »¹⁷² Cette déclaration d'Abigail Potter est représentative de l'attitude observée à peu près sans exception, par exemple dans la description de Retrotool à la KB¹⁷³.

¹⁶⁸ (POTTER, 2025).

¹⁶⁹ (JAILANI, 2025, p. 7).

¹⁷⁰ (LING et GOH, 2025, p. 10).

¹⁷¹ (LEHTINEN, 2025b).

¹⁷² (BRADOR 2024, p. 2) et (POTTER, 2025).

¹⁷³ (VAN DER VAART et SCHMIDT, 2026, p. 102-104).

Que les systèmes d'IA collaborent avec l'humain au lieu de l'évincer est une exigence prioritaire au-delà des bibliothèques¹⁷⁴. Les économistes du *Stanford Institute for Human-Centered*, alignés avec ceux du MIT, cherchent comme on l'a vu à orienter les politiques publiques vers une IA complémentaire du travail humain : « *How can policies encourage the types of AI that complement human labor instead of imitating and replacing it?* »¹⁷⁵

Le catalogueur ne se voit donc en aucun cas imposer le recours à l'IA, et encore moins ses prédictions : « *It is critical that human catalogers retain the ability to decide whether to use AI tools and accept their outputs as an integral part of metadata work.* »¹⁷⁶ De pareilles déclarations sont monnaie courante et l'unanimité est à peu près complète sur ce point¹⁷⁷. L'ambition des projets étudiés est d'étendre les capacités du catalogueur, non pas de le remplacer.

Il arrive qu'il se voie attribuer un apanage de tâches où la machine n'est pas encore assez efficace, apanage dont les contours changent selon les projets. Les publications anglo-saxonnes y rangent le catalogage réparateur ou attentif aux spécificités culturelles, comme l'indexation dans un vocabulaire contrôlé Māori¹⁷⁸. Il comprend, dans la présentation de l'outil SGCAT, la création des liens à d'autres entités ou l'indexation¹⁷⁹. C'est parfois la spécificité d'un corpus, dont les caractéristiques s'éloignent des données d'entraînement du modèle, qui délimite ce pré carré humain. Ainsi dans l'expérimentation menée par l'Abes, les établissements spécialisés comme le MNHN ou l'EFR tiraient un moindre profit des suggestions automatiques d'indexation Rameau¹⁸⁰. Néanmoins, comme le souligne Yann Nicolas, les tâches confiées à l'IA empiètent de plus en plus sur le domaine qu'on voudrait parfois réserver au catalogueur. L'indexation automatique à la ZBW est une réalité, et un LLM consultant une API au sein d'une chaîne agentique (Ordo, projet de la bibliothèque de la MMU) ou d'un circuit d'exécution (KBR) peut créer un point d'accès, en se fondant sur les publications de l'entité candidate. Cette opération de liage requiert, selon l'expression d'Ahava Cohen, une « codification de l'instinct humain »¹⁸¹ ; la Bibliothèque nationale d'Israël aussi y travaille.

Mais quand bien même certaines tâches érigées en *nec plus ultra* insurmontable pour l'IA deviendraient automatisables, la faisabilité technique n'est qu'un paramètre du maintien de l'humain dans la chaîne de signalement. Car l'autonomie du catalogueur est biface : elle tient aussi bien à la reconnaissance de l'expertise humaine qu'à la nécessité de faire porter la responsabilité des métadonnées produites sur quelqu'un, plutôt que sur ce *quelque chose* opaque qu'est un algorithme. La double origine de la centralité accordée à l'expert est manifeste dans la politique d'usage responsable de l'IA à la Bibliothèque nationale d'Autriche : « *As far as possible, we ensure the quality of results produced by the AI systems we use through human involvement. Our employees strive to ensure quality through regular reviews of AI results and process monitoring. Final decisions, especially when artificial intelligence is used in the decision-making process, always remain the responsibility of humans.* »¹⁸²

Il est parfois envisagé de laisser un plus haut degré d'initiative à l'IA. Dans le système agentique Ordo au Danemark, le catalogueur est sollicité lorsque le LLM est incertain de sa réponse¹⁸³. Dans les termes d'une récente taxonomie des interactions entre humains et LLM¹⁸⁴, on passerait à un régime où la rétroaction humaine guide la machine plutôt qu'elle ne l'évalue, pendant l'accomplissement de la tâche et non après. L'idée de donner à la machine un rôle plus actif et de constituer l'expertise en ressource mise à sa disposition se retrouve dans des projets qui n'intègrent pas encore de LLM. L'équipe d'AutoSE à la ZBW prévoit ainsi de permettre à l'IA de demander aux experts des jeux de données annotés ou des évaluations de cas spécifiques¹⁸⁵.

4.2. L'expert au cœur du cycle de vie des systèmes

La participation humaine est donc cruciale au cours du signalement. Elle ne l'est pas moins pour préserver une capacité interne à comprendre, vérifier, corriger et faire évoluer les outils, pendant

¹⁷⁴ (FISHER, 2025, p. 6-7).

¹⁷⁵ (BRYNJOLFSSON et UNGER, 2023b, p. 25).

¹⁷⁶ (PCC TASK GROUP ON AI AND MACHINE LEARNING FOR CATALOGING AND METADATA, 2025b, p. 5).

¹⁷⁷ (ENGEL, DO, SALEM et al., 2025, p. 8).

¹⁷⁸ (SUSSMEIER et HENRY, 2025, p. 13).

¹⁷⁹ (LING et GOH, 2025, p. 9).

¹⁸⁰ (NICOLAS, 2025a).

¹⁸¹ (COHEN, 2025).

¹⁸² (*Policy Artificial Intelligence* [sans date], p. 2). Nous soulignons.

¹⁸³ (MALMSTEN et LUNDBORG, 2025).

¹⁸⁴ (ZOU, HUANG, WU et al., 2025b, p. 3).

¹⁸⁵ (KASPRZIK, 2024a, p. 164), (KASPRZIK, 2024b, p. 210).

leur élaboration et après. La contribution des catalogueurs est indispensable pour affiner les prédictions du modèle, et cette démarche de comparaison entre signalement humain et automatique depuis la construction du prototype jusqu'au contrôle qualité postérieur à la mise en production s'observe dans tous les projets considérés. À travers les recommandations des organismes bibliothéconomiques, les rapports d'expérience et les entretiens se fait entendre une même exigence : l'expertise métier doit gouverner le cycle de vie du système, y compris une fois celui-ci mis en production. « *The PCC believes that catalogers and other metadata professionals have valuable expertise that plays an important role in reviewing, testing, and assessing AI features and tools in the library landscape.* »¹⁸⁶ Le point de vue individuel d'une *Metadata Librarian* aux États-Unis reflète la même conviction : les bibliothécaires joueront un rôle décisif dans l'entraînement et l'évaluation des modèles, ainsi que dans leur intégration dans les circuits de traitement¹⁸⁷. Le projet d'assistance à l'indexation de la Bibliothèque nationale d'Estonie est ainsi représentatif de l'approche générale : « *Librarians and cataloguers are actively involved in training, validating, and improving AI-generated metadata. Their involvement shall ensure that the AI-based indexing will be aligned with traditional cataloguing methodologies and provide real-world feedback to developers.* »¹⁸⁸

Cela peut se faire de plusieurs manières : sessions de travail spécifiques, télémétrie, intégration d'un module dédié dans la chaîne de traitement. Dès 2019 Annif organisait des ateliers de revue¹⁸⁹, et le Labo de l'Abes a procédé selon la même méthode, s'appuyant sur l'expertise catalographique de certains membres de l'établissement pendant l'élaboration de l'outil¹⁹⁰, puis sur les statistiques d'usage des bibliothèques pilotes pendant l'expérimentation¹⁹¹. La ZBW s'est en outre munie d'une interface intégrée à l'outil de catalogage, permettant de noter chacune des propositions faites par l'outil AutoSE¹⁹² ; il est prévu que ces évaluations servent à l'amélioration continue des modèles¹⁹³. À la DNB, un corpus de quelques centaines de notices revues par les catalogueurs sert d'étalon pour le contrôle qualité, y compris après passage en production, ce qui donne au travail expert une fonction à la fois d'audit et de calibrage¹⁹⁴.

Il est intéressant de noter que dans ces deux derniers cas, l'évaluation des suggestions par les catalogueurs confrontés à des cas réels *en production* servent à l'entraînement ultérieur des modèles. Grâce à cette boucle de rétroaction, leur précision s'accroîtra. Ce principe pourrait s'étendre à un système d'IA agentique, avec une temporalité différente (le correcteur repasserait sur des notices après leur génération) : « *Training the agent – by having a human review and correcting AI generated records we are giving the system a high-quality dataset, allowing it a chance to see high quality records and 'self-learn' by embedding these into the vector database with domain knowledge.* »¹⁹⁵ Le catalogage, outre son objectif premier, en accomplit simultanément un second, celui d'entraîner la machine. Ce glissement de finalité était aussi observé dans l'étude ethnographique mentionnée dans l'introduction, portant sur une compagnie de fret maritime : « *Another extant role, the manual work of classifying data, was not replaced, but repurposed to provide the auditing process with a ground truth against which the algorithmic outcome was compared.* »¹⁹⁶

Le cadre de planification à l'IA de la Bibliothèque du Congrès souligne la nécessité de surveiller la précision du modèle au cours du temps (à mesure que les données sur lesquelles il est entraîné changent), de s'assurer qu'il continue de fonctionner conformément à ce que les utilisateurs en attendent¹⁹⁷. Cela impose à l'établissement d'établir des critères de qualité et une méthodologie claire pour l'évaluer¹⁹⁸. Cette préconisation est relayée dans les conclusions du projet LibrarIN : « *It is therefore important—not just during the AI training phase—to keep human specialists' knowledge and expertise "in the loop" to continue to monitor the quality of the automatically generated output.* »¹⁹⁹ L'expérience de la DNB montre aussi que cette phase subséquente à la mise en production initiale ne doit pas être négligée : « *Die Etablierung maschineller Erschließungsverfahren in der DNB war und ist*

¹⁸⁶ (PCC TASK GROUP ON AI AND MACHINE LEARNING FOR CATALOGING AND METADATA, 2025b, p. 4).

¹⁸⁷ (ANDERSON et TIAN, 2025, p. 368).

¹⁸⁸ (SCHMIDT, 2026, p. 126).

¹⁸⁹ (SUOMINEN, INKINEN, et LEHTINEN, 2022, p. 272-273).

¹⁹⁰ (NICOLAS, 2025a).

¹⁹¹ (NICOLAS, 2025b, p. 11-13).

¹⁹² (KASPRZIK, 2024a, p. 162-164).

¹⁹³ (KASPRZIK, 2025a), (KASPRZIK, 2024cp. 12).

¹⁹⁴ (POLEY, UHLMANN, BUSSE et al., 2025, p. 18-19).

¹⁹⁵ (KORPET et REES, 2025, p. 32).

¹⁹⁶ (GRØNSUND et AANESTAD, 2020, p. 15).

¹⁹⁷ (MANCHESTER 2023).

¹⁹⁸ (SACCUCCI et POTTER, 2025, p. 12).

¹⁹⁹ (SCHMIDT, KÜHLER, SCHUMANN et al., 2026, p. 190).

kein einmaliger, abgeschlossener Vorgang. Vielmehr sind die laufende Fortentwicklung und Pflege maschineller Erschließungsverfahren zu neuen Daueraufgaben geworden. Wie alle Verfahren bibliothekarischen Arbeitens müssen sie regelmäßig auf den Prüfstand gestellt werden, auch, um neue (wissenschaftliche) Erkenntnisse aufzunehmen und umzusetzen.»²⁰⁰ Le contrôle qualité se fait par échantillonnage, en partant de vedettes-matière choisies aléatoirement, de publications parallèles (l'indexation automatique de la version électronique est mise en regard de l'indexation humaine de la version imprimée), ou d'un choix général de publications indexées automatiquement sur lesquelles se penchent des catalogueurs pour noter la pertinence des vedettes-matière²⁰¹. Hannes Lowagie décrit aussi l'importance de cette étape :

*Regularly reviewing and fine-tuning the dataset is crucial for refining the model's accuracy. This task demands significant time and resources as every update requires careful analysis and judgement. Despite the challenges, investing time in manually adjusting the training set can help ensure that the model learns from its mistakes, gradually reducing error rates and leading to a more reliable, accurate system over time. This ongoing cycle of monitoring, error correction and refinement is what ultimately transforms a good model into a highly effective one.*²⁰²

La même mise en garde se retrouve dans d'autres domaines – après sa mise en production, un LLM doit être testé, éventuellement ré-entraîné et perpétuellement surveillé pour garantir le niveau de performance escompté²⁰³. Cela vaut aussi pour l'apprentissage machine traditionnel.

L'expertise n'est pas seulement requise pour maintenir la machine sur les rails posés initialement, mais pour faire face à la nouveauté, comme le souligne la Bibliothèque Nationale d'Estonie²⁰⁴ à propos de l'indexation. Les vocabulaires contrôlés évoluent (réforme Rameau, FAST, des concepts apparaissent. Or les outils d'IA les plus répandus aujourd'hui pour l'indexation et les projets les plus ambitieux qui s'en servent (Annif, AutoSE) relèvent de l'apprentissage machine, et sont donc limités par les données d'entraînement. Ils dépendent d'un apport humain pour intégrer de nouvelles catégories ou de nouveaux usages (exemples de corpus d'entraînement ou de retours d'annotation) comme l'explique l'équipe d'Annif : « *In addition, the system usually needs curated training data in which a new term added to the vocabulary appears enough times for it to learn to use this new term correctly. Training material is the stumbling block of all machine learning and artificial intelligence: if the training material is incomplete or distorted in some way, the algorithms will not perform as expected.* »²⁰⁵ La nouveauté peut aussi venir de l'apparition de nouveaux formats de document nécessitant l'établissement de normes de signalement²⁰⁶. Cette perpétuelle adaptation au changement du champ catalographique empêche d'imaginer un outil figé *ne varietur*, même pour les chaînes les plus modernes, comme celle que décrit la Bibliothèque universitaire métropolitaine de Manchester : « *We will still need to retain staff to oversee the operation of the system. Records need to be checked, new standards implemented, books physically labelled, new training documents written and processed into the system's knowledge base. We should also monitor the functioning of the system.* »²⁰⁷

4.3. Un précédent favorable : le *data librarian*

Jusqu'à présent nous avons insisté sur la contribution des bibliothécaires à l'amélioration des modèles grâce à leur expertise catalographique. Ils sont amenés à corriger des notices traitées par l'IA, qui serviront à la recalibrer, ou à traiter des notices de documents peu ou pas représentés dans le corpus d'entraînement pour élargir le périmètre de compétence de l'outil (par exemple en indexant des ouvrages avec de nouvelles vedettes-matière). Un second atout, la compétence informatique, étend cette sphère d'action. En 2016, alors directeur du Département de l'information bibliographique numérique de la BnF, Gildas Illien écrivait de la Transition bibliographique qu'elle augmenterait la proportion de dérivations de métadonnées à la source, mais que la gestion de celles-ci demanderait de nouvelles compétences : « Le *data librarian* devra maîtriser les processus de traitement automatisé

²⁰⁰ (JUNGER et SCHOLZE, 2021, p. 68).

²⁰¹ (JUNGER, 2018, p. 7).

²⁰² (LOWAGIE, 2025a, p. 67).

²⁰³ (ZOU, HUANG, WU et al., 2025a, p. 7).

²⁰⁴ (SCHMIDT, 2026, p. 128).

²⁰⁵ (Semi-Automatic Subject Cataloging Using Finto AI [sans date]).

²⁰⁶ (COHEN, ALIVERTI, BARBUS et al., 2025, p. 2).

²⁰⁷ (KORPET et REES, 2025, p. 32).

des données, mais être aussi en mesure d'analyser leur structure, leur qualité, leur contenu.»²⁰⁸ De même, un projet d'IA profitera grandement de l'appui d'un bibliothécaire capable d'imaginer des critères pour isoler des notices susceptibles d'améliorer le modèle, puis d'extraire du catalogue les lots correspondants, avant de leur appliquer une cascade d'éventuelles modifications et de les remettre aux mains des ingénieurs. Dans un ouvrage récent de Véronique Mesguich sur les données en bibliothèque, un chapitre est consacré à leur analyse et réutilisation grâce à des outils informatiques idoines, un autre à la figure du *data librarian*, signe de l'ancrage de cette pratique dans la profession en France²⁰⁹. Toutes les configurations techniques examinées, depuis l'assemblage d'IA variées sur une plateforme propriétaire en usage à la KBR jusqu'à des chaînes à IA agentique au Danemark ou à Manchester, requièrent initialement des quantités, plus ou moins vastes de notices ; le PCC dresse la liste détaillée des étapes de sélection et préparation de ces jeux de données, pour souligner le rôle que les bibliothécaires doivent jouer à chacune d'entre elles²¹⁰. On trouve la même idée dans une étude récente mettant en regard les compétences énumérées dans les référentiels des bibliothécaires et dans ceux des praticiens de l'IA²¹¹.

La culture de certains établissements, tributaire de celle de leurs pays, conduit parfois à un taux d'adoption de l'informatique élevé chez les bibliothécaires. Il en est ainsi à la Bibliothèque nationale d'Israël : « *Within the library, management urges cataloguers not to see themselves as typists, but to use tools for batch corrections, OCR of documents, and similar tasks. New hires are encouraged to view technology not as a replacement, but as a way to handle repetitive work and save time.* »²¹² Lors des recrutements, Ahava Cohen recherche des profils dotés de l'esprit de système, curieux de la logique sous-jacente à leur activité et aux données qu'ils manipulent. Un bibliothécaire chargé des disques a ainsi créé un agent pour récupérer leur description sur un site de musique et la transformer en notice, un autre a construit une chaîne d'automatisation pour traduire des notices d'ouvrages français trouvées sur WorldCat du français à l'anglais. « *The expectation is to go beyond your job description.* »²¹³ Cette autonomie laissée dans l'exploration et la création d'outils d'automatisation légère, ciblant des tâches précises, rend les bibliothécaires certainement plus aptes à contribuer à des projets de grande envergure comme un assistant d'indexation ou de description bibliographique similaire à celui de la KBR.

L'avantage de la profession ne se limite pas à l'agilité que certains de ses membres ont acquis, selon les contextes institutionnels et les propensions individuelles, dans la manipulation des données. Elle tient aussi, plus généralement, à l'habitude de réfléchir en des termes fortement compatibles avec ceux qui entourent l'entraînement d'un modèle : qualité des données, représentativité, biais de sélection, normalisation, etc. Les bibliothécaires savent « penser données », avec la prise de distance et la réflexion critique que cela suppose²¹⁴. Cette adéquation a déjà été montrée, sous plusieurs angles : « *What does emerge is that since AI today is data driven, high level data skills are necessary, and the pre-existing knowledge and competencies of librarians and information professionals are of immense value, including knowledge of library-specific data for library projects or contributions to data governance and stewardship in wider institutional applications of AI.* »²¹⁵ L'étude comparative citée plus haut, qui étudie la façon dont les bibliothécaires peuvent contribuer à des projets d'IA, portent le même jugement : « *Given long-standing traditions in metadata design, search, and information retrieval, librarians are again well-positioned to assist in the development of robust search frameworks to help feed data to AI systems, grounding them and mitigating hallucination.* »²¹⁶

4.4. Vers une nouvelle application de l'expertise

Il y aurait donc bien persistance de l'expertise. Le cœur des métiers catalographiques réside dans la maîtrise fine d'écosystèmes normés (formats, codes, règles locales d'application, autorité, vocabulaires contrôlés)²¹⁷. Le passage progressif à une chaîne de signalement où l'IA, entraînée et perfectionnée grâce au travail humain, remplit une part importante de la notice, constitue une application nouvelle des mêmes compétences fondamentales. On trouverait là confirmation d'une enquête de 2019

²⁰⁸ (ILLIEN, 2016).

²⁰⁹ (MESGUICH, 2023, p. 111-125, 141-154).

²¹⁰ (PCC TASK GROUP ON AI AND MACHINE LEARNING FOR CATALOGING AND METADATA, 2025b, p. 24-25).

²¹¹ (HARPER et GROTH, 2025, p. 555).

²¹² (COHEN, 2025).

²¹³ (COHEN, 2025).

²¹⁴ (MESGUICH, 2023, p. 149).

²¹⁵ (COX, 2024, p. 66).

²¹⁶ (HARPER et GROTH, 2025, p. 554).

²¹⁷ (PHILIBERT, 2024, p. 38).

auprès de directeurs de bibliothèques, laquelle suggérait que les rôles de contrôle de la qualité des données et curation des données dérivées que les établissements assument déjà pourraient se trouver renforcer par l'IA²¹⁸.

Une nouvelle orientation se dessine peu à peu. Par sa compréhension des données et des structures où elles se coulent, par sa familiarité croissante avec l'informatique et les LLM, le bibliothécaire est dans une position privilégiée pour expliciter les règles de signalement à suivre dans l'établissement. À la BnF, par exemple, les consignes de catalogage fournies par Kitcat stipulent qu'il ne convient pas de créer un point d'accès pour l'auteur d'une préface si elle est longue de moins de six pages²¹⁹. La connaissance de ces principes est indispensable à la prochaine génération d'outils, dont la chaîne agentique prévue par la bibliothèque de la MMU donne une idée ; un LLM consulte une documentation vectorisée (RAG) pour cataloguer en adhérant au mieux aux consignes locales²²⁰. Cela suppose que ces principes soient disponibles dans un format aussi facile à exposer à l'humain ou à la machine. On disposerait alors d'une documentation favorisant l'interopérabilité des données, car un établissement qui les emploierait pourrait consulter les règles sous-jacentes à leur production. Cette vision a fait l'objet d'une présentation de Hannes Lowagie lors du dernier congrès de l'IFLA²²¹.

4.5. Avenir de l'indexation

4.5.1. Un primat humain contesté

Insister sur la centralité de la compétence humaine n'empêche pas de souligner les divergences de pratique entre professionnels, surtout dans le domaine de l'indexation. Deux catalogueurs donnés n'utiliseront pas les vedettes d'un vocabulaire contrôlé exactement de la même façon : « *As soon as more than one person is involved in classifying a subject area or searching for information, it becomes apparent that people have slightly, or wildly, differing ideas about the nature of a subject, both when choosing a heading within the system to apply to an item, or when using it to search.* »²²² L'indexation peut manquer d'exhaustivité ou de spécificité²²³. Et comment en serait-il autrement, disent en substance les spécialistes cités, dès lors que les catalogueurs sont humains, donc faillibles, et que les vocabulaires contrôlés sont vastes et complexes ?

Ces erreurs et hétérogénéités de l'indexation sont parfois présentées comme de bonnes raisons de recourir à l'IA²²⁴. La Bibliothèque nationale d'Estonie mentionne cet argument : « *The quality of description depends on the intellectual assumptions of a cataloger and is affected by some amount of subjectivity.* »²²⁵ Les vedettes-matière assignées ne sont pas toujours les plus précises, objectives, utiles, ni cohérentes, lit-on dans le compte rendu LibrarIN des entretiens menés avec cette institution²²⁶. L'équipe d'Annif rapporte que l'une des observations faites lors d'un atelier avec des catalogueurs en Finlande portait sur l'harmonisation accrue de l'indexation : « *Even professional indexers do not always necessarily agree on what the perfect description of the work; Annif could help with that. It might streamline the process and make the results more similar.* »²²⁷ Au-delà de l'indexation, l'IA présente un intérêt pour la précision et fiabilité des métadonnées en général²²⁸. Hannes Lowagie y voit l'opportunité d'améliorer « *efficiency, accuracy, and consistency in cataloging* »²²⁹ et dans son prototype prometteur de système agentique, la bibliothèque de la MMU observe que des champs habituellement négligés par l'humain se trouvent remplis²³⁰.

Ce qu'on note, c'est que l'insertion de l'IA dans une chaîne de traitement entraîne une réflexion sur la notion de qualité et un recalibrage des exigences : « *How is quality regarding subject cataloguing defined and what are proper criteria and instruments for measurement?* »²³¹ L'objectif de rendre les collections découvrables peut primer le respect des règles de catalogage, la fin justifiant les moyens.

²¹⁸ (COX, PINFIELD, et RUTTER, 2019, p. 14-15).

²¹⁹ (KOSKAS, 2025).

²²⁰ (KORPET et REES, 2025, p. 35).

²²¹ (LOWAGIE, 2025c).

²²² (MALMSTEN, LUNDBORG, FANO et al., 2025, p. 181).

²²³ (GOLUB, 2021, p. 712).

²²⁴ (CHOW, KAO, et LI, 2024, p. 575).

²²⁵ (ASULA, MAKKE, FREIENTHAL et al., 2021, p. 1).

²²⁶ (SCHMIDT, 2026, p. 124).

²²⁷ (LEHTINEN, 2025b).

²²⁸ (COHEN, ALIVERTI, BARBUS et al., 2025).

²²⁹ (LOWAGIE, 2025a, p. 8).

²³⁰ (KORPET et REES, 2025, p. 32).

²³¹ (JUNGER, 2018, p. 8).

Dans un entretien consigné par LibrarIN, un responsable de la DNB explique ce renversement : « *And the discussion as to whether the results of the automated processes are good enough was also conducted from the point of view of whether the bibliographic rules are being adhered to. No, they are not adhered to, but the aim of the automated processes was exclusively to find and support.* »²³² C'est la rencontre de deux milieux professionnels, celui des catalogueurs et des informaticiens ou spécialistes d'apprentissage machine, de statistique et science des données, autour d'un projet commun qui amène de telles discussions. Les attentes d'exactitude des uns et les méthodes de mesure des autres (entre autres la précision, le rappel, les scores de F1 et pAUC²³³) peuvent être difficiles à concilier. Argie Kasprzik insiste sur la nécessité d'un meilleur partage de la responsabilité entre les deux parties sur cette question²³⁴.

