

The Canadian Journal of Information and Library Science
La Revue canadienne des sciences de l'information et de
bibliothéconomie



Peut-on parler de l'automatisation comme cinquième
paradigme archivistique ?
Can we speak of automation as the fifth archival paradigm?

Siham Alaoui 

Volume 47, numéro 1, 2024

URI : <https://id.erudit.org/iderudit/1112106ar>

DOI : <https://doi.org/10.5206/cjils-rcsib.v47i1.17111>

[Aller au sommaire du numéro](#)

Éditeur(s)

Canadian Association for Information Science - Association canadienne des sciences de l'information

ISSN

1195-096X (imprimé)

1920-7239 (numérique)

[Découvrir la revue](#)

Citer cet article

Alaoui, S. (2024). Peut-on parler de l'automatisation comme cinquième paradigme archivistique ? *The Canadian Journal of Information and Library Science / La Revue canadienne des sciences de l'information et de bibliothéconomie*, 47(1), 18–34. <https://doi.org/10.5206/cjils-rcsib.v47i1.17111>

Résumé de l'article

Au cours des dernières années, on assiste à la montée en puissance d'un discours en faveur de l'intégration des fonctionnalités de l'intelligence artificielle aux pratiques de gestion des documents et des archives. Cela s'inscrit dans une optique d'automatisation de ces pratiques afin d'améliorer la qualité de l'information documentaire et multiplier les opportunités de son exploitation par les usagers, et ce, suivant leurs préoccupations. Cela nous amène à nous interroger sur la possibilité de parler d'un cinquième paradigme archivistique, soit celui de l'automatisation, lequel commande une révision des procédés par lesquels l'information documentaire est créée, traitée, diffusée et préservée à long terme. Sur la base de la cartographie des paradigmes archivistiques établie par Cook (2013), cet article vise à explorer la manière dont l'automatisation peut s'inscrire dans la continuité de ces paradigmes, tout en mettant de l'avant les nouveaux rôles à jouer par les archivistes et les gestionnaires de documents dans ce contexte.

© Siham Alaoui, 2024



Ce document est protégé par la loi sur le droit d'auteur. L'utilisation des services d'Érudit (y compris la reproduction) est assujettie à sa politique d'utilisation que vous pouvez consulter en ligne.

<https://apropos.erudit.org/fr/usagers/politique-dutilisation/>

érudit

Cet article est diffusé et préservé par Érudit.

Érudit est un consortium interuniversitaire sans but lucratif composé de l'Université de Montréal, l'Université Laval et l'Université du Québec à Montréal. Il a pour mission la promotion et la valorisation de la recherche.

<https://www.erudit.org/fr/>

Peut-on parler de l'automatisation comme cinquième paradigme archivistique?

Siham Alaoui 

Université Laval

Au cours des dernières années, on assiste à la montée en puissance d'un discours en faveur de l'intégration des fonctionnalités de l'intelligence artificielle aux pratiques de gestion des archives. Cela s'inscrit dans une optique d'automatisation de ces pratiques afin d'améliorer la qualité de ces traces documentaires et multiplier les opportunités de leur exploitation par les usagers, et ce, suivant leurs préoccupations. Cela nous amène à nous interroger sur la possibilité de parler d'un cinquième paradigme archivistique, soit celui de l'automatisation, lequel commande une révision des procédés par lesquels les archives sont créées, traitées, diffusées et préservées à long terme. Sur la base de la cartographie des paradigmes archivistiques établie par Cook (2013), cet article vise à explorer la manière dont l'automatisation peut s'inscrire dans la continuité de ces derniers, tout en mettant de l'avant les nouveaux rôles à jouer par les archivistes dans ce contexte.

Mots Clés: paradigme archivistique, intelligence artificielle, automatisation, transformation numérique, archivistes

Les développements technologiques ont modifié la manière dont les organisations actuelles conduisent leurs activités. Dans l'univers archivistique, on assiste à une redéfinition des modalités par lesquelles les documents d'archives sont créés, traités, diffusés et préservés à long terme dans des environnements numériques (Duranti et Rogers, 2012). Ils sont capturés par différents systèmes, agrégés à d'autres objets documentaires (ex. données, renseignements personnels) et décrits de plusieurs manières en fonction des activités dans lesquelles ils s'insèrent (Duranti et Thibodeau, 2006; Lappin et al., 2021). Par surcroît, les documents sont aisément créés et partagés par une pluralité d'acteurs qui possèdent des compétences archivistiques, numériques et intellectuelles variées. Cela engendre des pratiques hétérogènes de gestion qui compromettent la qualité des documents d'archives en regard de leur authenticité, fiabilité, intégrité et réexploitabilité (Bak, 2012; Lappin et al., 2021). Une piste prometteuse pour pallier en partie ces enjeux est le recours aux dispositifs de l'intelligence artificielle (IA). Renvoyant aux programmes informatiques mis à profit pour automatiser des processus en mobilisant des compétences similaires à celles de l'être humain, l'usage de l'IA mise sur la réduction des redondances dans l'exécution de diverses tâches (Mueller et Massaron, 2021). C'est dans cette lignée de pensée qu'on assiste, au cours des dernières années, au développement d'un