4.5.2. Maintien des vocabulaires contrôlés

Plutôt que de couler le projet d'indexation automatisée dans le moule préexistant, il peut être tentant de s'affranchir de ses imperfections²³⁵. Citons à nouveau la même autorité :

*Automatization solutions for subject indexing can either aim to imitate those rules as closely as possible or they can aim to exploit the full potential of the technologies that are currently available. The latter can lead to a situation where established subject indexing rules should ideally be reevaluated and possibly changed or dropped (e.g., because they introduce information which is no longer needed due to new retrieval techniques), which may require some change management.*²³⁶

Le KBLab de la Bibliothèque nationale de Suède adopte la même position. Ses membres remarquent que les vedettes-matières pertinentes pour l'indexation d'un document sont parfois dispersées dans les différents domaines d'un vocabulaire contrôlé, sans lien aucun. À l'heure de la recherche sémantique, on ne doit plus se contenter d'un tel système, mais chercher « *a leaner subject heading system that would be easier for both humans and AIs to apply consistently.* »²³⁷ La Bibliothèque du Congrès, dans son expérimentation, songe à employer FAST plutôt que LCSH, postulant que l'apprentissage machine donnera de meilleurs résultats avec des chaînes d'indexation construites *ex post* plutôt qu'*a priori*²³⁸.

Les projets étudiés se greffent sur une chaîne de traitement existante ou la prolongent, et composent donc avec la structure intellectuelle et organisationnelle existante : les modèles sont entraînés sur des notices indexées grâce au vocabulaire contrôlé en vigueur, et leurs résultats sont notés par des catalogueurs détenant la maîtrise des vedettes-matière assignées. Cette évaluation s'intègre parfois finement aux circuits de signalement en place²³⁹. Le souci de maintenir l'homogénéité de l'indexation entre les corpus traités par la machine et ceux traités par l'humain, ce qui garantit une expérience de recherche uniforme pour le lecteur final, dissuade de faire table rase du passé. L'hypothèse de l'abandon du vocabulaire contrôlé préexistant, lorsqu'on la mentionne, est rapidement écartée²⁴⁰.

Le KBLab a toutefois voulu explorer cette piste, grâce à un modèle de type Sentence-BERT entraîné sur des données suédoises et la méthode BERTopic²⁴¹. À grands traits, la procédure est la suivante : un modèle transformeur est utilisé pour vectoriser les paragraphes dont se composent des documents, les vecteurs ainsi obtenus sont regroupés en thèmes, des mots sont extraits des textes au sein de chaque thème pour assigner un label à celui-ci, et le sujet de chaque document est enfin déterminé par l'agrégat des thèmes de ses paragraphes²⁴². Plusieurs thèmes peuvent être unis du plus général au plus précis, comme dans un vocabulaire contrôlé classique. La conclusion de cette expérimentation est enthousiaste : « *It is possible to develop a machine-learning based approach that enables users to search and navigate according to their interests without relying on fixed headings from a*

²³² (MERGEL, KÜHLER, SCHUMANN et al., 2026, p. 36).

²³³ On trouvera des explications claires sur ces différents indicateurs dans (POLEY, UHLMANN, BUSSE et al., 2025, p. 17-19).

²³⁴ (KASPRZIK, 2025a), (KASPRZIK, 2024cp. 16).

²³⁵ (RONCAGLIA, 2026, p. 122) suggère, dans le même ordre d'idées, qu'il serait plus productif d'appliquer l'IA de pointe au signalement entités-relations qu'à des notices A/B.

²³⁶ (KASPRZIK, 2020, p. 4).

²³⁷ (MALMSTEN et LUNDBORG, 2025).

²³⁸ (SACCUCCI et POTTER, 2025, p. 13). Sur les avantages de FAST pour la recherche, voir (KAMAL et GOLUB, 2025p. 19).

²³⁹ (KASPRZIK, 2020, p. 4-5).

²⁴⁰ (GOLUB, 2021, p. 707).

²⁴¹ (MALMSTEN, LUNDBORG, FANO et al., 2025).

²⁴² Pour le détail de la procédure, voir (MALMSTEN, LUNDBORG, FANO et al., 2025, p. 185).

controlled vocabulary. Whether such a system is used to replace, extend or support existing practices will ultimately depend on the usage of the current subject heading system at any given institution. »²⁴³

Reste à déterminer quels seraient les avantages de ce nouveau vocabulaire *incontrôlé*, généré à la volée selon le contenu de chaque document, par rapport à FAST, RAMEAU, etc. Les perspectives ouvertes par cette technologie sont assurément impressionnantes, pour ne pas dire éblouissantes : « *Present possibilities of topic modeling allow better access and division into intuitive, fluid, and interconnected frameworks. They nurture a space where learning is not linear but richly layered and a multidimensional subject access journey.* »²⁴⁴ Dans les faits, les institutions se montrent mesurées : elles retravaillent (ou réfléchissent à retravailler) leur vocabulaire contrôlé et facilitent ainsi la réussite de l'IA, lors de la phase l'indexation²⁴⁵ ou lors d'une recherche de type RAG²⁴⁶. L'équipe d'Annif se montre optimiste quant au maintien de l'importance de l'indexation matière : « *I think subject indexing will keep being crucial, even in the age of AI, or rather especially in the age of AI, because LLMs need data and these subject indexes are valuable data to the language models as well.* »²⁴⁷

Le rôle du catalogueur, spécialiste de la procédure d'indexation, reste alors fondamental. Même si la recherche sémantique dans les catalogues se généralisait par vectorisation du texte disponible pour chaque document de la bibliothèque (contenu, notice, résumé, table des matières selon les cas), il est probable que l'expérience de navigation dans les collections reste plus cohérente et structurée si des vedettes-matière permettent au lecteur de passer d'une ressource particulière à un ensemble de ressources similaires *en comprenant quel est leur dénominateur commun*²⁴⁸. À moins que la granularité supérieure des techniques de *topic modelling* ne renverse la donne, il y a tout à parier que les sujets, de toute façon nécessaires, seront puisés dans un vocabulaire contrôlé construit et entretenu humainement.

Plutôt que de remplacer cette forme ancienne d'organisation de la connaissance, il est possible que les LLM la rendent plus utile encore, en aidant un lecteur peu familier des vedettes-matière en usage dans sa bibliothèque à construire une recherche qui en tire un plein parti²⁴⁹. Il est aussi envisageable que l'IA serve à analyser un corpus pour suggérer de nouveaux termes à verser dans un vocabulaire contrôlé : la ZBW a étudié le recours à des modèles de la famille BERT pour enrichir son thesaurus économique, après examen des candidats par les éditeurs²⁵⁰. On ne peut qu'être d'accord avec l'orientation proposée par Jim Hahn : « *New indexing techniques must acknowledge our old indexing cultures, while simultaneously co-creating new cultures of description that may be an ever more collaborative endeavor of knowledge machines and knowledge professionals.* »²⁵¹

On peut penser qu'une indexation épousant les spécificités thématiques de chaque texte se justifierait si elle descendait du niveau du document à celui de ses parties (pages d'un journal par exemple, voire paragraphes de ces pages), comme cela est envisagé dans la version future de Gallica. RAMEAU est assez vaste et détaillé pour rendre compte de l'immense majorité des livres, mais quel vocabulaire contrôlé peut prétendre à une ampleur et minutie telle qu'il rende possible l'indexation de tous les paragraphes de la moindre page ? Quand bien même l'on recourrait à une indexation fondée sur des vedettes générées sur-mesure, l'alignement dans un second temps avec les référentiels existants de la BnF resterait toutefois crucial, souligne Arnaud Laborderie, chef de projet Gallica²⁵². Même à considérer une bibliothèque numérique, en sortant donc du cadre strict de notre étude, on voit que la pertinence de ce domaine d'expertise catalographique traditionnel qu'est l'indexation n'est pas menacée.

²⁴³ (MALMSTEN, LUNDBORG, FANO et al., 2025, p. 195).

²⁴⁴ (UZWYSHYN, 2024, p. 177).

²⁴⁵ La DNB décrit ainsi l'ajout ou le retrait de vedettes dans le GND, voir (JUNGER, 2018, p. 8).

²⁴⁶ (KASPRZIK, 2024c, p. 29-30), (KASPRZIK, 2024ap. 164).

²⁴⁷ (LEHTINEN, 2025b).

²⁴⁸ On ne peut qu'être d'accord avec (MALMSTEN, LUNDBORG, FANO et al., 2025, p. 196) : une perspective et des recherches UX sont ici nécessaires.

²⁴⁹ (KAMAL et GOLUB, 2025, p. 21).

²⁵⁰ (KEMPF, 2025).

²⁵¹ (HAHN, 2024, p. 14).

²⁵² (LABORDERIE, 2025).

DU CATALOGUEUR PRODUCTEUR AU CATALOGUEUR RELECTEUR : UNE TRANSFORMATION AMBIVALENTE

FEUILLE DE ROUTE

Si son expertise garde toute sa valeur, le catalogueur œuvrant dans une chaîne de traitement plus ou moins automatisée s'achemine vers le rôle de relecteur. Il y a là une potentielle montée en généralité, mais aussi une éventuelle menace pour l'expertise, due aux effets négatifs de l'automatisation. Ce n'est pas seulement une question d'efficacité et de précision du signalement, mais de sens du travail : risque-t-il de se perdre si l'on s'éloigne des formes actuelles du signalement ? Les sources de satisfaction professionnelle au sein de la profession sont si variées qu'on ne peut répondre qu'en prenant les plus grandes précautions.

Liens des sections :

[↗ Montée en généralité](#)[↗ Menace pour l'expertise](#)[↗ Ironies de l'automatisation](#)[↗ Automatisation et sens du travail](#)[↗ Des sources de satisfaction professionnelle variées](#)

5.1. Une montée en généralité

Comme l'écrit Coralie Philibert, l'IA provoque une évolution des missions des catalogueurs : « Grâce à l'IA, le catalogueur obtient certes ensuite un gain de temps indéniable puisque la manipulation manuelle devient automatisée mais sa mission se transforme : son rôle est de relire, vérifier, comparer, valider et corriger le cas échéant les données produites. Il devient un relecteur-valideur. »²⁵³ La machine propose, l'humain tranche et corrige (acceptation, reprise manuelle, ajustement du modèle, correction par lot). C'est dans cette répartition des tâches que le groupe de travail RDA / RDF sur l'IA voit l'avenir de la profession : « *Therefore, catalogers will remain indispensable as overseers, interpreters, and innovators, working alongside AI rather than being replaced by it.* »²⁵⁴ Les éditeurs d'un numéro récent de *Cataloging & Classification Quarterly* consacré à l'IA soulignent que les implications de ce changement de rôle doivent être examinées du point de vue managérial²⁵⁵.

Des descriptions de chaînes récentes attestent que la transition est en cours. L'outil SGCAT mène les catalogueurs du rôle de créateur à celui de relecteur critique²⁵⁶. Dans l'outil de catalogage automatique fondé sur ChatGPT mis en production par l'Université autonome du Chili, c'est la supervision et la validation humaine finale qui garantissent la qualité des notices²⁵⁷. La même vision est présentée par Hannes Lowagie, à la KBR : « *Cataloguers will have to collaborate with AI systems to enrich, validate and refine the metadata these systems produce. This means that rather than creating metadata from scratch, cataloguers review the output of AI models, assess its quality and make any necessary adjustments to ensure accuracy.* »²⁵⁸

On peut dire qu'il se produit là un changement similaire à celui que la dérivation de notices a produit par le passé, tel que le décrit Benoît Tuleu, directeur du Département du Dépôt légal de la BnF : « Il y a 30, 40 ans un catalogueur savait saisir dans les bonnes zones la bonne information. Aujourd'hui, il contrôle des données venues de l'extérieur. »²⁵⁹ Il est vrai que, la dérivation d'une notice depuis une base externe, ou la génération de celle-ci par une IA (KB, KBR, Chili, Singapour), dès lors que les deux procédures requièrent une correction subséquente, produisent une expérience de travail semblable. Certes les erreurs qu'on peut s'attendre à trouver dans un cas ou dans l'autre ne sont pas de même

²⁵³ (PHILIBERT, 2024, p. 86).

²⁵⁴ (COHEN, ALIVERTI, BARBUS et al., 2025, p. 2).

²⁵⁵ (COX et HAJIBAYOVA, 2025, p. 406).

²⁵⁶ (LING et GOH, 2025, p. 10).

²⁵⁷ (DIAZ et VERGARA, 2025, p. 22).

²⁵⁸ (LOWAGIE, 2025a, p. 25).

²⁵⁹ (TULEU, 2025).

nature, mais il s'agit bien d'une procédure de traitement unitaire, impliquant une relecture attentive et une comparaison soutenue avec l'ouvrage physique entre les mains.

Le rôle de supervision par échantillonnage que l'IA pourrait conférer au catalogueur est en quelque sorte préfiguré par d'autres chaînes déjà en fonctionnement. Ainsi au Dépôt légal du son dématérialisé de la BnF, des flux de métadonnées sont envoyés par les diffuseurs, et celles-ci sont transformées automatiquement en notices structurées. Le Département Son, vidéo, multimédia est chargé de contrôler leur correction, et procède pour cela à des extractions hebdomadaires, servant à repérer les anomalies et apporter des corrections²⁶⁰. À ceci près que cette procédure de conversion est déterministe et n'implique pas de modèle à entraîner et évaluer, le travail rappelle celui que décrivent les équipes de la ZBW et de la DNB pour le contrôle des notices indexées automatiquement.

Le déplacement vers la relecture peut être interprété comme une montée en généralité : l'agent se concentre moins sur la saisie unitaire que sur la cohérence globale, sur les cas limites ou les anomalies. C'est ainsi que Hannes Lowagie présente cette transformation : « *In essence, this shift moves cataloguers from a traditional, hands-on role to a higher level position where they are responsible for ensuring the precision and reliability of large-scale automated cataloguing systems.* »²⁶¹ Le catalogueur devient ainsi garant de la qualité, et ses corrections nourrissent des boucles de rétroaction qui, dans certains dispositifs vus plus haut, améliorent le système. Le fait de traiter un plus grand nombre de notices avec moins de saisie manuelle, ainsi que la contribution ce faisant à un projet d'IA efficace, pourraient être sources de satisfaction professionnelle. Il n'est pas exclu que cette nouvelle modalité de travail, en collaboration avec l'IA, soit perçue comme une opportunité positive de renouveler ses pratiques et d'étendre ses compétences (comme le dit un rapport sur d'éventuels projets d'IA au SCD de l'Université Paris-Saclay²⁶²), un défi où la profession manifesterait la souplesse et l'aptitude à l'adaptation qui la caractérisent²⁶³. Ce n'est pas seulement la machine qui apprend, mais aussi le catalogueur, qui, confronté à des erreurs d'un certain type commises par l'IA, aurait à déterminer quels lots de notices corrigées lui fournir pour augmenter la précision des prédictions dans un domaine spécifique. Une telle démarche relèverait de ce que certains chercheurs appellent l'adaptation conjointe : « *[...] the human and the computer are both in the loop, and AI becomes a co-adaptive process, in which a human is changing AI behavior, but the human also adapts to use AI more effectively and adapts their data and goals in response to what is learned using machine learning.* »²⁶⁴

5.2. Une menace pour l'expertise ?

5.2.1. Une potentielle perte de compétences

Néanmoins ce rôle de relecteur ne va pas sans de nombreuses difficultés, bien repérées. La première touche à la constitution et au maintien des compétences catalographiques. Dans un circuit de traitement où les agents humains ne cataloguent presque plus, un problème se pose : « À être passif, ne perd-on pas les bases ? »²⁶⁵ Or poser ces bases requiert une très longue pratique, qui ne peut pas s'acquérir de façon théorique, livresque. Comme le dit Nuria Pastor-Martinez, citée par Coralie Philibert : « Le catalogage n'est pas une question de connaissances mais une question de pratique. C'est avec l'expérience qu'on devient expert, quand on est capable de passer à des automatismes. [...] L'activité du catalogueur ne serait ni rentable ni efficace si celui-ci cataloguait au compte-gouttes ; il ne serait pas sûr de lui. Pour qu'un catalogueur soit efficace et prenne conscience de ses compétences, il faut un certain temps. »²⁶⁶

Une relecture efficace exige une attention soutenue sur des objets dont la structure est complexe, et que la perte de la pratique quotidienne peut avoir rendu plus difficiles à appréhender. L'IFLA met en garde contre la perte de compétence que l'IA pourrait causer, sous la rubrique « *Threats to human agency* » de sa liste de risques : « *AI can make life easier, but can also lead to a loss of skills and confidence, creating technology dependence. AI is often being used in invisible or opaque ways, which*

²⁶⁰ Cécile Queffelec, communication orale, 29 janvier 2026.

²⁶¹ (LOWAGIE, 2025a, p. 25).

²⁶² (RUSSEIL-SALVAN, 2024, p. 105).

²⁶³ (PHILIBERT, 2024, p. 79-80).

²⁶⁴ (NATARAJAN, MATHUR, SIDHEEKH et al., 2024, p. 4).

²⁶⁵ (KOSKAS, 2025).

²⁶⁶ Voir (PHILIBERT, 2024, p. 28).

can result in the reduction of human choice»²⁶⁷. La ZBW souligne ce risque²⁶⁸, ainsi que la Bibliothèque nationale de Finlande : « *Over time, the model has become more accurate with fewer mistakes. However, there are concerns about hidden mistakes in assigning subject headings. Therefore, it's important for librarians to remain skilled in cataloguing and indexing subject headings.* »²⁶⁹ En l'absence d'experts ayant une longue expérience de l'indexation, on aboutirait à une situation où il est difficile de déterminer si les prédictions contiennent ou non des erreurs. Dans son examen des risques d'un circuit comme PowerApps pour le catalogage courant, Morgane Russeil-Salvan mentionne celui de la perte de maîtrise du format UNIMARC²⁷⁰. Un billet du blog d'OCLC *Hanging together* expose la même crainte, tant pour les catalogueurs entrant dans la profession que pour leurs pairs expérimentés : « *Key considerations include how to ensure that new professionals develop foundational skills traditionally gained through tasks like brief record creation—skills that become essential for effectively evaluating AI outputs later in their careers. Experienced catalogers wondered whether spending more time reviewing than creating might impact their ability to identify subtle errors or handle complex materials that require human insight.* »²⁷¹

5.2.2. Formation des recrues

Cette dernière réflexion met à jour une autre inquiétude sous-jacente, celle de la formation des futures recrues. Un article fondateur sur les conséquences de l'automatisation en contexte industriel l'énonçait déjà il y a plus de quarante ans : « *There is some concern that the present generation of automated systems, which are monitored by former manual operators, are riding on their skills, which later generations of operators cannot be expected to have.* »²⁷² Un autre billet du blog d'OCLC formule clairement le problème : « *If AI handles foundational tasks like creating brief records—work that traditionally serves as essential training for new catalogers—how do we ensure new professionals still develop the core skills they'll need to effectively evaluate AI outputs?* »²⁷³ Les institutions recourant à l'AI se posent la question, comme la KB dont l'outil Retrotool, rappelons-le, permet d'obtenir une notice en MARC 21 d'un ouvrage à partir de la page de titre de l'ouvrage placée sous une caméra :

*We shouldn't erase the need to know MARC 21 and cataloging rules just because you can put the book under the lens. How to preserve cataloging knowledge that's already scarce? Many employees are going in retirement, a lot of knowledge is going away. As a library, to carry out tasks in the future, we need to keep educating people on that.*²⁷⁴

La formation des catalogueurs en début de carrière est d'autant plus cruciale que l'expertise catalographique ne peut plus être considérée comme acquise en sortie de formation. Dans un tel contexte, on peut craindre que la réduction des opportunités d'appliquer codes et formats au sein de circuits de signalement où l'IA génère la plupart des notices ne fasse obstacle à l'entraînement de l'humain censé surveiller le système. À la KB, les nouvelles recrues doivent être entraînées *ex nihilo* au sein de l'institution²⁷⁵. En France, l'évolution du programme des concours a diminué la part du catalogage dans les formations bibliothéconomiques, et les nouveaux agents ont donc des compétences moins développées dans ce domaine²⁷⁶. Il en est de même à la Bibliothèque nationale d'Israël²⁷⁷, ou en Belgique encore :

With the declining availability of library and document education programmes, there are fewer opportunities for individuals to gain formal training in these essential areas. This trend means that fewer professionals are entering the field with the specialised knowledge required for roles in library science, document management and information organisation. As these educational programmes become scarcer, the industry faces challenges in ensuring that the next generation of professionals

²⁶⁷ (COX et DE BRASDEFER, 2025, p. 5).

²⁶⁸ (KASPRZIK, 2025b, p. 9).

²⁶⁹ (MERGEL, SCHMIDT, SCHUMANN et al., 2024, p. 24).

²⁷⁰ (RUSSEIL-SALVAN, 2024, p. 27).

²⁷¹ (BRYANT 2025).

²⁷² (BAINBRIDGE, 1983, p. 776). Pour une application à la programmation, voir (FRIEDRICHSEN 2025).

²⁷³ (PROFFITT 2025b).

²⁷⁴ (BUESINK, 2025b).

²⁷⁵ (BUESINK, 2025b).

²⁷⁶ (ROCHE et KOSKAS, 2025, p. 2).

²⁷⁷ (COHEN, 2025).

is equipped with the skills needed to maintain high standards in cataloguing, archival work and resource management.²⁷⁸

5.3. Ironies de l'automatisation

Ainsi, la compétence catalographique s'acquiert et se maintient en pratiquant ; si l'on supprime durablement les tâches fondatrices qui fournissent des occasions d'apprentissage (création de notices, édition après dérivation), on rend plus difficile la formation de nouveaux professionnels, et l'on expose les professionnels expérimentés eux-mêmes à une diminution graduelle de leur acuité. Il y a longtemps que ce risque de perte de compétence a été identifié dans d'autres domaines : « L'automatisation crée des problèmes de compétences (car elles sont moins mobilisées) et de vigilance qui peuvent interférer avec la capacité des humains à intervenir lorsqu'il le faudrait. »²⁷⁹ Il est au cœur de l'article de psychologie cité plus haut : les compétences manuelles ou cognitives s'érodent avec le temps si elles ne sont pas mises en pratique, cas où se trouverait un agent dont la seule mission consisterait à surveiller le résultat d'un processus automatique²⁸⁰. Pour contrecarrer cet effet négatif, il est par exemple proposé de laisser l'opérateur reprendre une activité directe pendant chaque période de travail de surveillance²⁸¹. Sans de telles mesures préventives, on risquerait d'aboutir à une situation où les normes de description ne sont appliquées qu'au travers de l'IA, sans que ses résultats ne puissent être évalués par l'humain, celui-ci s'étant peu à peu dessaisi de la connaissance intime qu'il en avait : « Surtout, si la complémentarité consiste seulement en une délégation progressive de tâches à l'IA dans une conception servile de cette dernière, au lieu de donner lieu à la création de plus d'intelligence, elle peut déboucher, si on n'y prend garde, sur une dynamique relevant de la dialectique hégélienne du maître et de l'esclave : l'humain perd progressivement l'expertise et devient de plus en plus dépendant de la machine. »²⁸²

L'ironie est ici que les erreurs subsistantes dans un système d'AI bien rodé, risquent de provenir de cas-limites compliqués à déceler, pour lesquels l'agent aura besoin d'une pleine maîtrise des normes applicables. On peut penser que cette couche de connaissances subtiles relatives à des spécificités documentaires rarement observées sera la première à s'émousser en cas de diminution de la pratique. Après des centaines de monographies cataloguées, un catalogueur n'oublie pas de sitôt dans quelle zone il convient de reporter le titre. Mais qu'en est-il de la procédure à suivre face à une couverture portant un titre différent de celui de la page de titre ? Ou du titre d'un document peu fréquent dans le service où il travaille, mettons un catalogue d'exposition ? Un système performant mais imparfait fait porter l'effort de détection sur des erreurs plus coûteuses cognitivement à repérer, liées qu'elles sont aux nœuds du graphe des connaissances catalographiques les plus vulnérables à l'atrophie par manque de pratique. « *Efficient retrieval of knowledge from long-term memory depends on frequency of use* »²⁸³, rappelle Lisanne Bainbridge.

Une menace supplémentaire est le biais nous portant à adopter trop volontiers les suggestions des systèmes automatisés, biais décrit dans un récent rapport du Contrôleur européen de la protection des données. Il est encore accentué dans un contexte où l'expertise humaine n'est pas à la mesure de la complexité des problèmes délégués par la machine : « *Operators' decisions are often shaped by a strong trust in the system's perceived expertise, especially when the system performs challenging calculations or inferences, or operates on complex or highly specialized data for which the human may feel unqualified to intervene.* »²⁸⁴. Cela impose d'imaginer des dispositifs incitant au doute²⁸⁵. Si les agents chargés de relecture ne bénéficient plus d'une expertise suffisante, ils risquent d'être particulièrement vulnérables, laissant les erreurs récurrentes de l'IA former la norme implicite, le critère faussé dont ils se servent pour l'évaluer. Dans cette spirale négative, à mesure que la compétence humaine s'éroderait, le sentiment de légitimité nécessaire pour apporter une correction dans une notice générée automatiquement diminuerait. Ce phénomène est lui aussi bien étudié : « *This perceived imbalance in knowledge can*

²⁷⁸ (LOWAGIE, 2025a, p. 16).

²⁷⁹ (VUARIN et STEYER, 2025, p. 246).

²⁸⁰ (BAINBRIDGE, 1983, p. 775).

²⁸¹ (BAINBRIDGE, 1983, p. 777).

²⁸² (GAUTIE et PÉREZ, 2025, p. 99).

²⁸³ (BAINBRIDGE, 1983, p. 776).

²⁸⁴ (BERNARDO et HERNÁNDEZ, 2025, p. 10).

²⁸⁵ (VUARIN et STEYER, 2025, p. 264).

discourage intervention and reinforce reliance on the system, even in situations where human judgment is necessary. »²⁸⁶

Un système où l'activité du catalogueur se bornerait à la seule relecture serait-il donc voué à l'enraiment à moyen terme ? Une politique de compétences cohérente doit, en cas d'adoption d'un tel système, garantir une pratique suffisante des gestes qui fondent l'expertise, et organiser la relecture pour qu'elle reste viable. S'exercerait-elle champ par champ ? Notice par notice ? Il faudrait se fonder sur l'expérience acquise dans l'organisation de la relecture humaine de notices rédigées par des humains au sein de services de grande envergure comme ceux de Dépôt légal. Ce qui vaut dans de tels contextes vaudra certainement dans une chaîne partiellement automatisée : « La relecture est une question de calibrage ; il y a une infinité d'erreurs possibles à corriger. »²⁸⁷ Il sera essentiel de déterminer le périmètre du contrôle qualité attendu, et de donner aux agents des moyens effectifs de se prononcer sur des notices dues à l'IA : facilité d'accès, dans l'environnement numérique de travail, aux numérisations sur laquelle se fondent la prédiction, interface permettant de modifier aisément la notice, etc.

Citant des études de psychologie antérieures, Lianne Bainbridge explique encore qu'il est humainement impossible de surveiller longtemps, de façon fiable, un système en vue de détecter des anomalies peu probables. La vigilance de l'opérateur chargé de surveiller un système automatique flanchera s'il a fonctionné correctement pendant une durée assez longue pour inspirer confiance²⁸⁸. La vigilance de l'humain au sein d'un dispositif de collaboration avec la machine, doit être organisée avec soin²⁸⁹. Le Contrôleur européen de la protection des données y insiste : « *Operators responsible for monitoring ADM [Automated Decision-Making] systems might be tasked with oversight processes that rarely exhibit obvious failures. While this may sound reassuring, the rarity of errors can actually lead to a decline in vigilance over time - a well-documented phenomenon in supervisory control settings. To counteract this and maintain high levels of attentiveness and decision quality, it is crucial to design workflows that actively support and challenge human oversight.* »²⁹⁰ Une idée intéressante à cet égard est celui des contrôles périodiques de l'attention, « *awareness checkpoints* »²⁹¹, consistant en des tâches de calibrage entremêlées au flot des tâches réelles. Dans le contexte catalographique, une chaîne de traitement largement automatisée imposerait donc d'imaginer une relecture non seulement entrecoupée de dispositifs de maintien de la vigilance et de l'expertise, mais encore soigneusement mesurée dans ses objectifs : ni pointilleuse au point de nuire exagérément au rendement, ni si lâche qu'elle succombe à cette dégradation graduelle de l'attention.

5.4. Automatisation et sens du travail

5.4.1. Un risque de dévalorisation

Au-delà de ce problème de faisabilité, il faut se demander comment la relecture, en devenant l'activité principale, affecterait la qualité de vie au travail, le sens et la valeur qu'on attribue à celui-ci. Notons qu'aujourd'hui, c'est un rôle vu parfois comme pénible et ingrat. L'expérience des services de signalement de vaste envergure comme celui de la Bibliographie nationale française, au Dépôt légal à la BnF, permet de s'en convaincre. La relecture des notices rédigées par les collègues du service fait l'objet d'une prime, car l'activité était perçue à son instauration comme un plus par rapport aux fonctions des Bibliothécaires adjoints spécialisés, à valoriser comme tel ; quoiqu'elle apporte une forme de reconnaissance officielle de l'expertise, c'est la partie de leur journée qu'aiment le moins les catalogueurs²⁹². Comme on l'a souligné, « ce travail est à la fois chronophage, fastidieux et essentiel puisqu'il est garant de la qualité finale des notices mises au catalogue »²⁹³.

Il ne s'agit pas là d'une anomalie. Une *Metadata Librarian* de la bibliothèque de droit de l'Université de Notre-Dame pose la question : « *Will our work shift from intellectually engaging tasks like original description to primarily quality control? And if so, will the work still feel meaningful?* »²⁹⁴

²⁸⁶ (BERNARDO et HERNANDEZ, 2025, p. 13).

²⁸⁷ (KOSKAS, 2025).

²⁸⁸ (BAINBRIDGE, 1983, p. 776).

²⁸⁹ (VUARIN et STEYER, 2025, p. 268).

²⁹⁰ (BERNARDO et HERNANDEZ, 2025, p. 29).

²⁹¹ (BERNARDO et HERNANDEZ, 2025, p. 30).

²⁹² (KOSKAS, 2025).

²⁹³ (PHILIBERT, 2024, p. 50).

²⁹⁴ (ANDERSON et TIAN, 2025, p. 368).

Les membres du groupe *Primary cataloging workflows* au sein du *Managing AI in Metadata Workflows Working Group* formé par l'OCLC Research Library Partnership perçoivent le même danger : « *Others wondered whether shifting from creating metadata to reviewing AI-generated records might make their work less engaging or meaningful.* »²⁹⁵ Cette passivité nouvelle est présentée comme potentiellement aliénante : « *Passare il proprio tempo a correggere errori nei metadati prodotti dall'IA sembra essere un compito assai più faticoso e alienante della produzione diretta degli stessi metadati da parte del bibliotecario, anche se alla fine i risultati sono quantitativamente di più e qualitativamente più ricchi.* »²⁹⁶ Ce qui était auparavant un travail gratifiant deviendrait alors corvée : « *If AI is more efficient, it may replace professionals in many roles. AI could be used to control human work and reduce rewarding work to drudgery.* »²⁹⁷

Passer du catalogage actif à la relecture pourrait miner la motivation professionnelle, en vidant l'activité de sa substance et en diminuant l'aura d'expertise qui s'y attache. Le statut même de l'agent qui s'en trouve affecté, comme le dit Lisanne Bainbridge : « *The level of skill that a worker has is also a major aspect of his status, both within and outside the working community. If the job is "deskilled" by being reduced to monitoring, this is difficult for the individuals involved to come to terms with.* »²⁹⁸ C'est bien l'une des craintes des bibliothécaires ; l'un d'eux, dans un sondage administré en ligne début 2024 sur des listes de diffusion professionnelles aux États-Unis, dit s'attendre à cette diminution de l'expertise du métier : « *I expect cataloging to be further de-professionalized as a result of AI.* »²⁹⁹

La menace qui se dessine ici concerne d'autres pans de l'application de l'IA dans la société. Comme l'écrit Robert B. Fisher, synthétisant une dizaine de cadres réglementaires et déclarations de principes relatives à l'IA, la société n'a aucun intérêt à ce que cette technologie la prive d'activités riches de sens à accomplir, cruciales pour la santé mentale et l'estime de soi des individus³⁰⁰. Le projet DIAL-IA (Dialoguer sur l'IA), porté et coordonné par l'Institut de recherches économiques et sociales (IRES), souligne dans son manifeste ce risque de perte de sens en cas d'application aveugle de solutions foulant aux pieds ce qui fait, aux yeux de l'agent, la valeur de son travail :

Les SIA (Systèmes d'Intelligence Artificielle) sont développés à partir du travail prescrit, des représentations du travail qui ne prennent pas en compte les nuances, les subtilités, la complexité qui font l'intérêt du travail réel et lui donnent du sens. Réduire le travail réel en données exploitables par des SIA les rendent en quelque sorte aveugles. Le risque est fort que ces SIA deviennent alors des contraintes supplémentaires pour les travailleurs pouvant contribuer à la perte de sens, voire conduire à l'obsolescence humaine.³⁰¹

5.4.2. L'attachement au travail bien fait

Le catalogage, par la méticulosité qu'il requiert dans le respect des normes, du format, appliqués au cas d'espèce d'un ouvrage spécifique, peut procurer, selon la nature des documents traités et les dispositions de l'agent, le sentiment gratifiant d'avoir œuvré *dans les règles de l'art*³⁰². L'aspiration à faire du beau travail se matérialise dans l'idéal de la « belle notice », ancré dans les représentations du métier³⁰³. Des travaux sociologiques récents peuvent éclairer ce faisceau de notions : « Le beau travail, ce n'est pas seulement de "la belle ouvrage", cela peut être aussi quelque chose réalisé selon les règles, un travail de qualité, reconnu par les pairs comme étant de valeur. »³⁰⁴ La notice bien faite d'un volume complexe, appartenant à un ensemble monographique hétérogène, dans une langue et une écriture rares, apporte un sentiment qui dépasse la seule utilité de ce travail. La perception de la qualité des champs remplis confine à l'esthétique : la translittération systématique des entrées d'une table des matières, par exemple, présente une régularité, une netteté harmonieuse qui donne lieu à

²⁹⁵ (PROFFITT 2025b).

²⁹⁶ (RONCAGLIA, 2026, p. 118).

²⁹⁷ (COX, 2024, p. 160).

²⁹⁸ (BAINBRIDGE, 1983, p. 776).

²⁹⁹ (CHEN et AND LI, 2024, p. 324).

³⁰⁰ (FISHER, 2025, p. 5).

³⁰¹ (*Pour un dialogue social au service des bons usages de l'IA et d'une nouvelle étape de progrès social dans les entreprises et les administrations* 2025).

³⁰² Je me fonde ici sur des observations faites à l'Humathèque, bibliothèque du Campus Condorcet, entre le 1^{er} 2023 et le 31 décembre 2025. J'y étais chargé des collections tibétaines et himalayennes au service des collections documentaires, où travaille une vingtaine d'acquéreurs catalogueurs, répartis en domaines thématiques et aires géographiques.

³⁰³ Emmanuel Jaslier, communication orale, 29 janvier 2026.

³⁰⁴ (BOUILLOUD, 2023, p. 61).

un « jugement de beauté », partagé avec les collègues. Ce jugement, dit Jean-Philippe Bouilloud, « est porté par les membres d'un groupe sur quelque chose que d'une certaine manière, ils sont souvent les seuls à apprécier, que ce soit un "beau béton" ou un "beau tableur électrique" »³⁰⁵. En devenant immatérielle lors de l'informatisation des catalogues, la description catalographique n'a pas perdu cet attrait. Au contraire, de nouvelles dimensions s'y sont ajoutées : le juste établissement des liens vers les autorités pertinentes, une plus grande ambition de complétude, peuvent renforcer cette satisfaction devant la qualité atteinte. L'identité professionnelle et le sentiment d'appartenance à un groupe, qui aident si puissamment à donner du sens au travail, se forment dans cette valorisation collective : « Le jugement esthétique porté par les pairs se transforme en support d'identité pour celui qui est reconnu. [...] L'ouvrier soudeur réputé dans son atelier pour la perfection de ses soudures, le magistrat pour la qualité de la rédaction de ses jugements, ou le maçon pour celle de ses constructions, ont dans la belle pratique de leur technique, dans le "jugement de beauté" défini par leurs pairs, quelque chose qui participe de leur identité. »³⁰⁶ Le souci du travail bien fait est à ce point reconnu comme caractéristique fondamentale de l'identité du catalogueur que l'imaginaire professionnel dispose d'un terme pour dénoncer en le désignant son envers négatif : la « sur-qualité »³⁰⁷, c'est-à-dire le déséquilibre entre la finesse du signalement unitaire et la masse documentaire à traiter ou le profit qu'en tire, *in fine*, le lecteur parcourant le catalogue.

À l'égard de cet attachement à la qualité du travail, l'introduction d'un système d'IA catalographique présente deux risques. Le premier, celui d'une diminution temporaire ou durable de la précision, relèverait du « travail empêché », cette situation où « l'organisation, à force de réduire les coûts et de rationaliser la production [...] empêche aussi cette dynamique positive de la valeur, et prive l'individu du sentiment de faire quelque chose de qualité »³⁰⁸. Le souci de qualité est au cœur du métier des catalogueurs ; un dispositif produisant des résultats incertains peut être vécu comme une dégradation des conditions d'exercice, contrevenant à l'éthique professionnelle. Par leur vivacité, de telles peurs et inquiétudes pourraient aller jusqu'à relever des risques psychosociaux³⁰⁹. Un témoignage à la DNB relève la tension entre l'aspiration à la qualité du travail et les approximations fréquentes de l'IA : « *The quality is getting worse, and you're embarrassing yourself to the outside world. So that has been a factor because, of course, we also have a lot of mistakes in there, so the processes are not perfect, and especially in the library context, this perfectionism was and is very pronounced, which is also right and important.* »³¹⁰

Mal géré, ce risque pourrait provoquer une résistance au changement, phénomène qui souvent résulte moins d'un refus de principe que d'appréhensions à l'idée de mal faire. Quoiqu'une méta-étude récente³¹¹ suggère que, dans de nombreux domaines, les systèmes hybrides humain-IA font mieux que l'humain seul, cette collaboration nécessite un calibrage minutieux. Et cela lorsque l'IA a effectivement des aperçus utiles à livrer ; lorsqu'un système de recommandation de vedettes-matière produit des propositions jugées peu pertinentes, il peut accroître la confusion et la charge mentale, entraînant un scepticisme ancré non dans une hostilité foncière à l'innovation mais dans le constat de défaillances objectives³¹².

Même si la notice générée par l'IA est aussi bonne que celle des experts humains, on peut douter que ceux-ci, en la corrigeant partiellement, éprouvent le sentiment de fierté que nous venons de considérer. Un travail de ce type ne se cristallise pas en un objet précis, résultat du processus de production, puisque les notices ajoutées au catalogue auront été, en grande partie, remplies sans intervention humaine. Dans une réorientation du métier vers la relecture l'acte de création artisanale disparaîtrait, et donc la possibilité de s'enorgueillir du résultat de son travail : « *La beauté du travail est un support, une revendication de valeur, et cette revendication de valeur va rebondir sur l'individu qui réalise ce travail.* »³¹³ Les implications profondes du passage d'une logique artisanale à une logique industrielle sont parfois perçues dans la littérature bibliothéconomique : « *Put differently, is the library community fine with reducing cataloging and classification to mechanized production, or is it a craft that draws on a singular set of knowledge, intuition, and attention to detail?* »³¹⁴

³⁰⁵ (BOUILLLOUD, 2023, p. 62).

³⁰⁶ (BOUILLLOUD, 2023, p. 62).

³⁰⁷ (BELLIER, 2017, p. 42), (PHILIBERT, 2024p. 31).

³⁰⁸ (BOUILLLOUD, 2023, p. 64).

³⁰⁹ (TULEU, 2025).

³¹⁰ (MERGEL, KÜHLER, SCHUMANN et al., 2026, p. 32).

³¹¹ Voir (VACCARO, ALMAATOUQ, et MALONE, 2024, p. 2-5).

³¹² (ASULA, MAKKE, FREIENTHAL et al., 2021, p. 9).

³¹³ (BOUILLLOUD, 2023, p. 64).

³¹⁴ (MOULAISON-SANDY et AND COBLE, 2024, p. 382).

Un système comme l'outil de recommandation de vedettes-matière de l'Abes est centré sur l'assistance au catalogueur ; dans la conception d'un circuit automatisé d'envergure qui romprait avec cette direction générale, à l'instar de la ZBW ou de la DNB, il faudrait être particulièrement attentif au danger de perte de sens et de satisfaction des agents. Il n'est pas exclu que l'allègement des tâches répétitives, le soulagement de voir un fonds ancien enfin traité ou un flux de nouveautés intégré sans retard aux collections de la bibliothèque contrebalancent la mutation que nous venons de décrire. Peut-être ce sentiment du travail bien fait se déplacerait vers le catalogue dans son ensemble, dont l'ordonnancement régularisé par de multiples chantiers qualité automatiques, donnerait lieu, chez le relecteur, à ce « jugement de beauté » que la création d'une notice individuelle provoque aujourd'hui.

5.5. Des sources de satisfaction professionnelle variées

On touche là un point crucial pour la qualité de vie au travail, dans sa dimension psychologique fondamentale. « *Automate the mundane to focus on the fun* »³¹⁵ est sans doute un objectif louable, encore faut-il parvenir à faire le départ entre les deux. Les catalogueurs n'accordent pas la même valeur aux mêmes tâches ; l'automatisation, selon les pans de l'activité qu'elle concerne, peut donc apporter un sentiment de libération à certains, mais en frustrer d'autres, selon l'axiologie de chacun. La stratégie de mise en œuvre d'un projet d'IA, le périmètre exact de celui-ci, doivent tenir compte de cette diversité, faute de quoi il risque de provoquer un sentiment de dépossession. La nécessité d'établir une cartographie de l'intérêt des tâches catalographiques est mentionnée par les porteurs du projet SGCAT de Singapour : « *What is cataloguing's "laundry"?* »³¹⁶ Quelle est la part ingrate du signalement ? Il est douteux qu'on puisse trancher universellement la question. Certes pour la plupart des documents, il existe des métadonnées descriptives sur lesquelles la majorité des catalogueurs s'accorderait, tandis que plusieurs vedettes-matière sont recevables³¹⁷. Mais cette distinction n'est pas superposable au potentiel d'automatisation : l'application de l'IA à l'indexation a été explorée bien avant qu'il ne devienne envisageable de convertir la numérisation d'une page de titre en notice MARC. Surtout, les critères d'objectivité ou de facilité d'application de l'IA ne tiennent pas compte de la satisfaction prise par les agents à accomplir les tâches touchées. Comme le dit Mélanie Roche, cheffe du service de gestion du Dépôt légal des périodiques :

Les bibliothécaires aiment des aspects différents de leur métier. L'appréciation varie du tout au tout. Certains diront : « Ma valeur ajoutée, c'est le choix d'une indexation adaptée au document. » D'autres affirmeront qu'ils n'ont pas les moyens de chercher les bonnes vedettes-matière, qu'ils hésitent parfois longtemps. Ceux-là seraient ravis que la machine s'en charge. Il y a autant d'avis que de catalogueurs, selon leurs affinités et ce qu'ils considèrent comme leurs atouts. Certains apprécient de maîtriser les règles sur le bout des doigts, de rectifier la ponctuation d'un titre, l'ajout d'une majuscule, tandis que d'autres s'en passeraient volontiers. Tous les catalogueurs du service adorent leur travail, mais ils en trouvent certains aspects plus ou moins enrichissants, et ces aspects ne sont jamais les mêmes d'un collègue à l'autre. C'est ce qui rend difficile la décision d'automatiser telle tâche au motif qu'elle serait unanimement ressentie comme rébarbative.³¹⁸

Au Département du Dépôt légal encore, Mathilde Koskas, cheffe du service de la Bibliographie nationale française, partage cette analyse :

Beaucoup de catalogueurs considèrent que l'indexation, la découverte des bonnes vedettes-matière, est la partie intellectuelle de leur travail, tandis que certains n'en font pas (tous les ouvrages reçus ne sont pas indexés). On peut trouver de la valeur intellectuelle et de la satisfaction professionnelle dans les finesses du catalogage signalétique, dans l'effort consenti pour rendre compte de l'ouvrage de la façon la plus fine. Pour lancer une application [d'IA catalographique]

³¹⁵ (LING et GOH, 2025, p. 9).

³¹⁶ (LING et GOH, 2025, p. 5).

³¹⁷ (LING et GOH, 2025, p. 6).

³¹⁸ (ROCHE, 2025).

dans le service, il faudrait donc veiller à ne pas soutenir péremptoirement que telles tâches sont répétitives, routinières, et d'autres non.³¹⁹

Dans une contribution dialoguée au *World Library and Information Congress* de l'IFLA à Astana à l'été 2025, Mélanie Roche et Mathilde Koskas ont abordé ce problème de la valeur assignée par chaque agent aux différents aspects du signalement. « *Technology's ideal contribution to human life is relieving us of tasks we don't want to do* » y dit Mélanie Roche, à quoi sa collègue répond : « *Which raises the question of what these tasks are: what is seen as not rewarding, not interesting, not cost-effective? The first two at least are subjective: some cataloguers revel in subject indexing while others favour the intricacies of a structured bibliographic description. Very few of them would like to stop cataloguing altogether.* »³²⁰ De telles réflexions représentent l'effort le plus abouti pour embrasser le point de vue des catalogueurs face à une éventuelle automatisation. Elles mettent l'accent non pas sur la productivité du circuit mais sur ce qu'éprouveraient les agents impliqués. Même si l'on se convainquait que les bibliothèques assurent des missions suffisamment vastes et variées pour que la crainte du chômage technique puisse être écartée, les agents qui ont consacré des années à affûter leurs compétences de signalement pourraient mal vivre un changement d'affectation : « *As managers we are not only custodians of collections and metadata but also responsible for the well-being and professional development of the people who report to us. Our ambition should be to make sure the right skills are used where they are needed, rather than deprive our colleagues of work they value and skills they and the institution have heavily invested in acquiring.* »³²¹

³¹⁹ (KOSKAS, 2025).

³²⁰ (ROCHE et KOSKAS, 2025, p. 3).

³²¹ (ROCHE et KOSKAS, 2025, p. 4).

FORMATION ET PROFILS TECHNIQUES

FEUILLE DE ROUTE

Cette partie – la cinquième et dernière – s'éloigne des catalogueurs pour se concentrer sur les détenteurs de l'expertise en ingénierie IA. Ce sont des profils différents, et la formation ne saurait suffire à faire passer un agent de l'un à l'autre. On observe que les équipes ayant réussi la mise en production d'un système semi-automatisé (DNB, ZBW) étaient pourvues d'une forte expertise interne. L'évolution de la technologie ne laisse pas penser qu'il en ira autrement à l'avenir. Le recours à des prestataires trace une ligne de partage des eaux, distinguant deux types de projets différents dans leur volume de traitement et champ d'application. La présence au sein de l'institution d'agents pourvus de ces compétences précieuses lui permet de conserver la pleine maîtrise de son catalogue.

Liens des sections :

↗ [Former en respectant les frontières professionnelles](#)

↗ [Des équipes pionnières pourvues d'une forte expertise](#)

↗ [Des profils techniques indispensables](#)

↗ [Prestataires ou compétences internes ?](#)

↗ [Des experts indispensables à la compréhension de l'IA](#)

6.1. Former en respectant les frontières professionnelles

La nécessité de proposer des formations à l'IA au personnel des bibliothèques est partout affirmée³²². L'une des conclusions du projet LibrarIN porte sur ce point : « *However, in order to be successful in the adoption and use of the tools, the countries need additional policies and funding to increase the digital literacy of library staff across the board—and not just for IT specialists. Similar to the growth of digital literacy among their patrons, libraries need the funds to invest in the digital literacy of their employees.* »³²³ Coralie Philibert concluait son récent mémoire sur une idée parallèle, affirmant que « les catégories B doivent être davantage formées aux *workflows* des données, à la veille technique, à la question de la souveraineté des données, à la maîtrise de l'environnement technique et fonctionnel de leur établissement documentaire »³²⁴.

Le besoin d'acquérir une maîtrise de la manipulation de données ou des enjeux de l'IA peut amener un remaniement des référentiels de compétences, mais la frontière avec le périmètre fonctionnel des services informatiques des établissements doit être respectée. Dans l'enquête de 2019 déjà citée, c'est un observateur extérieur au monde des bibliothèques qui propose de confondre les rôles : « *The modern librarian is a software engineer, who understands distribution mechanisms in the modern world.* »³²⁵ Au contraire, la littérature professionnelle insiste plutôt sur l'hybridation des compétences : le *data librarian* peut avoir une certaine connaissance de la programmation et des questions techniques nécessaires à ses missions, mais il ne doit pas endosser des fonctions étrangères différentes à sa profession³²⁶. Les témoignages recueillis par LibrarIN à la DNB soulignent cette même partition : « *[Librarians] don't have to be computer scientists, but they do need to be tech-savvy. We need people who can deal with large amounts of data and who have data management skills.* »³²⁷ De même, l'articulation avec le troisième pôle d'expertise nécessaire à tout projet d'IA, celui des *data scientists* capables d'entraîner ou mettre en œuvre le modèle, est nettement expliquée par le PCC : « *AI models in library metadata applications require expert validation from catalogers, metadata librarians, and archivists. Data scientists develop ML models, but SMEs define the metadata logic, classification rules, and quality standards.* »³²⁸ Dans un projet sortant de notre périmètre, consacré à la génération de résumés d'émissions télévisées, le succès est attribué à la collaboration entre ces trois corps de métier : « *Data*

³²² (BOATENG, 2025, p. 145), (YING-HSANG, 2025, p. 20), (COX, PINFIELD, et RUTTER, 2019) par exemple.

³²³ (MERGEL, KÜHLER, SCHUMANN et al., 2026, p. 193).

³²⁴ (PHILIBERT, 2024, p. 90).

³²⁵ (COX, PINFIELD, et RUTTER, 2019, p. 11).

³²⁶ (MESGUICH, 2023, p. 144).

³²⁷ (MERGEL, KÜHLER, SCHUMANN et al., 2026, p. 31).

³²⁸ (PCC TASK GROUP ON AI AND MACHINE LEARNING FOR CATALOGING AND METADATA, 2025b, p. 25).

scientists, developers, and cataloging experts have worked closely together throughout. »³²⁹ L'absence d'une expertise en IA au sein de l'établissement mène à des difficultés et au recours à des prestataires ou plateformes commerciales, conclut une méta-étude récente³³⁰ ; à son tour, la nécessité d'externaliser détermine la nature et l'amplitude des projets menés. Ce constat est fait par les établissements eux-mêmes : le NLB de Singapour cite parmi les freins à ses expérimentations l'absence d'expertise IA interne³³¹, que tempère quelque peu un service d'appui interministériel sur la question³³².