discours scientifique abordant la manière dont l'IA peut être intégrée aux pratiques archivistiques pour les automatiser (Colavizza et al., 2021; Darda, Carre et Cross, 2023; Han et al., 2022; Makhoul-Shabou et al., 2020; Pötzsch, 2018; Rolan et al., 2019; Terras, 2022). Ces écrits reflètent, certes, une conscience collective à l'égard de l'importance de l'intégration de l'IA aux pratiques archivistiques, notamment en revisitant les principes et les théories fondant l'archivistique contemporaine.

Par ailleurs, la collaboration entre les chercheurs en provenance de l'archivistique et de l'informatique a donné naissance à une nouvelle discipline scientifique, celle de l'archivistique computationnelle (*Computational archival science*) (Hedges, Marciano et Goudarouli, 2022; Williams, 2019). Celle-ci se veut le fruit de la transposition des approches et méthodes reconnues en informatique vers l'archivistique en vue d'automatiser les modalités par lesquelles la gestion des documents d'archives s'opère. Plus précisément, il est question du recours aux techniques de l'IA, couplées au traitement automatique de la langue naturelle (*Natural language processing*) en vue de traiter de larges corpus documentaires pour en extraire les caractéristiques (*patterns*) clés. Celles-ci servent à soutenir la classification, l'indexation et l'évaluation des documents d'archives (Colavizza et al., 2021; Duranti et al., 2021; Makhoul-Shabou et al., 2020). Il est aussi question de la mobilisation des techniques de reconnaissance de l'écriture manuscrite (*Handwritten Text Recognition* : HTR) à des fins de transcription des manuscrits pour en favoriser une meilleure accessibilité (Colavizza et al., 2021; Terras, 2022). D'autres chercheurs mettent de l'avant le recours aux

techniques d'apprentissage non-supervisé pour la classification des documents en fonction de leur degré de criticité, notamment dans l'optique du déploiement des technologies d'amélioration de la confidentialité (*Privacy Enhancing technologies*) (Lemieux et Werner, 2024).

Certains chercheurs rendent compte des effets de l'IA sur la redéfinition des modes de pensée et des pratiques archivistiques. Ils voient que l'automatisation repose sur la *datification* des archives, c'est-à-dire leur transformation en données massives traitables par la machine (Jaillant, 2021). Un discours lié à l'encadrement de l'IA est également en cours de développement, notamment pour pallier les enjeux éthiques liés à l'usage des données personnelles par les algorithmes (Jaillant et Caputo, 2022). Considérant la place de plus en plus importante que l'IA a commencé à se tailler dans l'univers des archives, il semble légitime de s'interroger sur la possibilité de désigner l'automatisation comme cinquième paradigme archivistique. Celui-ci s'inscrirait dans la continuité de la pensée de Cook (2013) dans son article sur les quatre paradigmes archivistiques, soit celui de la preuve, de la mémoire, de l'identité et de la communauté (Cook, 2013).

Dans certaines disciplines, des chercheurs ont d'ores et déjà commencé à désigner l'automatisation comme paradigme, par exemple en sciences de la gestion (Racanati, 2023), en informatique (Jin et al., 2023), en sciences des données (Balasubramanian, 2023) et en marketing (Silva et al., 2021). Un paradigme renvoie à l'ensemble des idéologies et des pratiques collectivement reconnues et acceptées au sein d'une communauté, et qui illustrent la manière dont un phénomène est appréhendé et approprié (Cook, 2013). Par transposition dans l'univers technologique, le paradigme de l'automatisation réfère aux modalités du recours aux dispositifs technologiques de l'IA pour automatiser les tâches, notamment dans une perspective d'efficacité et d'efficience (Balasubramanian, 2023; Silva et al., 2021). Il inclut également les questions sur le changement de la culture, des rôles et des responsabilités des acteurs qui réalisent ces tâches, et le degré de leur remplacement par la machine (Balasubramanian, 2023; Racanati, 2023). En outre, il s'agit d'y intégrer une vision socioculturelle, portant sur l'acceptabilité sociale de ces dispositifs d'automatisation, leur encadrement par des lois, des règlements et des cadres normatifs collectivement reconnus comme valides (Balasubramanian, 2023). Il est aussi question des compétences à acquérir pour faire le bon usage de ces dispositifs, ainsi que la manière dont ces dispositifs réinventent les modalités de communication entre les acteurs qui assument ces tâches (Balasubramanian, 2023).