Les formations organisées pour les bibliothécaires ne sauraient viser à les mettre en état de mener intégralement, par eux-mêmes, un projet d'IA catalographique. Comme le disait il y a une décennie Luc Bellier à propos des projets en Big Data, il faut « utiliser des ressources humaines principalement issues de la fonction publique et disposant des compétences acquises durant leur formation »³³³. Ce sont bien des professions différentes qui doivent joindre leurs efforts, soulignent des membres de la DNB³³⁴. Même une politique de formation très riche n'y suffirait pas. Les webinaires internes organisés par la ZBW, par exemple, ne prétendent pas expliquer en profondeur les soubassements techniques d'AutoSE³³⁵. Les interrogations émises en 2016 par Gildas Illien à propos des projets informatiques et de la place que les bibliothécaires peuvent y prendre ont toute leur pertinence pour notre cas : « Comment les bibliothécaires peuvent-ils intervenir dans les processus de conception et de réalisation ? Avec quelles compétences et selon quelles responsabilités ? Quelles interactions faut-il mettre en place entre les métiers qui concourent au développement, et avec quels – outils de management ? »³³⁶

6.2. Des équipes pionnières pourvues d'une forte expertise

Les équipes ayant réussi, avant l'apparition des LLM, à créer des outils et des circuits d'indexation automatique comprenaient des membres-clefs pourvus d'une compétence informatique poussée. Argie Kasprzik, responsable d'AutoSE à la ZBW depuis 2018, est titulaire d'un doctorat en informatique théorique³³⁷. L'automatisation de l'indexation, jusqu'alors conçue comme un projet, acquiert à son arrivée un statut pérenne. Un doctorant, accueilli au sein de l'établissement, avait consacré sa thèse à la question, mais le code de ce prototype n'était pas satisfaisant. Un architecte logiciel a donc été recruté. La mise en ligne du service date de mai 2021. Les recherches appliquées nécessaires à l'entraînement du modèle ont d'abord été menées par le doctorant présent dès l'origine, puis par une doctorante à partir de 2023.

Osma Suominen, à la Bibliothèque nationale de Finlande, est lui aussi titulaire d'un doctorat en informatique. Il a commencé le développement d'Annif en 2017, après avoir convaincu la direction de son établissement. Il concentre deux des trois spécialités nécessaires à un tel projet, la familiarité avec l'IA et le savoir-faire informatique requis pour concevoir l'architecture du système et le mettre en production. Annif a rapidement remporté une compétition et l'Université de Jyväskylä l'a adopté. Ces succès initiaux ont mené à deux recrutements, dont celui de Mona Lehtinen en 2019³³⁸, grâce auxquels le service d'indexation Finto AI, reposant sur Annif, a pu être mis en ligne.

Les similitudes structurales entre les deux équipes sont frappantes : petite taille, forte expertise technique des membres, et enfin insertion dans un établissement à la culture favorable à la recherche, ainsi que l'observe LibrarIN à propos d'Annif : « *The library's substantial IT background benefits the institution by allowing for experimentation and the use of AI.* »³³⁹ Mona Lehtinen confirme cette analyse : « *We do a lot of in-house development: we like to code our own projects. Our whole ethos is centered on open source and having our own developers work closely with the library staff. That's one of our core values, a key aspect of how we function.* »³⁴⁰ Quant à la ZBW, il s'agit non seulement d'une bibliothèque mais aussi d'une véritable institution de recherche, un centre d'expertise et de compétence qui considère la création de prototypes de pointe tel AutoSE comme relevant de ses missions. L'institution est dotée

³²⁹ (TOFTGÅRD et RENDE, 2025).

³³⁰ (PINAR et COX, 2025, p. 585).

³³¹ (JAILANI, 2025, p. 13).

³³² (NG, 2026).

³³³ (BELLIER, 2017, p. 9).

³³⁴ (MÉRGEL, KÜHLER, SCHUMANN et al., 2026, p. 29).

³³⁵ (KASPRZIK, 2025a).

³³⁶ (ILLIEN, 2015, p. 105).

³³⁷ (KASPRZIK, 2025a).

³³⁸ (LEHTINEN, 2025b).

³³⁹ (SCHUMANN et MÉRGEL, 2026, p. 47).

³⁴⁰ (LEHTINEN, 2025b).

de quatre chaires professorales. En somme, dans les deux équipes décrites, l'atmosphère serait celle d'un laboratoire de recherche appliquée. La description du *AI-Lab* de la Bibliothèque nationale de Norvège par son directeur, Javier de la Rosa, donne une idée de la liberté d'initiative que de telles unités instituent en leur sein³⁴¹.

L'équipe de la DNB responsable de l'indexation automatique présente une caractéristique intéressante³⁴². Menée depuis 2019 par Christoph Poley, elle comprend six membres, non seulement des informaticiens (il l'est lui-même) et des mathématiciens, mais aussi des catalogueurs. Cette mixité permet d'avancer plus vite que s'il fallait échanger par le truchement de spécifications techniques et de réunions hebdomadaires avec le service de catalogage (qui ne traite à la DNB que des publications physiques). Celui-ci est consulté, mais dans une bien moindre mesure que les catalogueurs internes à l'équipe. Une telle composition permet des progrès plus rapides, un travail plus dynamique.

Le Labo de l'Abes correspond à une typologie similaire. En collaboration avec un informaticien expérimenté, Christophe Bonnefond, le responsable de cette structure, Yann Nicolas, conservateur des bibliothèques et codeur, a créé au fil du temps plusieurs outils utiles au réseau SUDOC (Paprika, par exemple, pour l'alignement d'autorités personnes). Le service expérimental d'assistance à l'indexation Rameau repose sur des techniques innovantes, qui ont bénéficié de la présence de stagiaires du master d'informatique, filière « Intelligence Artificielle et Science des Données » de l'Université de Montpellier. L'algorithme initialement retenu, compris dans la trousse à outils Annif, a été délaissé au profit d'une solution originale³⁴³, consistant à comparer le vecteur du texte associé à un document (titre, résumé, etc.) non pas au vecteur tiré du seul libellé de la vedette-matière, mais construit « à partir du contenu textuel de l'ensemble des notices Sudoc liées à ce concept »³⁴⁴. Cette idée s'écarte de la méthode *Embedding Based Matching* conçue par la DNB³⁴⁵, où seul le libellé de la vedette-matière est vectorisé.

6.3. Des profils techniques indispensables

Le groupe de travail du PCC sur l'IA, l'apprentissage machine et le signalement a donc raison d'insister sur l'importance des compétences au sein de l'équipe portant un projet d'IA catalographique, mais est-il certain que le rôle de *Data Technology Expert* puisse être rempli par un catalogueur, comme il le suggère³⁴⁶ ? Dans un projet qui se fonderait uniquement sur les LLM et abandonnerait des formes antérieures de l'IA – et c'est bien le type d'initiative décrit dans cette section du rapport –, à seule fin de générer les métadonnées descriptives d'une notice, une à une, avec l'aide d'un prestataire et sur une plateforme d'automatisation commerciale, cela semble possible d'après l'exemple de la KBR. Mais l'indexation est autrement compliquée. Certes des méthodes existent qui n'ont pas recours à l'apprentissage machine, comme celle de l'Abes (vectorisation et LLM) ou celle d'une équipe de la DNB, qui utilise des LLM pour la génération libre de termes d'indexation, alignés ensuite avec les entrées du vocabulaire contrôlé par comparaison de vecteurs³⁴⁷. Mais est-il certain qu'un tel travail, dont on mesure aisément la complexité technique, relève des attributions et des compétences prévues dans les fiches de poste des bibliothécaires aujourd'hui ? La complexité s'accroît encore si l'on ambitionne de construire un circuit semi-automatisé, dont la précision sera évaluée par prélèvements statistiques, avec calcul d'indicateurs variés au cours du temps. Il s'agit clairement d'un chantier mobilisant un savoir-faire différent de celui des experts du signalement.

Il suffit de consulter les articles, d'une densité technique remarquable, que Géraldine Geoffroy consacre aux LLM en bibliothèque³⁴⁸, pour se convaincre que renoncer à l'apprentissage machine ne dispense pas d'une profonde expertise technique. Et même si le codage agentique peut donner plus d'autonomie au personnel n'ayant pas de formation informatique poussée³⁴⁹, cet atout ne saurait suffire.

La difficulté de recruter des experts d'un niveau proportionné à la complexité technique de la mission, extérieurs aux filières de recrutement habituelles des bibliothèques, explique en partie que les

³⁴¹ (DE LA ROSA, 2025, p. 62).

³⁴² (POLEY, 2026).

³⁴³ (NICOLAS, 2025b, p. 16).

³⁴⁴ (NICOLAS, 2025b, p. 5).

³⁴⁵ (KAHLER, 2025, p. 12).

³⁴⁶ (PCC TASK GROUP ON AI AND MACHINE LEARNING FOR CATALOGING AND METADATA, 2025b, p. 20).

³⁴⁷ (KLUGE et KAHLER, 2025).

³⁴⁸ (GEOFFROY 2025a) et (GEOFFROY 2025b), déjà cités.

³⁴⁹ (NICOLAS, 2025a).

projets mentionnés ici aient germé dans des établissements nationaux : outre que leur rôle d'agence bibliographique s'accompagne de l'obligation légale de décrire une masse documentaire croissante, ils sont sans doute les seuls à pouvoir attirer et retenir les spécialistes indispensables³⁵⁰. En France, les bibliothèques universitaires offrent souvent des contrats de court terme, avec un niveau de rémunération sans commune mesure avec ce que le secteur privé propose à profil égal³⁵¹.

6.4. Prestataires ou compétences internes ?

On pourrait alors se demander si le recours à des prestataires externes constitue une option praticable. Plusieurs des bibliothèques étudiées la choisissent, et se noue parfois une relation de collaboration étroite, relevant de la création conjointe, où l'expertise catalographique des uns se marie harmonieusement avec la compétence en IA des autres³⁵². Mais il faut observer que les projets opérationnels les plus anciens sont tous le fruit d'efforts internes³⁵³. L'indexation automatique selon les concepts du GND était en place à la DNB depuis 2014 grâce à un marché ; entre 2019 et 2022, la bibliothèque dut créer un nouveau système et une nouvelle infrastructure pour s'en affranchir³⁵⁴, l'entreprise n'ayant plus convenance à continuer de maintenir et développer l'outil³⁵⁵. La ZBW a dès les années 2000 exploré des outils commerciaux d'indexation automatique (un essai s'est tenu en 2007 par exemple), mais le niveau de précision n'était pas suffisant, et l'institution a décidé en 2014 de n'utiliser ou de ne produire que des logiciels libres. S'ensuivit l'arrivée d'un doctorant, puis d'Argie Kasprzik qui porta, formalisa et finalisa le projet comme on l'a dit³⁵⁶. À notre connaissance, l'intérêt pour l'indexation automatique de la Bibliothèque de Finlande date de l'initiative d'Osma Suominen. Là encore, les conclusions du projet LibrarIN semblent particulièrement pertinentes : « *In the future, National Libraries will need high-level IT knowledge and skills to keep up with the current speed of development. This also includes the task of [...] having access to training and education, and intentionally recruiting highly skilled specialists with expertise in technology [...].* »³⁵⁷

Toutefois dans deux cas récents, un partenariat a mené à un outil mûr. Après la pandémie, la KBR s'est appuyée sur l'entreprise Initium pour concevoir un circuit d'exécution grâce à une plateforme de Microsoft³⁵⁸ dont la KB se sert aussi³⁵⁹, là encore grâce à une entreprise intermédiaire. S'il a été décidé de trouver un prestataire pour concevoir Retrotools, c'est que la loi hollandaise incite à externaliser les projets demandant des compétences absentes de l'institution, qu'il est difficile de recruter des développeurs dans le secteur public, et qu'une fois entrés à la bibliothèque, ils auraient de toute façon couru le risque d'être affectés à d'autres projets³⁶⁰.

Plusieurs points communs rapprochent Retrotools de son prédécesseur belge :

- L'architecture, c'est-à-dire le flux d'événements menant de l'envoi de la numérisation de l'ouvrage à l'obtention de la notice, est déléguée à la plateforme d'automatisation de Microsoft.
- Ils ont été conçus pour le catalogage rétrospectif et portent pour l'instant seulement sur la description bibliographique, non sur l'indexation, de documents imprimés. La solution d'apprentissage machine de Microsoft sur Power Apps n'est compatible qu'avec un vocabulaire inférieur à 200 mots³⁶¹. La plateforme permet certes de recourir à ChatGPT pour l'indexation, mais si elle doit puiser dans un vocabulaire contrôlé, les résultats seront infructueux sans une stratégie très sophistiquée d'orchestration en plusieurs étapes, telle la méthode de la DNB³⁶². La KB envisage d'utiliser Annif³⁶³ ;

³⁵⁰ (KOSKAS, 2025) m'a aidé à dégager cette idée.

³⁵¹ (BOYER, 2025p. 40-41).

³⁵² (COHEN, 2025), (BUESINK, 2025b).

³⁵³ La Bibliothèque du Congrès mène depuis plusieurs années avec un partenaire commercial un projet d'IA catalographique comprenant un volet d'indexation automatique, et quoique les résultats soient intéressants, il ne semble pas qu'il ait pour l'instant été mis en production. Cela n'est pas sans susciter une certaine frustration parmi les agents impliqués, voir (SCHUMANN, KÜHLER, et MERGEL, 2026p. 173).

³⁵⁴ (POLEY, UHLMANN, BUSSE et al., 2025, p. 4).

³⁵⁵ (MERGEL, KÜHLER, SCHUMANN et al., 2026, p. 27).

³⁵⁶ (KASPRZIK, 2025a).

³⁵⁷ (SCHMIDT, KÜHLER, SCHUMANN et al., 2026, p. 186). Nous supprimons des erreurs d'impression qui nuisent à la compréhension du passage.

³⁵⁸ (SCHUMANN et MERGEL, 2026, p. 86).

³⁵⁹ (VAN DER VAART et SCHMIDT, 2026, p. 97-98).

³⁶⁰ (BUESINK, 2025b).

³⁶¹ (LOWAGIE, 2025a).

³⁶² (KAHLER, 2025).

³⁶³ (LOWAGIE, 2025b).

le manuel écrit par Hannes Lowagie, s'il mentionne l'intégration de cet outil, n'en fait pas une part essentielle de la stratégie de signalement.³⁶⁴

- Ils prévoient qu'un catalogueur examine chaque notice créée par l'IA. Il s'agit seulement d'accélérer son travail.

Ces projets bénéficient des facilités d'organisation du circuit de traitement qu'offre une solution clef en main, qui plus est paramétrée par un prestataire. L'entraînement d'un modèle (apprentissage machine traditionnel) pour identifier sur la page de titre les différents champs ne nécessite qu'une poignée d'exemples, qu'on dépose sur la plateforme³⁶⁵ ; depuis que les LLM sont pour la plupart dotés de la capacité d'analyser des images, on peut même s'affranchir de cette étape. Simple d'utilisation et doté d'une interface graphique (dont pour l'instant Annif est dépourvu, comme le regrettent certains de ses utilisateurs³⁶⁶), on peut penser que *Power Apps* diminue le coût d'entrée technique d'un projet d'IA. Surtout, la KB et la KBR se concentrent sur la description bibliographique, choix qui évite d'affronter le problème touffu à l'extrême de l'indexation dans un vocabulaire contrôlé. Elles s'en tiennent à un versant du signalement où les LLM donnent d'excellents résultats, alors que, comme le dit Osmo Suominen, l'adoption d'Annif exige d'importantes ressources humaines :

*There are upsides and downsides of the open source model for library systems. It allows for freedom and flexibility, but requires more technical expertise and resources than similar systems and services provided by commercial vendors. Organizations adopting an open source solution must be prepared to build the in-house expertise required to set up and maintain the systems.*³⁶⁷

La DNB et la ZBW ont elles d'emblée conçu leur dispositif d'indexation pour traiter de vastes volumes de documents *digitaux*, sans examen humain de chaque notice. Ces publications comprenaient déjà des métadonnées descriptives fournies par les éditeurs³⁶⁸ ; il n'était donc pas nécessaire de recourir à l'IA pour automatiser cette partie-là du signalement. Ces deux groupes de projets se distinguent donc à plusieurs égards :

- Support des documents concernés : physiques / digitaux.
- Dimension du signalement : description bibliographique / indexation matière.
- Orientation première de la chaîne de traitement : traiter des arriérés / faire face à une masse croissante de publications électroniques.
- Implication du catalogueur : pour chaque notice / pour mener des campagnes de contrôle qualité.
- Modèles et expertise : un prestataire organise le recours à un LLM pré-existant pour la conversion de l'information en zones MARC / des spécialistes au sein de la bibliothèque entraînent des modèles par apprentissage machine traditionnel ou créent une orchestration de LLM sur-mesure.
- Architecture : circuit hébergé sur une plateforme commerciale / circuit créé et entretenu au sein de l'institution.

6.5. Des experts indispensables à la compréhension de l'IA

Former un vivier d'experts en IA au sein de l'établissement permettrait de conjurer, dans une certaine mesure, un risque qui hante la littérature professionnelle : celui d'une perte de maîtrise du catalogue, peuplé de métadonnées dont on ne saurait plus rendre raison. Le pire scénario serait que les services de catalogographie dépendent à l'avenir de boîtes noires, de modèles dont les oracles, lorsqu'ils sont faux, ne peuvent être expliqués. Or le signalement, lorsqu'il est public et partagé, forme un engagement implicite : la bibliothèque, surtout si elle est nationale, assume une responsabilité et garantit que ses notices sont assez fiables pour soutenir la recherche, la découverte et toute forme généralement de réutilisation possible au cours du « cycle de vie des données »³⁶⁹. Dans des écosystèmes où les notices circulent par dérivation (catalogues partagés, échanges de notices, réutilisation des autorités), il y a

³⁶⁴ (LOWAGIE, 2025a).

³⁶⁵ (LOWAGIE, 2025a, p. 50-55).

³⁶⁶ (INKINEN, LEHTINEN, et SUOMINEN, 2025, p. 6).

³⁶⁷ (SUOMINEN, INKINEN, et LEHTINEN, 2022, p. 279).

³⁶⁸ (JUNGER, 2018, p. 2).

³⁶⁹ (MESGUICH, 2023, p. 39, 87-88).

là une question de confiance entre institutions et de maintien de l'idéal de contrôle bibliographique universel.

Cette notion recouvre la production et le partage de métadonnées fiables, en temps utile, dans des cadres coopératifs, entre agences bibliographiques nationales, éditeurs, bibliothèques consommatrices de notices et les usagers finaux³⁷⁰. Dans un service de dépôt légal, l'enjeu est d'assurer aux citoyens un accès à des informations de qualité et garanties par un agent public : le catalogage précis des entrées constitue une production culturelle nationale, au service de la démocratie³⁷¹. Or tant que des évaluations n'ont pas administré la preuve de l'efficacité d'une IA catalographique, l'incertitude plane. Une fois conduites, leur interprétation fine révèle parfois des résultats disparates : dans le projet d'assistance au catalogage de la Bibliothèque du Congrès, l'IA remplit correctement les champs de titre et d'auteur, mais échoue sur l'indexation³⁷².

Sous la plupart de ses formes, l'IA, souffrant d'une opacité structurelle, ne peut rendre compte de ses décisions³⁷³. Dans une liste récente de dix-huit principes que l'humain est en droit d'exiger de l'IA en général, la cohérence et la fiabilité des résultats, l'explicabilité et la transparence sur les degrés de confiance forment une section entière³⁷⁴. L'explicabilité revient comme un leitmotiv dans douze déclarations de principes et préconisations relatives à l'éthique de l'IA, choisies par Luciano Floridi pour leur haute réputation³⁷⁵. Un LLM peut gloser les choix qu'il a faits, mais rien n'assure que l'explication fournie corresponde à la réalité des étapes de « raisonnement » parcourues, et cela n'apporte aucun renseignement sur la cause de l'erreur, au niveau fondamental des *poids* qui forment la particule élémentaire de ces constructions³⁷⁶.

Dès lors que l'aspiration à une compréhension profonde, mathématique, des façons dont l'IA atteint ses conclusions ne peut être pour l'instant satisfaite, il faut s'attacher à garder toujours trace de la provenance des métadonnées : savoir quand et comment telle forme d'IA y a contribué, quels documents ont servi d'entrée, quelles règles ont été appliquées, etc. Cette traçabilité est cruciale à la fois pour le contrôle de qualité et pour la transparence vis-à-vis des réutilisateurs³⁷⁷. Elle est au cœur des préoccupations des concepteurs de chaînes de traitement agentique, tel le système Ordo mis en place par la Bibliothèque nationale de Suède, qui peut solliciter plusieurs API pour puiser l'information là où il en a besoin : « *The agentic approach in this case serves a dual purpose: it enables automation using multiple backend capabilities, but it also provides transparency and fine-grained insight into the process. It allows a user to take the reins when needed, perform an audit at any time or simply leave it as is while observing the inner workings of the process.* »³⁷⁸ Il en va ainsi aussi un circuit similaire en cours de construction à la bibliothèque universitaire métropolitaine de Manchester, où l'une des priorités est de permettre à un humain d'examiner la cascade de décisions prises par l'IA³⁷⁹. Dans un ambitieux système décrit par une spécialiste française de la question, les chaînes d'inférence d'une IA agentique préservées de façon structurée permettent un audit ultérieur³⁸⁰. On est là aux antipodes des solutions intégrées aux SIGB commerciaux. Comparant un système libre et auditable à la fonction AI Metadata Assistant introduit dans le SIGB Alma par Ex Libris, Géraldine Geoffroy écrit :

... outre le problème de gouvernance (cet assistant s'appuie en fait sur GPT 4o) et le manque à ce jour de transparence du dispositif (aucun prompt ni code publié, aucune documentation technique sur « l'entraînement » du modèle...), le problème induit par ce type d'outillage intégré dans les logiciels de gestion des bibliothèques est d'embarquer de fait les bibliothécaires dans un type d'usage de l'IA sans contrôle a priori ni sur le modèle utilisé, ni sur les modalités ou le périmètre applicatifs.³⁸¹

³⁷⁰ Voir (IFLA BIBLIOGRAPHY SECTION, 2025, p. 1).

³⁷¹ (KOSKAS, 2025).

³⁷² Voir (SACCUCCI et POTTER, 2025, p. 8-9).

³⁷³ (VUARIN et STEYER, 2025, p. 244) et (DEVILLERS, 2024, p. 137).

³⁷⁴ (FISHER, 2025, p. 6).

³⁷⁵ (FLORIDI, 2023, p. 116-130).

³⁷⁶ Grâce aux efforts considérables de recherche consacrés à l'interprétation du fonctionnement des LLM, peut-être en ira-t-il un jour autrement.

³⁷⁷ Voir (BRYANT 2025, p. 3).

³⁷⁸ (MALMSTEN et LUNDBORG, 2025, p. 2-3).

³⁷⁹ (KORPET et REES, 2025, p. 34).

³⁸⁰ (GEOFFROY 2025b, p. 1).

³⁸¹ (GEOFFROY 2025a, p. 2).

L'enregistrement d'un journal de bord d'exécution permet de parcourir les chaînes de raisonnement d'un modèle, de voir quels outils il a utilisés (consultation de la documentation, d'une API externe, etc.) ; un tel dispositif donne une prise concrète sur le problème de l'explicabilité.

CONCLUSION

7.1. Synthèse des observations

Au terme de notre étude, les questions posées initialement trouvent, selon les établissements, des réponses variées mais convergentes. Que l'IA serve à accélérer le traitement unitaire d'un lot rétrospectif physique en générant des métadonnées descriptives (KBR, KB), ou à indexer les centaines de milliers de documents digitaux que l'extension des obligations légales impose aux bibliothèques nationales de collecter (DNB, ZBW), l'outil sert à pallier des ressources humaines insuffisantes. Ce déséquilibre peut être conjoncturel, dû à des arriérés accumulés pour des causes ponctuelles (pandémie par exemple), ou structurel dans le cas du flot sans cesse croissant d'articles électroniques. On observe moins un redéploiement délibéré des agents, que des stratégies visant à continuer d'accomplir les missions de signalement malgré une stagnation ou perte d'effectifs, due par exemple au non-remplacement du personnel prenant sa retraite.

L'expertise catalographique reste indispensable, mais elle se manifeste sous d'autres formes : relecture et correction des notices produites par l'IA à partir des parties-clefs du livre numérisé, ou bien échantillonnage et annotation de corpus pour affiner les modèles. En cas de mise en place d'un circuit semi-automatisé, les catalogueurs ont un rôle crucial à jouer au sein de l'équipe technique qui en est responsable (DNB), ou en étroite collaboration avec celle-ci (ZBW). La question de la qualité est ancrée au cœur des deux types de chaînes : dans le premier, un humain contrôle une à une les notices produites, dans le second, la pertinence des vedettes-matière assignées est régulièrement évaluée. L'objectif dans cette dernière catégorie de projets est de proposer à l'utilisateur une indexation même marginalement perfectible plutôt que de ne lui en fournir aucune.

Nous avons vu que les chaînes de traitement semi-automatisées sont le fait d'équipes pourvues de fortes compétences en apprentissage machine. Cette expertise interne semble la condition indispensable à la réussite des projets, et constitue l'une des raisons pour lesquelles ils sont si coûteux en ressources humaines et financières. Ce personnel précieux garantit, à long terme, l'explicabilité des outils et facilite leur intégration dans le SI des établissements.

Nous avons fait état des inquiétudes, très présentes dans la littérature, que suscitent un éventuel abandon de la rédaction manuelle de notices et la mue du catalogueur en relecteur. Les problèmes subséquents (perte d'expertise, de vigilance, de sens) ne semblent pas se rencontrer dans les circuits étudiés, en raison de leur configuration : avec Retrotools ou l'outil de rétrocatalogage de la KBR, le catalogueur traite les ouvrages un à un, de sorte que l'expérience ressemble à la dérivation depuis une base externe, où la notice existerait déjà – il n'y a pas là de rupture majeure avec les pratiques actuelles. À la ZBW ou à la DNB, les équipes de catalogueurs existantes poursuivent leurs activités, uniquement sur des documents physiques pour cette dernière.

Le but des projets considérés n'est pas de réformer les codes de catalogage ni les formats, mais de remployer les métadonnées dans leur structure actuelle. On attend ainsi des modèles une indexation conforme aux vocabulaires contrôlés sur lesquels ils ont été entraînés ; il se peut que ces classifications du savoir évoluent pour faciliter le traitement automatisé, mais si l'on excepte telle expérience de pensée du KBLab en Suède, les établissements ne paraissent pas désireux de les abandonner.

7.2. Vers une génération semi-automatique des métadonnées descriptives ?

Avant de récapituler les enseignements opérationnels qu'on peut tirer de cette étude, permettons-nous une prédiction. Dans la matrice que nos observations dessinent, on a pu noter qu'il n'existait pas encore d'outil d'assistance ou d'automatisation partielle portant sur les métadonnées descriptives conçu par des équipes internes. L'augmentation continue des publications physiques ou le souhait de traiter des documents digitaux dépourvus de toute information bibliographique, combinés à l'érosion graduelle du nombre de catalogueurs, pourrait rendre l'approche illustrée par la KBR attrayante. Mais des établissements sourcilieux sur leur souveraineté technologique voudront sans doute contourner

la plateforme d'automatisation d'une entreprise comme Microsoft. Pour rester dans un circuit clos, le recours à des modèles suffisamment légers pour passer à l'échelle sans quitter des serveurs étatiques sera certainement envisagé.

De telles initiatives se trouveraient favorisées par le travail expérimental que mène la Bibliothèque nationale de Finlande autour de la génération de métadonnées descriptives grâce à des LLM de petite taille³⁸². La DNB prévoit par exemple d'explorer cette voie³⁸³, la moitié des ouvrages physiques reçus ne présentant pas de métadonnées.

7.3. Enseignements

Lorsqu'une chaîne de traitement est entièrement construite autour de l'IA, comme c'est le cas à la ZBW ou à la DNB, la place dévolue à l'humain, les compétences requises, le *modus operandi* changent au point de transformer l'organisation. De tels services de catalogage présentent un aspect très différent des services de dépôt légal traditionnels, où des dizaines de catalogueurs œuvrent côte à côte pour venir à bout, livre après livre, d'un labeur titanesque. Cette technologie est bien de celles qui reconfigurent et remodelent en profondeur l'institution qui y recourt³⁸⁴. Même appliquée au seul signalement, l'IA n'est pas sans affecter l'essence des bibliothèques³⁸⁵. C'est à très juste titre que les enquêtes de LibrarIN contenaient tout un volet sur les répercussions des projets d'IA catalographiques sur l'organisation³⁸⁶.

La magnitude de cette métamorphose, concomitante de l'extension du dépôt légal au domaine numérique, doit inciter qui voudrait conduire pareille entreprise à mener d'abord des études approfondies de faisabilité et d'impact. Les sommes à trouver, la durée de l'effort à soutenir, les ressources humaines à réunir sont considérables : le but poursuivi justifie-t-il une telle mobilisation ? Est-ce un dispositif d'assistance qu'on envisage, pour faciliter et accélérer un signalement unitaire, ou bien des opérations semi-automatisées, contrôlées par échantillonnage ? La difficulté dans ce dernier cas, nullement insurmontable, doit être mesurée avec lucidité : l'expérience accumulée par les services pionniers allemands, partagée dans de nombreux rapports, pave une voie encore facilitée par l'existence d'Annif.

Cette science ouverte, dûment défendue³⁸⁷, aplanit quelque peu les aspérités techniques, mais les versants humains, organisationnels du projet ne requièrent pas un examen moins attentif. Comment obtenir l'adhésion des catalogueurs, dont l'expertise sera capitale ? Risquent-ils, en se consacrant exclusivement à la relecture, de perdre le goût du travail et un savoir-faire que le nouveau circuit ne permettra plus à leurs successeurs d'acquérir ? Nous avons déduit ce danger théorique de nos lectures et de pratiques préventives observées dans des services où l'IA n'est pas encore entrée ; il ne constitue pas une préoccupation majeure de la ZBW ou de la DNB, où le traitement unitaire perdure dans d'autres branches de l'organisation. Il convient de le garder à l'esprit, sans s'exagérer la menace multiforme qu'il fait planer ni l'écarter hâtivement.

Ces considérations, relevant de la GPEC, doivent s'étendre à des profils moins courants en bibliothèque. À moins de se limiter à un outil de type Retrotool ou SGCAT et de recourir à un prestataire, l'apprentissage machine traditionnel a toute chance d'être indispensable. Comment obtenir, au sein de l'institution ou par un recrutement, les compétences en ingénierie IA dont on ne peut faire l'économie ? Les options techniques à l'étude conviennent-elles finement au problème à résoudre ? Les LLM prêts à l'emploi qu'on trouve dans le commerce, dans leur interface conversationnelle de base, ne se prêtent guère à l'indexation automatique efficace de centaines de milliers de documents, mais un LLM de petite taille, inséré dans un circuit sur mesure, peut suffire à identifier les métadonnées descriptives que porte une page de titre, etc.

Il n'y a là aucune impossibilité majeure, mais une série de questions à considérer méticuleusement. Nous espérons que notre travail fournira quelques éléments de réflexion utiles pour relever avec l'IA les défis du signalement.

³⁸² (BEAUGUITTE et SUOMINEN, 2024), (LEHTINEN, 2025ap. 11).

³⁸³ (POLEY, 2026).

³⁸⁴ (BAPTISTA, STEIN, KLEIN et al., 2020, p. 9).

³⁸⁵ (COX, PINFIELD, et RUTTER, 2019, p. 17) prévoyait il y a plusieurs années déjà que l'IA ne serait pas sans conséquence pour la définition des bibliothèques.

³⁸⁶ (MERGEL, SCHMIDT, SCHUMANN et al., 2024, p. 10).

³⁸⁷ (KASPRZIK, 2025b, p. 15).

SOURCES

Bibliographies

Awesome AI for LAM, [sans date]. *awesome-ai4lam* [en ligne] [Consulté le 30 mai 2025]. Disponible à l'adresse : <https://ai4lam.github.io/awesome-ai4lam/>

MEUNIER, Adrien, [sans date]. LibGuides: IA et les métiers de l'information, de la documentation et des bibliothèques: Cerner le sujet. [en ligne]. [Consulté le 29 décembre 2025]. Disponible à l'adresse : <https://enssib.libguides.com/c.php?g=716767&p=5193422>

PCC, [sans date]. PCC AI Resources. *PCC AI Resources* [en ligne]. [Consulté le 29 décembre 2025]. Disponible à l'adresse : https://docs.google.com/document/d/1muieHY-K2CvW_Y8wdJ6WbIXBe4eHnoPFLPv4TizWSOo/edit?usp=sharing&pli=1&authuser=0&usp=embed_facebook

PROJET DE RECHERCHE FAN (FULLY ALGORITHMIC LIBRARIAN), [sans date]. Bibliographie sur l'IA en bibliothèque. *Bibliographie sur l'IA en bibliothèque* [en ligne]. [Consulté le 29 décembre 2025]. Disponible à l'adresse : <https://fan.zib.de/>

IA en bibliothèque (transversal)

Généralités

BOATENG, Frank, 2025. The Transformative Potential of Generative AI in Academic Library Access Services: Opportunities and Challenges. *Information Services and Use* [en ligne]. 1 mai 2025. Vol. 45, n° 1-2, pp. 140-147. [Consulté le 22 juillet 2025]. DOI [10.1177/18758789251332800](https://doi.org/10.1177/18758789251332800). Disponible à l'adresse : <https://doi.org/10.1177/18758789251332800>

BRYANT, Rebecca, 2024. Futureproofing Library Teams. *Hanging Together* [en ligne]. 19 novembre 2024. [Consulté le 29 décembre 2025]. Disponible à l'adresse : <https://hangingtogether.org/futureproofing-library-teams/>

COX, Andrew M. et WANG, Xuemao, 2025. Artificial Intelligence in Libraries: The Emerging Research Agenda. *IFLA Journal*. 2025. Vol. 51, n° 3, pp. 567-569.

COX, Andrew, 2024a. Ethics Case Studies of Artificial Intelligence for Library and Information Professionals. In : BALNAVES, Edmund, BULTRINI, Leda, COX, Andrew M., *et al.* (éd.), *New Horizons in Artificial Intelligence in Libraries* [en ligne]. De Gruyter. pp. 156-168. [Consulté le 3 juin 2025]. Disponible à l'adresse : <https://library.oapen.org/handle/20.500.12657/98734>

COX, Andrew M., 2022. How Artificial Intelligence Might Change Academic Library Work: Applying the Competencies Literature and the Theory of the Professions. *Journal of the Association for Information Science and Technology* [en ligne]. 7 mars 2022. Vol. 74, n° 3, pp. 367-380. [Consulté le 20 juillet 2025]. DOI [10.1002/asi.24635](https://doi.org/10.1002/asi.24635). Disponible à l'adresse : <https://onlinelibrary.wiley.com/doi/10.1002/asi.24635>

COX, Andrew M., PINFIELD, Stephen et RUTTER, Sophie, 2019. The Intelligent Library: Thought Leaders' Views on the Likely Impact of Artificial Intelligence on Academic Libraries. *Library Hi Tech* [en ligne]. 16 septembre 2019. Vol. 37, n° 3, pp. 418-435. [Consulté le 24 juillet 2025]. DOI [10.1108/lht-08-2018-0105](https://doi.org/10.1108/lht-08-2018-0105). Disponible à l'adresse : <https://www.emerald.com/insight/content/doi/10.1108/LHT-08-2018-0105/full/html>

- COX, Andrew et HAJIBAYOVA, Lala, 2025. Introduction: Artificial Intelligence (AI) and Ethical Representation, Organization, and Discovery of Knowledge. *Cataloging & Classification Quarterly* [en ligne]. 3 octobre 2025. Vol. 63, n° 6–7, pp. 403-407. [Consulté le 29 janvier 2026]. DOI [10.1080/01639374.2025.2558657](https://doi.org/10.1080/01639374.2025.2558657). Disponible à l'adresse : <https://doi.org/10.1080/01639374.2025.2558657>
- HANEGAN, Michael et ROSSER, Chris, 2025. *Generative AI and Libraries: Claiming Our Place in the Center of a Shared Future*. Chicago : ALA Editions. ISBN 979-8-89255-310-0.
- HARPER, Corey et GROTH, Paul, 2025. The Librarian Skillset and the Needs of Artificial Intelligence. *Cataloging & Classification Quarterly* [en ligne]. 3 octobre 2025. Vol. 63, n° 6–7, pp. 546-565. [Consulté le 15 novembre 2025]. DOI [10.1080/01639374.2025.2539787](https://doi.org/10.1080/01639374.2025.2539787). Disponible à l'adresse : <https://doi.org/10.1080/01639374.2025.2539787>
- KONCH, Prasanna Kumar, HAZARIKA, Hirak Jyoti, HAZARIKA, Pranjali, *et al.*, 2025. Enhancing Library Services with AI Solutions: Empowering Librarians Through Cutting-Edge Tools—A Study. In : NAGAR, Atulya, JAT, Dharm Singh, MISHRA, Durgesh, *et al.* (éd.), *Intelligent Sustainable Systems*. Singapore : Springer Nature. 2025. pp. 85-95. ISBN 978-981-97-9327-3.
- MANNHEIMER, Sara, BOND, Natalie, YOUNG, Scott W. H., *et al.*, 2024. Responsible AI Practice in Libraries and Archives: A Review of the Literature. *Information Technology and Libraries* [en ligne]. 23 septembre 2024. Vol. 43, n° 3. [Consulté le 4 février 2026]. DOI [10.5860/ital.v43i3.17245](https://doi.org/10.5860/ital.v43i3.17245). Disponible à l'adresse : <https://ital.corejournals.org/index.php/ital/article/view/17245>
- MEESAD, Phayung et MINGKHWAN, Anirach, 2024. AI-Powered Smart Digital Libraries. In : MEESAD, Phayung et MINGKHWAN, Anirach (éd.), *Libraries in Transformation: Navigating to AI-Powered Libraries* [en ligne]. Cham : Springer Nature Switzerland. pp. 391-428. [Consulté le 29 juillet 2025]. ISBN 978-3-031-69216-1. Disponible à l'adresse : https://doi.org/10.1007/978-3-031-69216-1_10
- MERGEL, Ines, 2026. Studying the Co-production of AI Implementation in National Public Libraries. In : MERGEL, Ines et SCHMIDT, Carsten (éd.), *AI Innovations in Public Services: The Case of National Libraries* [en ligne]. Cham : Springer Nature Switzerland. pp. 3-10. [Consulté le 23 décembre 2025]. ISBN 978-3-032-01344-6. Disponible à l'adresse : https://doi.org/10.1007/978-3-032-01344-6_1
- MURPH, Nicole Lucero, 2025. Beyond Information Literacy: Exploring AI's Impact on Labor in Academic Libraries. *Library Trends* [en ligne]. 2025. Vol. 73, n° 4, pp. 609-627. [Consulté le 29 décembre 2025]. Disponible à l'adresse : <https://muse.jhu.edu/pub/1/article/968498>
- MWANTIMWA, Kelefa et MSOFFE, Grace, 2025. Application of Generative Artificial Intelligence in Library Operations and Service Delivery: A Scoping Review. *Technical Services Quarterly* [en ligne]. 3 avril 2025. Vol. 42, n° 2, pp. 139-168. [Consulté le 29 janvier 2026]. DOI [10.1080/07317131.2025.2467574](https://doi.org/10.1080/07317131.2025.2467574). Disponible à l'adresse : <https://doi.org/10.1080/07317131.2025.2467574>
- NOGUEIRA, Leticia Antunes, MOLTUBAKK, Stine Thordarson, FAGERVIK, Andreas, *et al.*, 2025. Cutting through the Noise: Assessing Tools That Employ Artificial Intelligence. *IFLA Journal* [en ligne]. 1 octobre 2025. Vol. 51, n° 3, pp. 735-749. [Consulté le 15 novembre 2025]. DOI [10.1177/03400352241304121](https://doi.org/10.1177/03400352241304121). Disponible à l'adresse : <https://doi.org/10.1177/03400352241304121>
- PINAR, Abdulhalik et COX, Andrew, 2025. An Analysis of Artificial Intelligence (AI) Capability in Libraries and Archives. *Cataloging & Classification Quarterly* [en ligne]. 3 octobre 2025. Vol. 63, n° 6–7, pp. 566-599. [Consulté le 30 janvier 2026]. DOI [10.1080/01639374.2025.2539790](https://doi.org/10.1080/01639374.2025.2539790). Disponible à l'adresse : <https://doi.org/10.1080/01639374.2025.2539790>

SCHMIDT, Carsten, KÜHLER, Justus, SCHUMANN, Anna-Lea, *et al.*, 2026. Recommendations for the AI Implementation in Libraries. In : MERGEL, Ines et SCHMIDT, Carsten (éd.), *AI Innovations in Public Services: The Case of National Libraries* [en ligne]. Cham : Springer Nature Switzerland. pp. 181-197. [Consulté le 23 décembre 2025]. ISBN 978-3-032-01344-6. Disponible à l'adresse : https://doi.org/10.1007/978-3-032-01344-6_14

SLATER, Kailyn, 2025. Against AI: Critical Refusal in the Library. *Library Trends* [en ligne]. 2025. Vol. 73, n° 4, pp. 588-608. [Consulté le 29 décembre 2025]. Disponible à l'adresse : <https://muse.jhu.edu/pub/1/article/968497>

UZWYSHYN, Raymond, 2024. Projects in Machine Learning and Natural Language Processing in Libraries: An Introductory Overview. In : BALNAVES, Edmund, BULTRINI, Leda, COX, Andrew M., *et al.* (éd.), *New Horizons in Artificial Intelligence in Libraries* [en ligne]. De Gruyter. pp. 171-178. [Consulté le 3 juin 2025]. Disponible à l'adresse : <https://library.oapen.org/handle/20.500.12657/98734>

WILLEMS, Marieke, ANDUCAS, Marta et SILVESTRINI, Elena, 2026. Innovating Libraries: The LibrarIN Toolkit for Policy and Practice. In : MERGEL, Ines et SCHMIDT, Carsten (éd.), *AI Innovations in Public Services: The Case of National Libraries* [en ligne]. Cham : Springer Nature Switzerland. pp. 199-222. [Consulté le 23 décembre 2025]. ISBN 978-3-032-01344-6. Disponible à l'adresse : https://doi.org/10.1007/978-3-032-01344-6_15

YING-HSANG, Liu, 2025. Navigating the AI-Driven Metadata Landscape: A Human Centered Approach. In : *Metadata Standards Matter : Building a Sustainable Future for Information* [en ligne]. Athènes. 2025. Disponible à l'adresse : <https://iflacos.eu/index.php/abstracts/#N>

Signalement

SemEval'25 Task 5, [sans date]. [en ligne] [Consulté le 7 juillet 2025]. Disponible à l'adresse : <https://sites.google.com/view/llms4subjects/home>

ANDERSON, Martha et TIAN, Tang Cindy, 2025. Harnessing AI for Metadata Innovation: An Interview with Cindy Tian. *Technical Services Quarterly* [en ligne]. 2 octobre 2025. Vol. 42, n° 3-4, pp. 361-369. [Consulté le 29 janvier 2026]. DOI [10.1080/07317131.2025.2571141](https://doi.org/10.1080/07317131.2025.2571141). Disponible à l'adresse : <https://doi.org/10.1080/07317131.2025.2571141>

BRZUSTOWICZ, Richard, 2023. From ChatGPT to CatGPT: The Implications of Artificial Intelligence on Library Cataloging. *Information Technology and Libraries* [en ligne]. 18 septembre 2023. Vol. 42, n° 3. [Consulté le 20 juillet 2025]. DOI [10.5860/ital.v42i3.16295](https://doi.org/10.5860/ital.v42i3.16295). Disponible à l'adresse : <https://ital.corejournals.org/index.php/ital/article/view/16295>

CHEN, Suzhen et AND LI, Mingyan, 2024. AI for Cataloging and Metadata Creation: Perspectives and Future Opportunities from Cataloging and Metadata Professionals. *Technical Services Quarterly* [en ligne]. 1 octobre 2024. Vol. 41, n° 4, pp. 317-332. [Consulté le 22 juin 2025]. DOI [10.1080/07317131.2024.2394919](https://doi.org/10.1080/07317131.2024.2394919). Disponible à l'adresse : <https://doi.org/10.1080/07317131.2024.2394919>

DENG, Sai, 2023. AI, Cataloging & Metadata. *Faculty Scholarship and Creative Works* [en ligne]. 15 novembre 2023. Disponible à l'adresse : <https://stars.library.ucf.edu/ucfscholar/1251>

D'SOUZA, Jennifer, SADRUDDIN, Sameer, ISRAEL, Holger, *et al.*, 2025. SemEval-2025 Task 5: LLMs4Subjects – LLM-based Automated Subject Tagging for a National Technical Library's Open-Access Catalog. *arXiv* [en ligne]. mai 2025. [Consulté le 22 juillet 2025]. Disponible à l'adresse : <http://arxiv.org/abs/2504.07199>

ENGEL, Jonathan Yehuda, DO, Dan Tam, SALEM, Brenda, *et al.*, 2025. Artificial Intelligence in Library Cataloging: A Review of Literature. *Journal of Library Metadata* [en ligne]. 10 juillet 2025. Vol. 25, n° 4, pp. 1-16. [Consulté le 20 juillet 2025]. DOI [10.1080/19386389.2025.2526913](https://doi.org/10.1080/19386389.2025.2526913). Disponible à l'adresse : <https://doi.org/10.1080/19386389.2025.2526913>

FREEMAN, Samantha Baine, 2025. Ghost in the Machine: Metadata, Commodification, and the Specter of Artificial Intelligence. *Journal of Library Metadata* [en ligne]. 3 juillet 2025. Vol. 25, n° 3, pp. 223-239. [Consulté le 20 juillet 2025]. DOI [10.1080/19386389.2025.2523717](https://doi.org/10.1080/19386389.2025.2523717). Disponible à l'adresse : <https://doi.org/10.1080/19386389.2025.2523717>

FRENCH, Rebecca B., 2025. A Framework for Metadata Workflows. *Cataloging & Classification Quarterly* [en ligne]. 2 janvier 2025. Vol. 63, n° 1, pp. 1-11. [Consulté le 22 juillet 2025]. DOI [10.1080/01639374.2025.2450391](https://doi.org/10.1080/01639374.2025.2450391). Disponible à l'adresse : <https://doi.org/10.1080/01639374.2025.2450391>

GEOFFROY, Géraldine, 2025a. Penser avant de cataloguer : entraîner un LLM de raisonnement sur des métadonnées bibliographiques. *IA & Bibliothèques* [en ligne]. 20 mai 2025. [Consulté le 12 janvier 2026]. Disponible à l'adresse : <https://iaetbibliotheques.fr/2025/05/penser-avant-de-cataloguer-entraîner-un-llm-de-raisonnement-sur-des-metadonnees-bibliographiques>

GEOFFROY, Géraldine, [sans date]. SmartBibl.IA - Ingénierie documentaire et Intelligence Artificielle. [en ligne]. [Consulté le 29 décembre 2025]. Disponible à l'adresse : <https://smartbiblia.fr/>

GEOFFROY, Géraldine, 2025b. Standards pour une IA distribuée et auditable par design. *IA & Bibliothèques* [en ligne]. 10 octobre 2025. [Consulté le 27 décembre 2025]. Disponible à l'adresse : <https://iaetbibliotheques.fr/2025/10/standards-pour-une-ia-distribuee-et-auditable-par-design>

ILLIEN, Gildas, 2016. Transition Bibliographique : La Revanche Des Données ?. *Archimag*. mars 2016. Vol. 292, pp. 20-23.

MERGEL, Ines, SCHMIDT, Carsten, SCHUMANN, Anna-Lea, *et al.*, 2024. *D3.2 Digital Transformation and ICT v2.0*. Horizon Europe Project LibrarIN.

MOULAISON-SANDY, Heather et AND COBLE, Zach, 2024. Leveraging AI in Cataloging: What Works, and Why?. *Technical Services Quarterly* [en ligne]. 1 octobre 2024. Vol. 41, n° 4, pp. 375-383. [Consulté le 22 juin 2025]. DOI [10.1080/07317131.2024.2394912](https://doi.org/10.1080/07317131.2024.2394912). Disponible à l'adresse : <https://doi.org/10.1080/07317131.2024.2394912>

OYIGHAN, Diseiye, UKUBEYINJE, Ejio Sandra, DAVID-WEST, Boma T., *et al.*, 2024. The Role of AI in Transforming Metadata Management: Insights on Challenges, Opportunities, and Emerging Trends. *Asian Journal of Information Science and Technology* [en ligne]. 26 septembre 2024. Vol. 14, n° 2, pp. 20-26. [Consulté le 20 juillet 2025]. DOI [10.70112/ajist-2024.14.2.4277](https://ajist.co/index.php/ajist/article/view/4277). Disponible à l'adresse : <https://ajist.co/index.php/ajist/article/view/4277>

RONCAGLIA, Gino, 2026. Cataloguing, Metadata, and Generative AI. Early Experiences and Future Perspectives. *JLIS.it* [en ligne]. 15 janvier 2026. Vol. 17, n° 1, pp. 106-127. [Consulté le 30 janvier 2026]. DOI [10.36253/jlis.it-693](https://doi.org/10.36253/jlis.it-693). Disponible à l'adresse : <https://jlis.it/index.php/jlis/article/view/693>

ROY, Soumendu, MALLIKRAJ, S. Vijay, MORADIA, Shivam, *et al.*, 2024. The Impact of Artificial Intelligence on Cataloging and Classification Systems in Modern Libraries. *Library Progress International* [en ligne]. 6 septembre 2024. Vol. 44, n° 3, pp. 769-773. [Consulté le 20 juillet 2025]. DOI [10.48165/bapas.2024.44.2.1](https://doi.org/10.48165/bapas.2024.44.2.1). Disponible à l'adresse : <https://bapasjournals.com/library-science/index.php/journal/article/view/468>

SNOW, Karen et BOTHMANN, Bobby, 2025. Toward a Unified Framework of Global Cataloging & Metadata Competencies. In : *Metadata Standards Matter : Building a Sustainable Future for Information* [en ligne]. Athènes. 2025. Disponible à l'adresse : <https://iflacos.eu/index.php/abstracts/#K>

SUSSMEIER, Stephanie et HENRY, Joshua A., 2025. Mind the Gap!: How Do We Ethically Bridge the Divide Between the Cataloging/Metadata Community and the World of AI?. *Journal of Library Metadata* [en ligne]. 30 juin 2025. Vol. 25, n° 4, pp. 1-19. [Consulté le 20 juillet 2025]. DOI [10.1080/19386389.2025.2525720](https://doi.org/10.1080/19386389.2025.2525720). Disponible à l'adresse : <https://doi.org/10.1080/19386389.2025.2525720>

WEIBEL, Stuart L., 1990. Automated Cataloging: Implications for Libraries and Patrons. *Clinic on Library Applications of Data Processing (27th : 1990)* [en ligne]. 1990. [Consulté le 20 juillet 2025]. Disponible à l'adresse : <https://hdl.handle.net/2142/1294>

Indexation matière

BARREAUX, Sabine, DAILLE, Béatrice et JACQUEY, Évelyne, 2017. Indexation automatique en SHS : bilan d'une expérimentation. *I2D - Information, données & documents* [en ligne]. 1 avril 2017. Vol. 54, n° 1, pp. 15-17. [Consulté le 2 janvier 2026]. DOI [10.3917/i2d.171.0015](https://doi.org/10.3917/i2d.171.0015). Disponible à l'adresse : <https://shs.cairn.info/revue-i2d-information-donnees-et-documents-2017-1-page-15>

BUSCH, Dimitri et LANDE, Dmytro, 2025. Semantische Dokumentenindexierung mit generativer KI. *Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen und Bibliothekare* [en ligne]. 17 avril 2025. Vol. 78, n° 1, pp. 34-49. [Consulté le 31 décembre 2025]. DOI [10.31263/voebm.v78i1.9251](https://doi.org/10.31263/voebm.v78i1.9251). Disponible à l'adresse : <https://journals.univie.ac.at/index.php/voebm/article/view/9251>

CAVALIÉ, Étienne, 2019. Les chantiers d'indexation rétrospective à la BNF. In : CAVALIÉ, Étienne (éd.), *L'indexation matière en transition : de la réforme de Rameau à l'indexation automatique*. Paris : Éditions du Cercle de la librairie. pp. 99-112. Collection Bibliothèques. ISBN 978-2-7654-1623-4.

CAVALIÉ, Étienne (éd.), 2019. Paris : Éditions du Cercle de la librairie. Collection Bibliothèques. ISBN 978-2-7654-1623-4.

CHOW, Eric H. C., KAO, T. J. et LI, Xiaoli, 2024. An Experiment with the Use of ChatGPT for LCSH Subject Assignment on Electronic Theses and Dissertations. *Cataloging & Classification Quarterly* [en ligne]. 3 juillet 2024. Vol. 62, n° 5, pp. 574-588. [Consulté le 22 juillet 2025]. DOI [10.1080/01639374.2024.2394516](https://doi.org/10.1080/01639374.2024.2394516). Disponible à l'adresse : <https://doi.org/10.1080/01639374.2024.2394516>

GOLUB, Koraljka, 2021. Automated Subject Indexing: An Overview. *Cataloging & Classification Quarterly* [en ligne]. 29 novembre 2021. Vol. 59, n° 8, pp. 702-719. [Consulté le 22 juillet 2025]. DOI [10.1080/01639374.2021.2012311](https://doi.org/10.1080/01639374.2021.2012311). Disponible à l'adresse : <https://doi.org/10.1080/01639374.2021.2012311>

GOLUB, Koraljka et PESTANA, Olivia, 2023. Subject-Based Knowledge Organisation: An OER for Supporting (Digital) Humanities Research. *Education for Information* [en ligne]. 1 mai 2023. Vol. 39, n° 2, pp. 203-216. [Consulté le 22 juillet 2025]. DOI [10.3233/EFI-230037](https://doi.org/10.3233/EFI-230037). Disponible à l'adresse : <https://journals.sagepub.com/action/showAbstract>

GOLUB, Koraljka, WANG, Jue et WIDEGREN, Johannes, 2024. Using ChatGPT for (Semi-) Automatic Subject Indexing of Different Document Types. In : *Digital Humanities in the Nordic and Baltic Countries 8th Conference Reykjavík 27-31 May 2024* [en ligne]. 2024. [Consulté le 22 juillet 2025]. Disponible à l'adresse : <https://urn.kb.se/resolve?urn=urn:nbn:se:lnu:diva-129484>

HAHN, Jim, 2024a. Bifurcation of Semi-Automated Subject Indexing Services. *Library Resources & Technical Services* [en ligne]. 12 août 2024. Vol. 68, n° 3. [Consulté le 22 juillet 2025]. DOI [10.5860/lrts.68n3.8066](https://doi.org/10.5860/lrts.68n3.8066). Disponible à l'adresse : <https://journals.ala.org/index.php/lrts/article/view/8066>

KAMAL, Ahmad M. et GOLUB, Koraljka, 2025. Subject Matters : Metadata Standards and Subject Access for Library and Museum Catalogues. In : HANSSEN, Jens-Morten et FURUSETH, Sessel (éd.), *The Hermeneutics of Bibliographic Data and Cultural Metadata* [en ligne]. Oslo : National Library of Norway. pp. 204-239. [Consulté le 22 juillet 2025]. Disponible à l'adresse : <https://urn.kb.se/resolve?urn=urn:nbn:se:lnu:diva-136374>

Par bibliothèque et pays

Allemagne – Autriche

Policy Artificial Intelligence, [sans date]. *Österreichische Nationalbibliothek* [en ligne] [Consulté le 30 mai 2025]. Disponible à l'adresse : <https://www.onb.ac.at/en/more/about-us/policy-artificial-intelligence>

DNB

JUNGER, Ulrike, 2018. Automation First. The Subject Cataloguing Policy of the Deutsche Nationalbibliothek. [en ligne]. 19 juillet 2018. [Consulté le 25 janvier 2026]. Disponible à l'adresse : <https://repository.ifla.org/handle/20.500.14598/6361>

JUNGER, Ulrike et SCHOLZE, Frank, 2021. Neue Wege und Qualitäten – Die Inhaltserschließungspolitik der Deutschen Nationalbibliothek. In : FRANKE-MAIER, Michael, KASPRZIK, Anna, LEDL, Andreas, *et al.* (éd.), *Qualität in der Inhaltserschließung* [en ligne]. De Gruyter Saur. pp. 55-70. [Consulté le 22 juillet 2025]. ISBN 978-3-11-069159-7. Disponible à l'adresse : <https://www.degruyterbrill.com/document/doi/10.1515/9783110691597-004/html>

KÄHLER, Maximilian, 2025. *Benchmarking Automatic Indexing Methods on German Scientific Literature*. CENL "AI in Libraries" Network Group webinars. 16 septembre 2025.

KLUGE, Lisa et KÄHLER, Maximilian, 2025. DNB-AI-Project at SemEval-2025 Task 5: An LLM-Ensemble Approach for Automated Subject Indexing. *arXiv* [en ligne]. avril 2025. [Consulté le 1 janvier 2026]. Disponible à l'adresse : <http://arxiv.org/abs/2504.21589>

MERGEL, Ines, KÜHLER, Justus, SCHUMANN, Anna-Lea, *et al.*, 2026. Automated Subject Cataloguing at the German National Library. In : MERGEL, Ines et SCHMIDT, Carsten (éd.), *AI Innovations in Public Services: The Case of National Libraries* [en ligne]. Cham : Springer Nature Switzerland. pp. 25-41. [Consulté le 23 décembre 2025]. ISBN 978-3-032-01344-6. Disponible à l'adresse : https://doi.org/10.1007/978-3-032-01344-6_3

POLEY, Christoph, UHLMANN, Sandro, BUSSE, Frank, *et al.*, 2025. Automatic Subject Cataloguing at the German National Library. *LIBER Quarterly: The Journal of the Association of European Research Libraries* [en ligne]. 8 avril 2025. Vol. 35, n° 1, pp. 1-29. [Consulté le 30 mai 2025]. DOI [10.53377/lq.19422](https://doi.org/10.53377/lq.19422). Disponible à l'adresse : <https://liberquarterly.eu/article/view/19422>

POLEY, Christoph, 2026. *Entretien*. 10 février 2026.

POLEY, Christoph, 2022. Insight into the Machine-Based Subject Cataloguing at the German National Library. [en ligne]. 1 décembre 2022. [Consulté le 25 janvier 2026]. Disponible à l'adresse : https://swib.org/swib22/slides/20221201_poley_dnb_final.pdf

POLEY, Christoph, 2025. *The Usage of Hardware Resources for Automatic Subject Cataloguing at the German National Library – an Analysis and Outlook for Future Challenges*. Fantastic Futures 2025. Londres. 4 décembre 2025.

ZBW

KASPRZIK, Anna, 2023. Automating Subject Indexing at ZBW: Making Research Results Stick in Practice. *LIBER Quarterly: The Journal of the Association of European Research Libraries* [en ligne]. 1 décembre 2023. Vol. 33, n° 1. [Consulté le 30 mai 2025]. DOI [10.53377/lq.13579](https://doi.org/10.53377/lq.13579). Disponible à l'adresse : <https://liberquarterly.eu/article/view/13579>

KASPRZIK, Anna, 2024a. The Automation of Subject Indexing at ZBW and the Role of Metadata in Times of Large Language Models. *Procedia Computer Science*. 2024. Vol. 249, pp. 160-166.

KASPRZIK, Argie, 2025a. *Entretien*. 8 octobre 2025.

KASPRZIK, Argie, 2025b. Die KI(rche) im Dorf lassen – Wege zu einer klima- und sozialverträglichen Nutzung von Künstlicher Intelligenz. [en ligne]. 18 juin 2025. [Consulté le 9 octobre 2025]. Disponible à l'adresse : <https://opus4.kobv.de/opus4-bib-info/frontdoor/index/index/docId/19841>

KASPRZIK, Anna, 2020. Putting Research-based Machine Learning Solutions for Subject Indexing into Practice. In : PASCHKE, Adrian, NEUDECKER, Clemens, REHM, Georg, *et al.* (éd.), *Proceedings of the Conference on Digital Curation Technologies (Qurator 2020)* [en ligne]. Berlin, Germany : CEUR. janvier 2020. [Consulté le 22 juillet 2025]. CEUR Workshop Proceedings. Disponible à l'adresse : https://ceur-ws.org/Vol-2535/#paper_1

KASPRZIK, Anna, 2024b. Transferring Applied Machine Learning Research into Subject Indexing Practice. In : BALNAVES, Edmund, BULTRINI, Leda, COX, Andrew M., *et al.* (éd.), *New Horizons in Artificial Intelligence in Libraries* [en ligne]. De Gruyter. pp. 199-212. [Consulté le 3 juin 2025]. Disponible à l'adresse : <https://library.oapen.org/handle/20.500.12657/98734>

KASPRZIK, Anna, 2024c. *Zwischen Hype und heiligen Kühen – Künstliche Intelligenz in Bibliotheken am Beispiel der Inhaltserschließung*. Künstliche Intelligenz & ihre Auswirkungen auf wissenschaftliche Bibliotheken. 2024.

KEMPF, Andreas Oskar, 2025. Applying Language Models for Knowledge Organization System Maintenance and Development: Approaching Roles and Competencies for Editorial Review Boards. In : *Metadata Standards Matter : Building a Sustainable Future for Information* [en ligne]. Athènes. 2025. Disponible à l'adresse : <https://iflacos.eu/index.php/abstracts/#J>

LIEBETRUTH, Claudia, 2022. Interdisciplinary Teamwork as a Success Factor. Competencies Needed for Adaptation and Implementation of a Subject Indexing Support Software at the Leibniz Information Center for Economics (ZBW). [en ligne]. 9 novembre 2022. [Consulté le 21 janvier 2026]. Disponible à l'adresse : https://www.ifla.org/wp-content/uploads/4.Liebetruth_ZBW_IFLA-webinar.pdf

Chili

DIAZ, Claudio Daniel Henriquez et VERGARA, María Paz Rioseco, 2025. AI-Powered Bibliographic Control: Automating Cataloging, Standardization, and Data Capture. [en ligne]. 15 août 2025. [Consulté le 15 novembre 2025]. Disponible à l'adresse : <https://www.ifla.org/events/artificial-intelligence-bibliographic-control-and-legal-matters-navigating-new-horizons/>

Estonie

ASULA, Marit, MAKKE, Jane, FREIENTHAL, Linda, *et al.*, 2021. Kratt: Developing an Automatic Subject Indexing Tool for the National Library of Estonia. *Cataloging & Classification Quarterly* [en ligne]. 25 octobre 2021. Vol. 59, n° 8, pp. 775-793. [Consulté le 22 juillet 2025].

DOI [10.1080/01639374.2021.1998283](https://doi.org/10.1080/01639374.2021.1998283). Disponible à l'adresse : <https://doi.org/10.1080/01639374.2021.1998283>

SCHMIDT, Carsten, 2026. From Metadata to Meaning: Applying AI in the National Library of Estonia. In : MERGEL, Ines et SCHMIDT, Carsten (éd.), *AI Innovations in Public Services: The Case of National Libraries* [en ligne]. Cham : Springer Nature Switzerland. pp. 121-133.

[Consulté le 23 décembre 2025]. ISBN 978-3-032-01344-6. Disponible à l'adresse : https://doi.org/10.1007/978-3-032-01344-6_10

SINISALU, mas, 2025. KRATT-automated Cataloguing. [en ligne]. 30 janvier 2025.

[Consulté le 30 décembre 2025]. Disponible à l'adresse : https://www.cenl.org/wp-content/uploads/2025/02/Automatic_cataloguing_CENL_AI_30_01_2025.pdf

États-Unis – Royaume-Uni – Australie

Bibliothèque du Congrès

Exploring Computational Description Experiments Work Library of Congress Labs Library of Congress, [sans date]. *Library of Congress, Washington, D.C. 20540 USA* [en ligne]

[Consulté le 20 juin 2025]. Disponible à l'adresse : <https://labs.loc.gov/work/experiments/ECD/?locrlr=blogsig>

BRADOR, Isabel, 2024. Could Artificial Intelligence Help Catalog Thousands of Digital Library Books? An Interview with Abigail Potter and Caroline Saccucci The Signal. *The Library of Congress* [en ligne]. 19 novembre 2024. [Consulté le 22 juin 2025]. Disponible à l'adresse : <https://blogs.loc.gov/thesignal/2024/11/could-artificial-intelligence-help-catalog-thousands-of-digital-library-books-an-interview-with-abigail-potter-and-caroline-saccucci>

MANCHESTER, Eileen J., 2023. Introducing the LC Labs Artificial Intelligence Planning Framework The Signal. *The Library of Congress* [en ligne]. 15 novembre 2023. [Consulté le 30 décembre 2025]. Disponible à l'adresse : <https://blogs.loc.gov/thesignal/2023/11/introducing-the-lc-labs-artificial-intelligence-planning-framework>

POTTER, Abigail et MCGRATTAN, Matthew, 2025. *Computational Description: An End-to-End Case Study*. Fantastic Futures 2025. Londres. 4 décembre 2025.

POTTER, Abigail, 2025. *Entretien*. 8 décembre 2025.

SACCUCCI, Caroline et POTTER, Abigail, 2024. Assessing Machine Learning for Cataloging at the Library of Congress. In : BALNAVES, Edmund, BULTRINI, Leda, COX, Andrew M., *et al.* (éd.), *New Horizons in Artificial Intelligence in Libraries* [en ligne]. De Gruyter. pp. 227-238.

[Consulté le 3 juin 2025]. Disponible à l'adresse : <https://library.oapen.org/handle/20.500.12657/98734>

SACCUCCI, Caroline et POTTER, Abigail, 2025. Results of AI Experimentation for Cataloging at the Library of Congress. [en ligne]. 15 août 2025. [Consulté le 15 novembre 2025]. Disponible à l'adresse : <https://www.ifla.org/events/artificial-intelligence-bibliographic-control-and-legal-matters-navigating-new-horizons/>

SCHUMANN, Anna-Lea, KÜHLER, Justus et MERGEL, Ines, 2026. Exploring Computational Descriptions for Metadata Creation for E-Books at the Library of Congress, United States of America. In : MERGEL, Ines et SCHMIDT, Carsten (éd.), *AI Innovations in Public Services: The Case of National Libraries* [en ligne]. Cham : Springer Nature Switzerland. pp. 163-178. [Consulté le 23 décembre 2025]. ISBN 978-3-032-01344-6. Disponible à l'adresse : https://doi.org/10.1007/978-3-032-01344-6_13

Autres

Inventory of NARA Artificial Intelligence (AI) Use Cases, 2024. *National Archives* [en ligne] [Consulté le 30 mai 2025]. Disponible à l'adresse : <https://www.archives.gov/ai>

BERMAN, Glen, SMITHIES, James, TAIURU, Karaitiana, *et al.*, 2025. *Evaluation Everywhere, But Not All At Once*. Fantastic Futures 2025. Londres. 4 décembre 2025.

BROADWELL, Peter, ENGEL, Claudia, WHITMIRE, Amanda, *et al.*, 2025. *AIInTheCards: Surfacing Environmental and Cultural History from Observational Records via VLMs and a Web-Based Review Interface*. Fantastic Futures 2025. Londres. 4 décembre 2025.

HAHN, Jim, 2024b. Sociotechnical Automation Science: A Case Study in Developing and Augmenting an Ensemble Neural Network with Multiple LLMs for Subject Cataloging at the Penn Libraries. [en ligne]. 26 juin 2024. [Consulté le 22 juillet 2025]. Disponible à l'adresse : <https://repository.upenn.edu/handle/20.500.14332/60308>

JANCO, Andrew, TUBB, Daniel, FARNSWORTH-ALVEAR, Ann, *et al.*, 2025. *Auto-Cataloging Research Materials from the Endangered Archives Programme*. Fantastic Futures 2025. Londres. 4 décembre 2025.

KORPET, Sheldon et REES, Nathalie, 2025. Augmenting Cataloguers: Planning an AI Agent to Generate MARC21 Records. *Catalogue and Index* [en ligne]. 17 juin 2025. Vol. 211, pp. 29-39. [Consulté le 7 janvier 2026]. Disponible à l'adresse : <https://journals.cilip.org.uk/catalogue-and-index/article/view/750>

PCC TASK GROUP ON AI AND MACHINE LEARNING FOR CATALOGING AND METADATA, 2025. Draft PCC Guiding Principles for Use of AI and Machine Learning Technologies in Cataloging and Metadata Work. [en ligne]. 2025. [Consulté le 7 novembre 2025]. Disponible à l'adresse : https://docs.google.com/document/d/1UIDqZGU-1gVfRmWb_-UKRhZRUYsyr1SdrU6qSA2f-80

SWAIN, Eric, 2025. *Unlocking Rare Collections in Trove*. Fantastic Futures 2025. Londres. 4 décembre 2025.

WOLFE, Erin, 2025. *From Microfilm to Metadata: AI-Powered Indexing of 139 Years of a Student Newspaper*. Fantastic Futures 2025. Londres. 4 décembre 2025.

France

BELLIER, Luc, 2017. *Organisation des données, organisation du travail en bibliothèques universitaires à l'heure du Big Data*. Mémoire d'étude. Villeurbanne : Enssib.

BOYER, Léopold, 2025. *IA et collections patrimoniales : usages actuels et potentiels*. Mémoire d'étude. Villeurbanne : Enssib.

DELHAYE, Marc et PERRIN, Sébastien, 2025. Bibliothèque de recherche et IA : vers de nouveaux espaces documentaires ?. *Bulletin des bibliothèques de France* [en ligne]. 2 juillet 2025. [Consulté le 5 septembre 2025]. Disponible à l'adresse : https://bbf.enssib.fr/matieres-a-penser/bibliotheque-de-recherche-et-ia-vers-de-nouveaux-espaces-documentaires_73426

PHILIBERT, Coralie, 2024. *Catégorie B et catalogage en bibliothèque : histoire d'un statut et d'un cœur de métier, état des lieux actuel et perspectives d'avenir*. Mémoire d'étude. Villeurbanne : Enssib.

WENZ, Romain, 2009. *L'avenir des catalogues: formats, données, outils, usages*. Mémoire d'étude. Villeurbanne : Enssib.

Abes

FIL'ABES, 2025. L'indexation RAMEAU assistée par IA : retour sur une expérimentation prometteuse. *Fil'Abes* [en ligne]. 10 avril 2025. [Consulté le 18 juillet 2025]. Disponible à l'adresse : <https://fil.abes.fr/2025/04/10/lindexation-rameau-assistee-par-ia-retour-sur-une-experimentation-prometteuse/>

NICOLAS, Yann, 2025a. *Entretien*. 13 mars 2025.

NICOLAS, Yann, 2025b. *Indexation RAMEAU Assistée Par IA. Expérimentation Réalisée Avec 12 Établissements Testeurs* [en ligne]. Montpellier : Abes. [Consulté le 18 juillet 2025]. Disponible à l'adresse : <https://abes.fr/wp-content/uploads/2025/04/etude-experimentation-indexation-rameau-assistee-par-ia.pdf>

NICOLAS, Yann, ZARAGOZA, Thomas et LE PROVOST, Aline, 2024. From Text to Data Inside Bibliographic Records: Entity Recognition and Linking. In : BALNAVES, Edmund, BULTRINI, Leda, COX, Andrew M., *et al.* (éd.), *New Horizons in Artificial Intelligence in Libraries* [en ligne]. De Gruyter. pp. 239-255. [Consulté le 3 juin 2025]. Disponible à l'adresse : <https://library.oapen.org/handle/20.500.12657/98734>

BNF

BERMÈS, Emmanuelle, 2021a. *Intelligence Artificielle. Feuille de Route 2021-2026*. Paris : Bibliothèque Nationale de France.

BERMÈS, Emmanuelle, 2021b. *Intelligence Artificielle. Feuille de Route 2021-2026. Extrait*. Paris : Bibliothèque Nationale de France.

HENRY, Antoine, 2026. How to Make Millions of Documents Accessible Through AI and a Digital Library? The Case of Gallica, French National Library. In : MERGEL, Ines et SCHMIDT, Carsten (éd.), *AI Innovations in Public Services: The Case of National Libraries* [en ligne]. Cham : Springer Nature Switzerland. pp. 109-120. [Consulté le 23 décembre 2025]. ISBN 978-3-032-01344-6. Disponible à l'adresse : https://doi.org/10.1007/978-3-032-01344-6_9

ILLIEN, Gildas, 2015. Une BnF agile ? Quand le développement logiciel fait bouger l'organisation du travail. In : PÉRALES, Christophe (éd.), *Conduire le changement en bibliothèque : Vers des organisations apprenantes* [en ligne]. Villeurbanne : Presses de l'enssib. pp. 102-120. La Boîte à outils. [Consulté le 2 janvier 2026]. ISBN 978-2-37546-067-2. Disponible à l'adresse : <https://books.openedition.org/pressesenssib/3743>

KOSKAS, Mathilde, 2025. *Entretien*. 6 novembre 2025.

LABORDERIE, Arnaud, 2025. *Entretien*. 26 novembre 2025.

MOREUX, Jean-Philippe, 2025a. *Entretien*. 25 septembre 2025.

MOREUX, Jean-Philippe, 2025b. *Entretien*. 30 octobre 2025.

MOREUX, Jean-Philippe, 2025c. *Entretien*. 9 décembre 2025.

MOREUX, Jean-Philippe, 2019. Indexation automatique des documents : de nouvelles métadonnées pour de nouveaux services. In : CAVALIÉ, Étienne (éd.), *L'indexation matière en transition : de la réforme de Rameau à l'indexation automatique*. Paris : Éditions du Cercle de la librairie. pp. 113-133. Collection Bibliothèques. ISBN 978-2-7654-1623-4.

MOREUX, Jean-Philippe, [sans date]. *Le Programme de Projets IA de La BnF*.

ROCHE, Mélanie et KOSKAS, Mathilde, 2025. Artificial Intelligence vs. Smart Cataloguing: A Prospective Dialogue. [en ligne]. 19 août 2025. [Consulté le 24 décembre 2025]. Disponible à l'adresse : <https://www.ifla.org/news/cos-at-wlic-2025-sessions-about-ifla-standards/>

ROCHE, Mélanie, 2025. *Entretien*. 26 novembre 2025.

TULEU, Benoît, 2025. *Entretien*. 20 octobre 2025.

Université Paris-Saclay

BELLIER, Luc et MINQUILAN, Dominique, 2025. *Entretien*. 27 juin 2025.

BRETEL, Henri et BELLIER, Luc, 2021. La Bibliométrie à l'université Paris-Saclay : vers une symbiose avec la science ouverte. *Arabesques* [en ligne]. 1 octobre 2021. N° 103, pp. 6-7.

[Consulté le 21 juillet 2025]. DOI [10.35562/arabesques.2723](https://doi.org/10.35562/arabesques.2723). Disponible à l'adresse : <https://publications-prairial.fr/arabesques/index.php?id=2723>

RUSSEIL-SALVAN, Morgane, 2024. *Intégration de l'intelligence Artificielle Aux Services de La DiBISO. Étude de Faisabilité et d'impact*. Gif-sur-Yvette : Université Paris-Saclay.

VIVIER, Emmanuelle, 2024. "BiblioLabs" Application permettant d'agréger, de comparer et d'analyser les données bibliographiques et bibliométriques concernant la littérature scientifique et académique des structures de recherche, Luc Bellier (Université Paris-Saclay). *CSIESR* [en ligne]. 17 janvier 2024. [Consulté le 7 juillet 2025]. Disponible à l'adresse : <https://www.csiesr.eu/bibliolabs-application-permettant-dagreger-de-comparer-et-danalyser-les-donnees-bibliographiques-et-bibliometriques-concernant-la-litterature-scientifique-et-academique-des-structures-de-rech>

Israel

COHEN, Ahava, 2025. *Entretien*. 26 décembre 2025.

Luxembourg – Suisse – Belgique

IA dans les bibliothèques, [sans date]. *Bibliosuisse* [en ligne] [Consulté le 30 mai 2025]. Disponible à l'adresse : <https://www.bibliosuisse.ch/fr/apropos/commissions/iadanslesbibliotheques>

LOWAGIE, Hannes, 2025a. *AI-powered Cataloguing a Practical Guide to Building a Cataloguing Application with Power Apps*. London : Facet Publishing. ISBN 978-1-78330-807-1.

LOWAGIE, Hannes, 2025b. *Entretien*. 10 octobre 2025.

LOWAGIE, Hannes, 2024. Harnessing PowerApps and AI for Automated Cataloguing: Innovations in Bibliographic Record Creation. *Catalogue and Index* [en ligne]. 16 décembre 2024. Vol. 209, pp. 3-9. [Consulté le 29 juillet 2025]. Disponible à l'adresse : <https://journals.cilip.org.uk/catalogue-and-index/article/view/697>

LOWAGIE, Hannes, 2025c. From PDF to Prompt. Toward Universal Bibliographic Control Through Machine-Readable Cataloging Rules. [en ligne]. 15 août 2025. [Consulté le 15 novembre 2025]. Disponible à l'adresse : <https://www.ifla.org/events/artificial-intelligence-bibliographic-control-and-legal-matters-navigating-new-horizons/>

LOWAGIE, Hannes et BERCEGEAY, Marie-Sophie, 2024. Rétro-Catalogage à La KBR. [en ligne]. 2024. [Consulté le 29 janvier 2026]. Disponible à l'adresse : https://halpartenaires2024.sciencesconf.org/data/pages/B_1_2024_HAL_AI_KBR.pdf

LOWAGIE, Hannes et WOENSEL, Julie Van, 2024. Simplifying Subject Indexing: A Python-Powered Approach in KBR, the National Library of Belgium. *The Code4Lib Journal* [en ligne]. 7 octobre 2024. N° 59. [Consulté le 22 juin 2025]. Disponible à l'adresse : <https://journal.code4lib.org/articles/18103>

MAURER, Yves, 2025. *Autocat Cataloguing Assistant*. Fantastic Futures 2025. Londres. 4 décembre 2025.

SCHNEIDER, Pit, MAURER, Yves et MARSCHALL, Ralph, 2023. Nautilus: An End-To-End METS/ALTO OCR Enhancement Pipeline. *LIBER Quarterly: The Journal of the Association of European Research Libraries* [en ligne]. 26 avril 2023. Vol. 33, n° 1, pp. 1-19. [Consulté le 30 mai 2025]. DOI [10.53377/lq.13330](https://doi.org/10.53377/lq.13330). Disponible à l'adresse : <https://liberquarterly.eu/article/view/13330>

SCHUMANN, Anna-Lea et MERGEL, Ines, 2026. The Retrocatalography Project: Combining AI with the Microsoft Power Platform at the Royal National Library of Belgium. In : MERGEL, Ines et SCHMIDT, Carsten (éd.), *AI Innovations in Public Services: The Case of National Libraries* [en ligne]. Cham : Springer Nature Switzerland. pp. 83-93. [Consulté le 23 décembre 2025]. ISBN 978-3-032-01344-6. Disponible à l'adresse : https://doi.org/10.1007/978-3-032-01344-6_7

Pays-Bas

Automatic Metadata Generation at the KB, National Library of the Netherlands, [sans date]. [en ligne] [Consulté le 20 juillet 2025]. Disponible à l'adresse : <https://www.kb.nl/en/over-ons/projecten/automatisch-metadateren>

KB Onderzoekskroniek - IP2024-05, [sans date]. *IP vakblad voor informatieprofessionals* [en ligne] [Consulté le 21 juillet 2025]. Disponible à l'adresse : <https://informatieprofessional.maglr.com/kb-ondzoekskroniek12>

BUESINK, Marie, 2025a. *Automated Metadata Creation in Practice: The KB's 'Retro Tool'*. LIBER. Lausanne. 3 juillet 2025.

BUESINK, Marie, 2025b. *Entretien*. 8 septembre 2025.

CELLUCCI, Vincent et VAN DE GEER, Jaap, 2025. *Starting AI at Your Institution From Scratch*. Fantastic Futures 2025. Londres. 4 décembre 2025.

KLEPPE, Martijn, VELDHOF, Sara, WAAL-GENTENAAR, Meta van der, *et al.*, 2019. *Exploration possibilities Automated Generation of Metadata* [en ligne]. Zenodo. [Consulté le 22 juillet 2025]. Disponible à l'adresse : <https://zenodo.org/records/3375192>

VAN DER VAART, Barbara et SCHMIDT, Carsten, 2026. AI for Access: Strategic Innovation at the Dutch National Library. In : MERGEL, Ines et SCHMIDT, Carsten (éd.), *AI Innovations in Public Services: The Case of National Libraries* [en ligne]. Cham : Springer Nature Switzerland. pp. 95-107. [Consulté le 23 décembre 2025]. ISBN 978-3-032-01344-6. Disponible à l'adresse : https://doi.org/10.1007/978-3-032-01344-6_8

Scandinavie

Finlande

Semi-Automatic Subject Cataloging Using Finto AI, [sans date]. [en ligne]
[Consulté le 7 juillet 2025]. Disponible à l'adresse : <https://www.kiwi.fi/spaces/Finto/pages/356352444/Semi-automatic+subject+cataloging+using+Finto+AI>

INKINEN, Juho, LEHTINEN, Mona et SUOMINEN, Osmo, 2025. *Annif Users Survey. Understanding Usage and Challenges*. Helsinki, Finlande : National Library of Finland.

LEHTINEN, Mona, 2025a. Annif, the Automated Subject Indexing Tool: Introduction and Future Development. [en ligne]. 30 janvier 2025. [Consulté le 30 mai 2025]. Disponible à l'adresse : <https://www.cenl.org/automatic-cataloguing-webinar-by-ai-in-libraries-network-group/>

LEHTINEN, Mona, 2025b. *Entretien*. 30 juin 2025.

SCHUMANN, Anna-Lea et MERGEL, Ines, 2026. AI-Supported Automated Subject Indexing and Metadata Management with Annif and Finto AI at the National Library of Finland. In : MERGEL, Ines et SCHMIDT, Carsten (éd.), *AI Innovations in Public Services: The Case of National Libraries* [en ligne]. Cham : Springer Nature Switzerland. pp. 43-57. [Consulté le 23 décembre 2025]. ISBN 978-3-032-01344-6. Disponible à l'adresse : https://doi.org/10.1007/978-3-032-01344-6_4

SUOMINEN, Osmo, INKINEN, Juho et LEHTINEN, Mona, 2022. Annif and Finto AI: Developing and Implementing Automated Subject Indexing. *JLIS.it* [en ligne]. 13 janvier 2022. Vol. 13, n° 1, pp. 265-282. [Consulté le 30 mai 2025]. DOI [10.4403/jlis.it-12740](https://doi.org/10.4403/jlis.it-12740). Disponible à l'adresse : <https://www.jlis.it/index.php/jlis/article/view/437>

Norvège

BEAUGUITTE, Pierre et SUOMINEN, Osmo, 2024. Automating Metadata Extraction and Cataloguing: Experiences from the National Libraries of Norway and Finland - SWIB24. [en ligne]. 20 août 2024. [Consulté le 25 juin 2025]. Disponible à l'adresse : <https://forum.swib.org/t/automating-metadata-extraction-and-cataloguing-experiences-from-the-national-libraries-of-norway-and-finland/121>

DE LA ROSA, Javier, 2025. Machine Learning at the National Library of Norway. In : JAILLANT, Lise, WARWICK, Claire, GOODING, Paul, *et al.* (éd.), *Navigating Artificial Intelligence for Cultural Heritage Organisations* [en ligne]. UCL Press. pp. 61-90. [Consulté le 29 décembre 2025]. ISBN 978-1-80008-835-1. Disponible à l'adresse : <https://www.jstor.org/stable/jj.24215718.10>

Artificial Intelligence in Libraries: Case Study of National Library of Norway, 2021. [en ligne]
[Consulté le 21 juillet 2025]. Disponible à l'adresse : <https://www.youtube.com/watch?v=-W8Z1dZO1t8>

Suède

MALMSTEN, Martin, LUNDBORG, Viktoria, FANO, Elena, *et al.*, 2025. Without Heading? : Automatic Creation of a Linked Subject System. In : BALNAVES, Edmund, BULTRINI, Leda, COX, Andrew M., *et al.* (éd.), *De Gruyter* [en ligne]. De Gruyter. pp. 179-198. [Consulté le 20 juin 2025]. Disponible à l'adresse : <https://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-540448>

MALMSTEN, Martin et LUNDBORG, Viktoria, 2025. *Ordo - Agentic AI for Library Workflows*. Fantastic Futures 2025. Londres. 4 décembre 2025.

TOFTGÅRD, Agnes et RENDE, Emma, 2025. *Old System, New Tricks. Using AI to Improve AV Metadata in a Legacy System*. Fantastic Futures 2025. Londres. 4 décembre 2025.

Singapour

JAILANI, Haliza, 2025. Catalysing Change: An AI Roadmap for Cataloguers at the National Library Board of Singapore (NLB). [en ligne]. 15 août 2025. [Consulté le 15 novembre 2025]. Disponible à l'adresse : <https://wlic2025.astanait.edu.kz/>

LING, Ng Hui et GOH, Jeremy, 2025. SGCAT: Using AI to Facilitate Cataloguing. *International Federation of Library Associations, Institutions (IFLA)* [en ligne]. 15 août 2025. [Consulté le 15 novembre 2025]. Disponible à l'adresse : <https://www.ifla.org/events/artificial-intelligence-bibliographic-control-and-legal-matters-navigating-new-horizons/>

NG, Hui Ling, 2026. *Entretien*. 27 janvier 2026.

Par organisation professionnelle ou entreprise

ADBU

ADBU, 2025. *Synthèse Sur Les Expérimentations IA Dans Les SCD*. mars 2025.

ADBU, COMMISSION SDSI, 2025. État des lieux de l'intelligence artificielle dans l'offre logicielle informatique documentaire – ADBU. *Association française des directeurs et personnels de direction des bibliothèques universitaires et de la documentation* [en ligne]. 1 octobre 2025. [Consulté le 29 décembre 2025]. Disponible à l'adresse : <https://adbu.fr/actualites/etat-des-lieux-de-lintelligence-artificielle-dans-loffre-logicielle-informatique-documentaire>

ARL – CNI

Association of Research Libraries Releases Guiding Principles for Artificial Intelligence, [sans date]. *Association of Research Libraries* [en ligne] [Consulté le 21 novembre 2025]. Disponible à l'adresse : <https://www.arl.org/news/association-of-research-libraries-releases-guiding-principles-for-artificial-intelligence/>

ARL et CNI, 2024a. *ARL/CNI AI Scenarios: AI-Influenced Futures* [en ligne]. Washington, DC : Association of Research Libraries, Coalition for Networked Information – Stratus Inc. [Consulté le 21 novembre 2025]. Disponible à l'adresse : <https://www.arl.org/resources/the-arl-cni-2035-scenarios-ai-influenced-futures-in-the-research-environment/>

ARL et CNI, 2024b. *ARL/CNI AI Scenarios: Strategic Context* [en ligne]. Washington, DC : Association of Research Libraries, Coalition for Networked Information – Stratus Inc. [Consulté le 21 novembre 2025]. Disponible à l'adresse : <https://www.arl.org/resources/arlcni-ai-scenarios-strategic-context/>

WEBSTER, Keith, 2025. *Futurescape Libraries AI Toolkit* [en ligne]. Washington, DC : Association of Research Libraries. Disponible à l'adresse : <https://www.arl.org/futurescape-libraries-toolkit/>

Ex Libris

AI in Academic Libraries : Trusted Solutions by Ex Libris, 2024. [en ligne] [Consulté le 29 décembre 2025]. Disponible à l'adresse : <https://exlibrisgroup.com/ai-we-can-use-and-trust/>

AI Metadata Enrichment for Libraries, 2024. *Ex Libris Knowledge Center* [en ligne] [Consulté le 29 décembre 2025]. Disponible à l'adresse : https://knowledge.exlibrisgroup.com/Alma/Product_Materials/010Roadmap/AI_Metadata_Enrichment_for_Libraries

LILL.DAIE, 2025. Library Open Workflows: Working with AI Tools Agents. *Ex Libris Developer Network* [en ligne]. 4 août 2025. [Consulté le 23 décembre 2025]. Disponible à l'adresse : <https://developers.exlibrisgroup.com/blog/library-open-workflows-working-with-ai-tools-agents/>

VELTZMAN, Itai et ELSTEIN, Rael, 2024. How Ex Libris Uses Artificial Intelligence and Smart Services to Transform Libraries. In : BALNAVES, Edmund, BULTRINI, Leda, COX, Andrew M., *et al.* (éd.), *New Horizons in Artificial Intelligence in Libraries* [en ligne]. De Gruyter. pp. 323-335. [Consulté le 3 juin 2025]. Disponible à l'adresse : <https://library.oapen.org/handle/20.500.12657/98734>

IFLA

Exploratory Meeting : Formation of an IFLA Special Interest Group on Artificial Intelligence, 2021. *IFLA* [en ligne] [Consulté le 18 juillet 2025]. Disponible à l'adresse : <https://www.ifla.org/news/exploratory-meeting-formation-of-an-ifla-special-interest-group-on-artificial-intelligence/>

BEHRENS, Renate, 2025. Retour sur le Symposium des Métadonnées de l'IFLA. *Arabesques* [en ligne]. 5 mai 2025. N° 117, pp. 11. [Consulté le 27 juin 2025]. DOI [10.35562/arabesques.4362](https://doi.org/10.35562/arabesques.4362). Disponible à l'adresse : <https://publications-prairial.fr/arabesques/index.php?id=4362>

COX, Andrew M. et DE BRASDEFER, Maria, 2025. IFLA AI Entry Point for Libraries and AI. *International Federation of Library Associations and Institutions (IFLA)* [en ligne]. 5 juin 2025. [Consulté le 9 septembre 2025]. Disponible à l'adresse : <https://repository.ifla.org/handle/20.500.14598/4034>

COX, Andrew, 2024b. The Implications for Use of Artificial Intelligence in Libraries and Education: An Introductory Overview. In : BALNAVES, Edmund, BULTRINI, Leda, COX, Andrew M., *et al.* (éd.), *New Horizons in Artificial Intelligence in Libraries* [en ligne]. De Gruyter. pp. 61-70. [Consulté le 3 juin 2025]. Disponible à l'adresse : <https://library.oapen.org/handle/20.500.12657/98734>

IFLA BIBLIOGRAPHY SECTION, 2025. Statement on Universal Bibliographic Control. *International Federation of Library Associations, Institutions (IFLA)* [en ligne]. juillet 2025. [Consulté le 21 novembre 2025]. Disponible à l'adresse : <https://repository.ifla.org/handle/20.500.14598/4259>

OCLC

Implementing AI to Further Scale and Accelerate WorldCat De-Duplication, 2025. *OCLC* [en ligne] [Consulté le 30 mai 2025]. Disponible à l'adresse : <https://www.oclc.org/en/news/announcements/2025/ai-worldcat-deduplication.html>

WorldShare Record Manager Release Notes, 2025. *OCLC Support* [en ligne] [Consulté le 29 décembre 2025]. Disponible à l'adresse : https://help.oclc.org/Metadata_Services/WorldShare_Record_Manager/Release_notes_and_known_issues/2025_release_notes/055WorldShare_Record_Manager_release_notes_September_2025

BRYANT, Rebecca, 2025. Striking the Right Balance: Opportunities and Challenges of AI in Metadata Workflows. *Hanging Together* [en ligne]. 20 octobre 2025. [Consulté le 31 décembre 2025]. Disponible à l'adresse : <https://hangingtogether.org/striking-the-right-balance-opportunities-and-challenges-of-ai-in-metadata-workflows/>

PROFFITT, Merrilee, 2025. Artificial Intelligence to Support Metadata Workflows: An OCLC RLP Working Group. *Hanging Together* [en ligne]. 14 mai 2025. [Consulté le 30 mai 2025]. Disponible à l'adresse : <https://hangingtogether.org/artificial-intelligence-to-support-metadata-workflows-an-oclc-rlp-working-group/>

PROFFITT, Merrilee, 2025. Backlogs and beyond: AI in Primary Cataloging Workflows. *Hanging Together* [en ligne]. 9 octobre 2025. [Consulté le 31 décembre 2025]. Disponible à l'adresse : <https://hangingtogether.org/backlogs-and-beyond-ai-in-primary-cataloging-workflows/>

RAJAPATIRANA, Bemal, 2025. Scaling De-Duplication in WorldCat: Balancing AI Innovation with Cataloging Care. *Hanging Together* [en ligne]. 4 juin 2025. [Consulté le 22 juin 2025]. Disponible à l'adresse : <https://hangingtogether.org/scaling-de-duplication-in-worldcat-balancing-ai-innovation-with-cataloging-care/>

Autres

AI in Libraries Network Group – The Conference of European National Librarians, [sans date]. [en ligne] [Consulté le 30 décembre 2025]. Disponible à l'adresse : <https://www.cenl.org/networkgroups/ai-in-libraries-network-group/>

LIBER Launches a Taskforce on Artificial Intelligence, 2025. *LIBER Europe* [en ligne] [Consulté le 18 juillet 2025]. Disponible à l'adresse : <https://libereurope.eu/article/liber-launches-a-taskforce-on-artificial-intelligence/>

COHEN, Ahava, ALIVERTI, Christian, BARBUS, Colleen, *et al.*, 2025. *Considering AI in the Growth of RDA* [en ligne]. RDA Steering Committee. Disponible à l'adresse : <https://hdl.handle.net/11213/22898>

JOHNSON, Pete, 2023. *Strategy 2023-2027* [en ligne]. LIBER. [Consulté le 14 janvier 2026]. Disponible à l'adresse : https://libereurope.eu/wp-content/uploads/2022/01/LIBER_STRAT_A5_digital-final-1.pdf

PCC TASK GROUP ON AI AND MACHINE LEARNING FOR CATALOGING AND METADATA, 2025. Draft PCC Guiding Principles for Use of AI and Machine Learning Technologies in Cataloging and Metadata Work. [en ligne]. 2025. [Consulté le 7 novembre 2025]. Disponible à l'adresse : https://docs.google.com/document/d/1UIDqZGU-1gVfRmWb_-UKRhZRUySyr1SdrU6qSA2f-80

PCC TASK GROUP ON AI AND MACHINE LEARNING FOR CATALOGING AND METADATA, 2025. *Final Report*. Washington, DC : Program for Cooperative Cataloging.

RAG en bibliothèque

AI-as-Infrastructure/Aiinfra-Atlas, 2025. [en ligne] [Consulté le 25 décembre 2025]. Disponible à l'adresse : <https://github.com/AI-as-Infrastructure/aiinfra-atlas>

Program, [sans date]. [en ligne] [Consulté le 15 novembre 2025]. Disponible à l'adresse : <http://wlic2025.astanait.edu.kz/program>

What Are Hierarchical Embeddings?, [sans date]. [en ligne] [Consulté le 7 novembre 2025]. Disponible à l'adresse : <https://milvus.io/ai-quick-reference/what-are-hierarchical-embeddings>

APPLEBY, Michael et FITZGERALD, Chelsea, 2025. *Connecting AI Models & Users to Yale Library Data*. Fantastic Futures 2025. Londres. 4 décembre 2025.

BEVARA, Ravi Varma Kumar, LUND, Brady D., MANNURU, Nishith Reddy, *et al.*, 2025. Prospects of Retrieval Augmented Generation (RAG) for Academic Library Search and Retrieval. *Information Technology and Libraries* [en ligne]. 16 juin 2025. Vol. 44, n° 2. [Consulté le 4 février 2026]. DOI [10.5860/ital.v44i2.17361](https://doi.org/10.5860/ital.v44i2.17361). Disponible à l'adresse : <https://ital.corejournals.org/index.php/ital/article/view/17361>

DESCHENES, Any, SNYDMAN, Stu et KILSDONK, Angela, 2025. *Collections Explorer: Building a User-Centered, AI-Driven Discovery Platform at Harvard Library*. Fantastic Futures 2025. Londres. 4 décembre 2025.

EVENSTEIN SIGALOV, Shani, 2025. *Building Bridges, Not Black Boxes: Integrating Generative AI and Linked Open Data for Cultural Heritage Futures*. Fantastic Futures 2025. Londres. 4 décembre 2025.

GALBREATH, Blake L., ENGLAND, Erica, JOHNSON, Corey M., *et al.*, 2025. AI-Infused Discovery Environments: Information Retrieval Boon or Overpromised Hype?. *Information Technology and Libraries* [en ligne]. 15 décembre 2025. Vol. 44, n° 4. [Consulté le 4 février 2026]. DOI [10.5860/ital.v44i4.17465](https://doi.org/10.5860/ital.v44i4.17465). Disponible à l'adresse : <https://ital.corejournals.org/index.php/ital/article/view/17465>

GORGA, Angelo La, TESTA, Roberto et VERNA, Lorenzo, 2025. LLM e Retrieval Augmented Generation (RAG) per le biblioteche. *DigitCult - Scientific Journal on Digital Cultures* [en ligne]. 5 août 2025. Vol. 10, n° 1, pp. 59-73. [Consulté le 30 janvier 2026]. DOI [10.36158/97912566920714](https://doi.org/10.36158/97912566920714). Disponible à l'adresse : <http://www.digitcult.it/index.php/dc/article/view/292>

LUND, Brady, 2025. Prospects of Retrieval Augmented Generation (RAG) for Academic Library Search and Retrieval. *Social Science Research Network* [en ligne]. juin 2025. [Consulté le 15 novembre 2025]. Disponible à l'adresse : <https://papers.ssrn.com/abstract=5295044>

MARTINEZ, Alberto, 2025. *Challenges of Building a RAG System for Library Reference*. Fantastic Futures 2025. Londres. 4 décembre 2025.

Sciences humaines et sociales

Philosophie

FISHER, Robert B., 2025. What Do We Want from AI?. *AI Magazine* [en ligne]. 2025. Vol. 46, n° 4, pp. e70042. [Consulté le 13 novembre 2025]. DOI [10.1002/aaai.70042](https://doi.org/10.1002/aaai.70042). Disponible à l'adresse : <https://onlinelibrary.wiley.com/doi/abs/10.1002/aaai.70042>

FLORIDI, Luciano, 2023. *L'éthique de l'intelligence Artificielle. Principes, Défis et Opportunités*. Mimésis. Philosophie.

Économie et sociologie

Pour un dialogue social au service des bons usages de l'IA et d'une nouvelle étape de progrès social dans les entreprises et les administrations, 2025. *Dialia* [en ligne] [Consulté le 20 janvier 2026]. Disponible à l'adresse : <https://dialia.alwaysdata.net/>

ACEMOGLU, Daron, AUTOR, David et JOHNSON, Simon, 2023. *Can We Have Pro-Worker AI?*. MIT Shaping the future of work initiative.

ACEMOGLU, Daron et JOHNSON, Simon, 2024. Learning From Ricardo and Thompson: Machinery and Labor in the Early Industrial Revolution and in the Age of Artificial Intelligence. *Annual Review of Economics* [en ligne]. 22 août 2024. Vol. 16, n° Volume16, 2024, pp. 597-621. [Consulté le 16 juillet 2025]. DOI [10.1146/annurev-economics-091823-025129](https://doi.org/10.1146/annurev-economics-091823-025129). Disponible à l'adresse : <https://www.annualreviews.org/content/journals/10.1146/annurev-economics-091823-025129>

- ACEMOGLU, Daron et JOHNSON, Simon, 2023. Rebalancing AI. *Finance & Development* [en ligne]. 30 novembre 2023. Vol. 60, n° 4, pp. 26-29. [Consulté le 16 juillet 2025]. DOI [10.5089/9798400260179.001.A007](https://www.elibrary.imf.org/view/journals/022/0060/004/article-A007-en.xml). Disponible à l'adresse : <https://www.elibrary.imf.org/view/journals/022/0060/004/article-A007-en.xml>
- AUTOR, David, 2024. Applying AI to Rebuild Middle Class Jobs. *National Bureau of Economic Research* [en ligne]. février 2024. [Consulté le 14 juillet 2025]. Disponible à l'adresse : <https://www.nber.org/papers/w32140>
- AUTOR, David et THOMPSON, Neil, 2025. Expertise. *National Bureau of Economic Research* [en ligne]. juin 2025. [Consulté le 14 juillet 2025]. Disponible à l'adresse : <https://www.nber.org/papers/w33941>
- AUTOR, David, MINDELL, David A. et REYNOLDS, Elisabeth B., 2022. *The Work of the Future: Building Better Jobs in an Age of Intelligent Machines* [en ligne]. The MIT Press. [Consulté le 14 juillet 2025]. ISBN 978-0-262-36775-2. Disponible à l'adresse : <https://direct.mit.edu/books/book/5246/The-Work-of-the-FutureBuilding-Better-Jobs-in-an>
- BERG, Andrew, PAPAGEORGIOU, Chris et VAZIRI, Maryam, 2023. Technology's Bifurcated Bite. *Finance & Development* [en ligne]. 30 novembre 2023. Vol. 60, n° 4, pp. 34-37. [Consulté le 16 juillet 2025]. DOI [10.5089/9798400260179.001.A009](https://www.elibrary.imf.org/view/journals/022/0060/004/article-A009-en.xml). Disponible à l'adresse : <https://www.elibrary.imf.org/view/journals/022/0060/004/article-A009-en.xml>
- BOUILLOUD, Jean-Philippe, 2023. *Pouvoir faire un beau travail. Une revendication professionnelle*. Toulouse : Érès. Sociologie clinique. ISBN 978-2-7492-7667-0.
- BROWN, Sara, 2019. The Lure of 'so-so Technology,' and How to Avoid It. *MIT Management Sloan School* [en ligne]. 31 octobre 2019. [Consulté le 19 juillet 2025]. Disponible à l'adresse : <https://mitsloan.mit.edu/ideas-made-to-matter/lure-so-so-technology-and-how-to-avoid-it>
- BRYNJOLFSSON, Erik et UNGER, Gabriel, 2023a. The Macroeconomic Impact of Artificial Intelligence: Productivity, Inequality, and Concentration. *Finance & Development*. décembre 2023.
- BRYNJOLFSSON, Erik et UNGER, Gabriel, 2023b. The Macroeconomics Of Artificial Intelligence. *Finance & Development* [en ligne]. 30 novembre 2023. Vol. 60, n° 4, pp. 20-25. [Consulté le 16 juillet 2025]. DOI [10.5089/9798400260179.001.A006](https://www.elibrary.imf.org/view/journals/022/0060/004/article-A006-en.xml). Disponible à l'adresse : <https://www.elibrary.imf.org/view/journals/022/0060/004/article-A006-en.xml>
- BULLOCK, Justin B., CHEN, Yu-Che, HIMMELREICH, Johannes, *et al.* (éd.), 2024. Oxford University Press. [Consulté le 17 juillet 2025]. ISBN 978-0-19-757932-9. Disponible à l'adresse : <https://doi.org/10.1093/oxfordhb/9780197579329.001.0001>
- CAPRARO, Valerio, LENTSCH, Austin, ACEMOGLU, Daron, *et al.*, 2024. The Impact of Generative Artificial Intelligence on Socioeconomic Inequalities and Policy Making. *PNAS Nexus* [en ligne]. 1 juin 2024. Vol. 3, n° 6. [Consulté le 16 juillet 2025]. DOI [10.1093/pnasnexus/pgae191](https://doi.org/10.1093/pnasnexus/pgae191). Disponible à l'adresse : <https://doi.org/10.1093/pnasnexus/pgae191>
- CASILLI, Antonio A., 2024. Il Paradosso Dell'IA: Il Lavoro Nascosto per Creare Un Futuro Senza Lavoro. *Micromega* [en ligne]. novembre 2024. N° 6. [Consulté le 26 juin 2025]. Disponible à l'adresse : <https://hal.science/hal-04747044>
- COURTIER-ORGOGOZO, Virginie et DEVILLERS, Laurence, 2024. La société face aux avancées des sciences et des techniques : Le cas de l'intelligence artificielle et de la génétique. *Futuribles* [en ligne]. 4 janvier 2024. Vol. 458, n° 1, pp. 25-44. [Consulté le 14 juillet 2025]. DOI [10.3917/futur.458.0025](https://doi.org/10.3917/futur.458.0025). Disponible à l'adresse : <https://shs-cairn-info.docelec.enssib.fr/revue-futuribles-2024-1-page-25>

DEVILLERS, Laurence, 2024. Le langage non responsable des systèmes d'intelligence artificielle (IA) générative. *Champ lacanien* [en ligne]. 15 octobre 2024. Vol. 28, n° 1, pp. 133-138. [Consulté le 14 juillet 2025]. DOI [10.3917/chla.028.0133](https://doi.org/10.3917/chla.028.0133). Disponible à l'adresse : <https://shs-cairn-info.docelec.enssib.fr/revue-champ-lacanien-2024-1-page-133>

GAUTIE, Jerome et PEREZ, Coralie, 2025. Pour Une IA En Faveur Des Travailleurs et Des Travailleuses. In : ERHEL, Christine et PALIER, Bruno (éd.), *Travailler Mieux*. Presses Universitaires de France – Humensis. pp. 97-118. ISBN 978-2-13-087544-4.

GOPINATH, Gita, 2023. Straight Talk: Harnessing AI for Global Good. *Finance & Development* [en ligne]. 30 novembre 2023. Vol. 60, n° 4. [Consulté le 16 juillet 2025]. DOI [10.5089/9798400260179.001.A018](https://doi.org/10.5089/9798400260179.001.A018). Disponible à l'adresse : <https://www.elibrary.imf.org/view/journals/022/0060/004/article-A018-en.xml>

GRINBAUM, Alexei, CHATILA, Raja, DEVILLERS, Laurence, *et al.*, 2023. *Systèmes d'intelligence artificielle générative : enjeux d'éthique* [en ligne]. report. Comité national pilote d'éthique du numérique. [Consulté le 14 juillet 2025]. Disponible à l'adresse : <https://cea.hal.science/cea-04153216>

KORINEK, Anton, 2023. Scenario Planning for an A(G)I Future. *Finance & Development* [en ligne]. 30 novembre 2023. Vol. 60, n° 4, pp. 30-33. [Consulté le 16 juillet 2025]. DOI [10.5089/9798400260179.001.A008](https://doi.org/10.5089/9798400260179.001.A008). Disponible à l'adresse : <https://www.elibrary.imf.org/view/journals/022/0060/004/article-A008-en.xml>

TAVARES, Marina Mendes et JARAVEL, Xavier, 2025. A Place for Human Talent in The AI Age. *Finance & Development* [en ligne]. 3 mars 2025. Vol. 62, n° 1. [Consulté le 16 juillet 2025]. DOI [10.5089/9798400297144.022.A008](https://doi.org/10.5089/9798400297144.022.A008). Disponible à l'adresse : <https://www.elibrary.imf.org/view/journals/022/0062/001/article-A008-en.xml>

Collaboration homme-machine

ANTHONY, Callen, BECHKY, Beth A. et FAYARD, Anne-Laure, 2023. "Collaborating" with AI: Taking a System View to Explore the Future of Work. *Organization Science* [en ligne]. 9 janvier 2023. [Consulté le 19 juillet 2025]. DOI [10.1287/orsc.2022.1651](https://doi.org/10.1287/orsc.2022.1651). Disponible à l'adresse : <https://pubsonline.informs.org/doi/abs/10.1287/orsc.2022.1651>

ARMSTRONG, Ben, CHEN, Valerie K., CUELLAR, Alex, *et al.*, 2024. Automation from the Worker's Perspective. *arXiv* [en ligne]. septembre 2024. [Consulté le 18 juillet 2025]. Disponible à l'adresse : <http://arxiv.org/abs/2409.20387>

BAINBRIDGE, Lisanne, 1983. Ironies of Automation. *Automatica* [en ligne]. 1 novembre 1983. Vol. 19, n° 6, pp. 775-779. [Consulté le 19 janvier 2026]. DOI [10.1016/0005-1098\(83\)90046-8](https://doi.org/10.1016/0005-1098(83)90046-8). Disponible à l'adresse : <https://www.sciencedirect.com/science/article/pii/0005109883900468>

BAPTISTA, João, STEIN, Mari-Klara, KLEIN, Stefan, *et al.*, 2020. Digital Work and Organisational Transformation: Emergent Digital/Human Work Configurations in Modern Organisations. *The Journal of Strategic Information Systems* [en ligne]. 1 juin 2020. Vol. 29, n° 2, pp. 101618. [Consulté le 26 juin 2025]. DOI [10.1016/j.jsis.2020.101618](https://doi.org/10.1016/j.jsis.2020.101618). Disponible à l'adresse : <https://www.sciencedirect.com/science/article/pii/S0963868720300263>

BERNARDO, Vitor et HERNÁNDEZ, Laura, 2025. *TechDispatch on Human Oversight of Automated Decision-Making* [en ligne]. Luxembourg : Technology, Privacy Unit of the European Data Protection Supervisor. [Consulté le 20 janvier 2026]. Disponible à l'adresse : <https://www.edps.europa.eu/data-protection/our-work/our-work-by-type/techdispatch>

- BRYNJOLFSSON, Erik, 2022. The Turing Trap: The Promise & Peril of Human-Like Artificial Intelligence. *Daedalus* [en ligne]. 2022. Vol. 151, pp. 272-287. [Consulté le 16 juillet 2025]. DOI [10.1162/daed_a_01915](https://doi.org/10.1162/daed_a_01915). Disponible à l'adresse : <https://consensus.app/papers/the-turing-trap-the-promise-peril-of-humanlike-artificial-brynjolfsson/366ba6811d9e5d19a2315c21a387ad4f/>
- CHRISTENSEN, Emma, 2025. Organizing Expectations: The Co-Constitutive Dynamics of AI Imaginaries. *Human Communication Research* [en ligne]. 4 juin 2025. pp. hqaf14. [Consulté le 17 juillet 2025]. DOI [10.1093/hcr/hqaf014](https://doi.org/10.1093/hcr/hqaf014). Disponible à l'adresse : <https://doi.org/10.1093/hcr/hqaf014>
- FRIEDRICHSEN, Uwe, 2025. AI and the Ironies of Automation - Part 1. *Uwe Friedrichsen* [en ligne]. 21 novembre 2025. [Consulté le 24 décembre 2025]. Disponible à l'adresse : https://ufried.com/blog/ironies_of_ai_1/
- GRØNSUND, Tor et AANESTAD, Margunn, 2020. Augmenting the Algorithm: Emerging Human-in-the-Loop Work Configurations. *The Journal of Strategic Information Systems* [en ligne]. 1 juin 2020. Vol. 29, n° 2, pp. 101614. [Consulté le 26 juin 2025]. DOI [10.1016/j.jsis.2020.101614](https://doi.org/10.1016/j.jsis.2020.101614). Disponible à l'adresse : <https://www.sciencedirect.com/science/article/pii/S0963868720300226>
- JARRAHI, Mohammad Hossein, LUTZ, Christoph, BOYD, Karen, *et al.*, 2023. Artificial Intelligence in the Work Context. *Journal of the Association for Information Science and Technology* [en ligne]. 2023. Vol. 74, n° 3, pp. 303-310. [Consulté le 20 juillet 2025]. DOI [10.1002/asi.24730](https://doi.org/10.1002/asi.24730). Disponible à l'adresse : <https://onlinelibrary.wiley.com/doi/abs/10.1002/asi.24730>
- NATARAJAN, Sriraam, MATHUR, Saurabh, SIDHEEKH, Sahil, *et al.*, 2024. Human-in-the-Loop or AI-in-the-loop? Automate or Collaborate?. *arXiv* [en ligne]. décembre 2024. [Consulté le 16 juillet 2025]. Disponible à l'adresse : <http://arxiv.org/abs/2412.14232>
- SHAO, Yijia, ZOPE, Humishka, JIANG, Yucheng, *et al.*, 2025. Future of Work with AI Agents: Auditing Automation and Augmentation Potential across the U.S. Workforce. *arXiv* [en ligne]. juin 2025. [Consulté le 16 juillet 2025]. Disponible à l'adresse : <http://arxiv.org/abs/2506.06576>
- TE'ENI, Dov, YAHAV, Inbal, ZAGALSKY, Alexely, *et al.*, 2023. Reciprocal Human-Machine Learning: A Theory and an Instantiation for the Case of Message Classification. *Management Science* [en ligne]. 14 novembre 2023. [Consulté le 16 juillet 2025]. DOI [10.1287/mnsc.2022.03518](https://doi.org/10.1287/mnsc.2022.03518). Disponible à l'adresse : <https://pubsonline.informs.org/doi/10.1287/mnsc.2022.03518>
- VACCARO, Michelle, ALMAATOUQ, Abdullah et MALONE, Thomas, 2024. When Combinations of Humans and AI Are Useful: A Systematic Review and Meta-Analysis. *Nature Human Behaviour* [en ligne]. 28 octobre 2024. Vol. 8, n° 12, pp. 2293-2303. [Consulté le 16 juillet 2025]. DOI [10.1038/s41562-024-02024-1](https://doi.org/10.1038/s41562-024-02024-1). Disponible à l'adresse : <http://arxiv.org/abs/2405.06087>
- VUARIN, Louis et STEYER, Véronique, 2025. La résilience organisationnelle face à l'intelligence artificielle : Examen critique du concept de human in the loop en action. *Innovations* [en ligne]. 31 janvier 2025. Vol. 76, n° 1, pp. 239-276. [Consulté le 26 juin 2025]. DOI [10.3917/inno.pr.2.0183](https://doi.org/10.3917/inno.pr.2.0183). Disponible à l'adresse : <https://shs.cairn.info/revue-innovations-2025-1-page-239>
- YAU, K., SALEEM, Yasir, CHONG, Yung-Wey, *et al.*, 2024. The Augmented Intelligence Perspective on Human-in-the-Loop Reinforcement Learning: Review, Concept Designs, and Future Directions. *IEEE Transactions on Human-Machine Systems* [en ligne]. 2024. Vol. 54, pp. 762-777. [Consulté le 16 juillet 2025]. DOI [10.1109/THMS.2024.3467370](https://doi.org/10.1109/THMS.2024.3467370). Disponible à l'adresse : <https://consensus.app/papers/the-augmented-intelligence-perspective-on-chieng-eyu/2e4c71d794d85777a559c5ccef015287/>

LLM

ZOU, Henry Peng, HUANG, Wei-Chieh, WU, Yaozu, MIAO, *et al.*, 2025. A Call for Collaborative Intelligence: Why Human-Agent Systems Should Precede AI Autonomy. *Revue inconnue* [en ligne]. 2025. [Consulté le 16 juillet 2025]. DOI [10.48550/ARXIV.2506.09420](https://doi.org/10.48550/ARXIV.2506.09420). Disponible à l'adresse : <https://arxiv.org/abs/2506.09420>

ZOU, Henry Peng, HUANG, Wei-Chieh, WU, Yaozu, CHEN, *et al.*, 2025. LLM-Based Human-Agent Collaboration and Interaction Systems: A Survey. *arXiv* [en ligne]. 2025. [Consulté le 16 juillet 2025]. DOI [10.48550/ARXIV.2505.00753](https://doi.org/10.48550/ARXIV.2505.00753). Disponible à l'adresse : <https://arxiv.org/abs/2505.00753>

Varia

MESGUICH, Véronique, 2023. *Les bibliothèques face au monde des données* [en ligne]. Villeurbanne : Presses de l'enssib. [Consulté le 22 janvier 2026]. Papiers. ISBN 978-2-37546-148-8. Disponible à l'adresse : <https://books.openedition.org/pressesenssib/17655>

Codage agentique

ALDERSON, Martin, 2025. Has the Cost of Building Software Just Dropped 90%?. *Martin Alderson* [en ligne]. 8 décembre 2025. [Consulté le 29 décembre 2025]. Disponible à l'adresse : <https://martinalderson.com/posts/has-the-cost-of-software-just-dropped-90-percent/>

METR, 2025. How Does Time Horizon Vary Across Domains?. *METR Blog* [en ligne]. 14 juillet 2025. [Consulté le 16 juillet 2025]. Disponible à l'adresse : <https://metr.org/blog/2025-07-14-how-does-time-horizon-vary-across-domains/>

RONACHER, Armin, 2025. 90%. *Armin Ronacher's Thoughts, Writings* [en ligne]. 29 septembre 2025. [Consulté le 29 décembre 2025]. Disponible à l'adresse : <https://lucumr.pocoo.org/2025/9/29/90-percent/>

GLOSSAIRE

Chaîne agentique Suite d'opérations menées à bien par un LLM, lequel utilise les outils à sa disposition pour enchaîner plusieurs étapes de traitement. Il prend les décisions intermédiaires nécessaires à l'accomplissement de la fin dernière qui lui a été assignée.

Chaîne de traitement Enchaînement d'opérations humaines et techniques menant à la production ou mise à jour de métadonnées.

Données d'entraînement Données utilisées pour entraîner un modèle, à même de faire des prédictions (apprentissage machine traditionnel) ou de générer du texte (LLM).

Dédoublonnage Détection et résolution de doublons dans un catalogue ou une base mutualisée.

Fichier d'autorités Référentiel normalisé d'entités (agents, sujets) fournissant formes retenues, variantes, relations, identifiants, et règles d'usage.

Fine-tuning Spécialisation d'un LLM pré-entraîné par un entraînement supplémentaire sur un corpus ciblé.

Humain dans la boucle Au sein d'une chaîne de traitement partiellement automatisée, maintien d'une vérification humaine afin de maîtriser les risques et d'obtenir le niveau de qualité souhaité.

Indexation automatique Recours à l'apprentissage machine traditionnel, à des LLM, ou à la vectorisation pour assigner à des documents leurs métadonnées sémantiques, à partir de leurs métadonnées descriptives ou de leur contenu.

Indexation matière Attribution à une ressource de vedettes-matière renseignant sur son contenu et la classant.

Modèle transformeur Famille de modèles fondée sur l'attention (angl. transformers). Les modèles générant des vecteurs (angl. embeddings), les modèles de langue (BERT, GPT, etc.) appartiennent à cette catégorie.

Métadonnées descriptives Métadonnées tirées de la ressource elle-même, consignées sans guère de place pour l'interprétation, et qui la décrivent dans les caractéristiques essentielles qui la singularisent (titre, auteur, etc.).

Métadonnées sémantiques Métadonnées le plus souvent conformes à des vocabulaires contrôlés, renseignant sur le contenu de la ressource et la classant, qui procèdent d'une opération d'indexation humaine ou algorithmique.

Recherche sémantique Recherche s'appuyant sur des représentations vectorielles (angl. embeddings) pour identifier des contenus de sens similaire qui ne présentent pas nécessairement des correspondances lexicales exactes.

Relecture par échantillonnage Contrôle qualité fondé sur un prélèvement de notices (aléatoire ou ciblé) afin d'estimer la qualité globale et de détecter des dérives.

Topic modelling Famille de méthodes identifiant automatiquement des thèmes dans un corpus, pouvant servir à produire des regroupements ou des descripteurs.

Vecteur (angl. embedding) Représentation numérique dense d'un texte de toute longueur, utilisée pour calculer des proximités sémantiques.

Vectorisation Conversion d'un texte de toute longueur en vecteur numérique (angl. embedding) permettant recherche et comparaison sémantiques.

Vedette-matière Mot ou groupe de mots représentant un sujet dans un vocabulaire d'indexation (ex. RAMEAU, LCSH, FAST).

Vocabulaire contrôlé Ensemble structuré de vedettes conçues pour l'indexation, géré par une autorité centrale afin d'assurer cohérence, stabilité et facilité de réutilisation.

INDEX

A

Abes 10, 18, 21, 24, 25, 28, 29, 42, 46
ADBU 10
AI4LAM 9
 Fantastic Futures 9
Alma 10, 21, 49
Annif 10, 20, 25, 27, 29, 30, 32, 34, 45,
 46, 47, 48
API 10, 17, 28, 49, 50
assistance au catalogage 12
automatisation du signalement 9
 conséquences pour la
 spécialisation linguistique et
 disciplinaire 19
 conséquences sur la productivité
 20, 23
 défis du passage en production 21
 intérêt des agents 24 sq.
 nécessite des préparatifs
 minutieux 20 sq.
 opportunités de redéploiement
 des agents 25 sq.
 projet au long cours 20
 requiert d'importants
 investissements 19 sq.
AutoSE 10, 20, 26, 28, 29, 30, 45

B

BnF 18, 21, 24, 25, 30, 32, 34, 35, 36,
 39

C

CENL 9
chaîne agentique 10, 28, 32
contrôle qualité 12, 29, 30, 39, 48
craintes pour l'emploi 23 sq.

D

dédoublonnage 10, 26
DNB 10, 18, 19, 20, 24, 26, 29, 33, 34,
 36, 41, 42, 44, 45, 46, 47, 48, 51, 52
données d'entraînement 12, 17, 30

E

environnementaux, enjeux 22
éthiques, enjeux 14, 21 sq.
étiquetage des métadonnées 14
Ex Libris 10, 22, 26, 49
expertise catalographique
 amplifiée par l'aptitude à
 manipuler des données 30 sq.
 explicitée dans une chaîne
 agentique 31 sq.
 menacée par l'IA 36 sq., 38 sq.
 nécessaire au cours du
 signalement 27 sq.
 nécessaire pour évaluer et faire
 évoluer le système 28–30
 transmission après adoption de
 l'IA 37 sq.
expertise interne en apprentissage
 machine
 indispensable au maintien de la
 maîtrise sur le catalogue 48–50
 présente dans les projets réussis
 45 sq.
 requiert des profils spécifiques 46
 sq.

F

fichier d'autorités 9, 10
fine-tuning 17, 22, 30
formation 13
 est une nécessité 44
 ne saurait gommer les périmètres
 fonctionnels 44 sq.

G

GND 34, 47

I

indexation matière 9, 21, 26, 48
classification multilabel extrême
12
harmonisée par l'IA 32
réévaluation de la qualité à la
lumière de l'IA 32 sq.
variabilité des résultats humains
32
vraisemblance d'une refonte 33 sq.
interopérabilité 17, 32

J

juridiques, enjeux 14, 22

K

KB 12, 19, 20, 21, 24, 26, 27, 35, 37, 47,
48, 51
KBR 10, 17, 19, 28, 31, 35, 46, 47, 48,
51

L

LC 10, 20
LIBER 9
LLM 10, 17, 19, 20, 21, 22, 27, 28, 30,
32, 34, 45, 46, 48, 49, 52

M

MARC 12, 19, 21, 37, 42, 48
masse documentaire croissante 18
conséquences sur la notion de
qualité 18
spécificité des agences
bibliographiques 18
métadonnées descriptives 9, 19, 42,
46, 48, 51, 52
métadonnées sémantiques 9, 17
modèle transformeur 20, 33

N

NIST 20
Noemi 21

O

OCLC 10, 19, 37, 40
WorldCat 10, 17, 26, 27, 31
OCR 31
open source 15, 45, 48

P

PCC 9, 29, 31, 44, 46
prestataires externes 19, 22, 24, 45
supplantés par l'expertise interne
dans les projets d'indexation
automatique 47
utiles à la création d'une chaîne
semi-automatisée de description
bibliographique 47 sq.

R

RAG 32, 34
RDA 9, 35
RDF 35
relecture 35 sq.
menace l'expertise
catalographique 36 sq., 38 sq.
menace potentielle pour la fierté
tirée du travail 40–42
menace potentielle pour la
motivation professionnelle 39 sq.
vue comme une montée en
généralité 36

S

sens du travail
a des sources variées selon les
agents 42 sq.
SIGB 21, 22, 49

T

topic modelling	34
typologie des projets	
deux ensembles se dégagent	48

V

vectorisation	20, 33, 34, 46
vocabulaire contrôlé	46, 47, 48
FAST	11, 30, 33, 34
LCSH	11, 33
RAMEAU	21, 25, 28, 30, 34, 46

Z

ZBW	20, 22, 24, 26, 28, 29, 34, 36, 37, 42, 45, 47, 48, 51
-----	---

TABLE DES MATIÈRES

SOMMAIRE	5
SIGLES ET ABRÉVIATIONS	7
INTRODUCTION	9
1.1. Accélération récente d'un phénomène ancien	9
1.1.1. Mobilisation des réseaux professionnels	9
1.1.2. Une cadence précipitée depuis 2022	10
1.2. Périmètre de l'étude	11
1.3. Des chaînes de signalement hétérogènes : typologie et exemples	12
1.4. Quelle bascule organisationnelle ?	13
1.5. Principes de sélection des cas	14
1.6. Plan schématique	14
ADOPTION INSTITUTIONNELLE DE L'IA : QUELS DÉFIS ?	17
2.1. Un domaine propice à l'automatisation contrôlée	17
2.2. Un problème d'échelle : volumes et arriérés	18
2.3. Un substitut à la spécialisation linguistique et disciplinaire ?	19
2.4. Mobiliser les ressources nécessaires	19
2.5. Interopérabilité, passage à l'échelle et mise en production	21
2.6. Enjeux éthiques	21
QUELLES PERSPECTIVES POUR L'EMPLOI DES AGENTS CONCERNÉS ?	23
3.1. Des craintes pour l'emploi	23
3.2. Adhésion des agents : curiosité, prudence, co-construction	24
3.3. Redéployer le travail humain	25
UNE PERMANENCE INDISPENSABLE DE L'EXPERTISE CATALOGRAPHIQUE	27
4.1. Implication de l'expert dans le signalement	27
4.2. L'expert au cœur du cycle de vie des systèmes	28
4.3. Un précédent favorable : le <i>data librarian</i>	30
4.4. Vers une nouvelle application de l'expertise	31
4.5. Avenir de l'indexation	32
4.5.1. Un primat humain contesté	32
4.5.2. Maintien des vocabulaires contrôlés	33
DU CATALOGUEUR PRODUCTEUR AU CATALOGUEUR RELECTEUR : UNE TRANSFORMATION AMBIVALENTE	35
5.1. Une montée en généralité	35
5.2. Une menace pour l'expertise ?	36
5.2.1. Une potentielle perte de compétences	36
5.2.2. Formation des recrues	37
5.3. Ironies de l'automatisation	38
5.4. Automatisation et sens du travail	39
5.4.1. Un risque de dévalorisation	39
5.4.2. L'attachement au travail bien fait	40
5.5. Des sources de satisfaction professionnelle variées	42

FORMATION ET PROFILS TECHNIQUES	44
6.1. Former en respectant les frontières professionnelles	44
6.2. Des équipes pionnières pourvues d’une forte expertise	45
6.3. Des profils techniques indispensables	46
6.4. Prestataires ou compétences internes ?	47
6.5. Des experts indispensable à la compréhension de l’IA	48
CONCLUSION	51
7.1. Synthèse des observations	51
7.2. Vers une génération semi-automatique des métadonnées descriptives ?	51
7.3. Enseignements	52
SOURCES	54
Bibliographies	54
IA en bibliothèque (transversal)	54
Généralités	54
Signalement	56
Indexation matière	58
Par bibliothèque et pays	59
Allemagne – Autriche	59
DNB	59
ZBW	60
Chili	60
Estonie	61
États-Unis – Royaume-Uni – Australie	61
Bibliothèque du Congrès	61
Autres	62
France	62
Abes	63
BNF	63
Université Paris-Saclay	64
Israel	64
Luxembourg – Suisse – Belgique	64
Pays-Bas	65
Scandinavie	66
Finlande	66
Norvège	66
Suède	66
Singapour	67
Par organisation professionnelle ou entreprise	67
ADBU	67
ARL – CNI	67
Ex Libris	67
IFLA	68
OCLC	68
Autres	69
RAG en bibliothèque	69

Sciences humaines et sociales	70
Philosophie	70
Économie et sociologie	70
Collaboration homme-machine	72
LLM	74
Varia	74
Codage agentique	74
GLOSSAIRE	76
INDEX	77
TABLE DES MATIÈRES	81