Les aspects mobilisés par les chercheurs dans la définition du paradigme de l'automatisation, soit la recherche de l'efficacité et de l'efficience, l'agentivité, l'encadrement, l'acceptabilité sociale, le consensus, le développement des compétences et la communication humaine, rejoignent ceux mis à profit par Cook (2013) dans son argumentaire sur l'évolution

des paradigmes archivistiques. Cependant, la conceptualisation de l'automatisation en tant que paradigme archivistique en tenant compte de ces aspects reste à explorer. Ainsi, nous formulons la question de recherche suivante : pourrait-on parler de l'automatisation comme un nouveau paradigme archivistique ? Si oui, quelles en seraient les principales caractéristiques ? Quelles en seraient des intersections avec les quatre paradigmes conceptualisés par Cook (2013) ? Comment cette mise en dialogue redéfinirait-elle le rôle de l'archiviste ? En vue de répondre à ces questions de recherche, nous allons mobiliser des corpus scientifiques en archivistique, notamment en identifiant les besoins soulevés dans la littérature en regard de la gestion des documents d'archives et en explorant la manière dont les quatre paradigmes de Cook (2013) y répondent. Ces derniers, après avoir été rappelés, seront mis en dialogue avec l'automatisation comme un cinquième paradigme archivistique qui ajoute un nouveau profil à l'archiviste : celui de l'archiviste-informaticien. Dans chacun des quatre paradigmes, nous brosserons un portrait des écrits scientifiques (ex. réflexions théoriques, études de cas et retours d'expériences) abordés par les chercheurs quant à l'intégration de l'IA aux pratiques archivistiques. L'idée est d'évaluer le degré du consensus scientifique à l'égard de l'automatisation et de la manière dont elle soutiendrait l'archiviste. Enfin, nous allons décrire la manière dont le rôle de l'archiviste-informaticien compléterait celui de l'archiviste tel qu'illustré dans les quatre paradigmes de Cook.

Les quatre paradigmes archivistiques de Cook (2013)

Dans son article publié dans la revue *Archival Science*, Cook (2013) retrace l'évolution de la pensée archivistique et la manière dont celle-ci réinvente les pratiques archivistiques. Il met l'accent sur le passage d'une gestion documentaire axée sur les besoins du producteur vers celle qui prend de plus en plus en considération les besoins des usagers des archives. Cette évolution s'étend en dehors des frontières institutionnelles, où l'archiviste s'implique auprès de diverses communautés pour enrichir le patrimoine archivistique et améliorer sa représentativité fidèle de la société (Cook, 2013). Cook (2013) représente cette évolution sous forme de quatre paradigmes archivistiques, soit (1) la preuve, (2) la mémoire, (3) l'identité et (4) la communauté (Cook, 2013).

Le paradigme de la *preuve* renvoie, selon Cook (2013) au rôle de l'archiviste dans la constitution des archives en tant que preuves des activités des producteurs. Sous l'influence d'un tel paradigme, l'archiviste, incarnant le rôle d'un juriste, met à profit les procédés archivistiques qui visent à constituer des traces documentaires fidèles des activités du producteur. Il préserve toutes les archives et maintient leur arrangement, car dans celui-ci réside la traçabilité des activités assumées par le producteur. Autrement dit, il est question d'appliquer le principe de respect des fonds (Duchéin, 1977) en respectant l'ordre primitif des documents d'archives conservés dans leur

totalité. Cependant, compte tenu de la nécessaire optimisation de la gestion des ressources allouées aux organisations publiques pour mener à bien leurs activités, le besoin de rationaliser la production documentaire a émergé.

Le paradigme de la *mémoire* illustre, selon l'auteur, le passage vers une culture de sélection des archives, fondée à la fois sur les besoins pragmatiques, sociétaux et légaux du producteur et sur les attentes des usagers des archives. Cook (2013) se sert de ces considérations pour justifier le développement de la réflexion sur l'évaluation des archives (la macro et la micro évaluation) et la manière dont l'archiviste intervient pour déterminer les valeurs des archives. En empruntant le rôle d'un historien, l'archiviste sélectionne ces traces en fonction de leurs propriétés intrinsèques et les circonstances dans le cadre desquelles elles ont été produites ou reçues. Par l'application de ces critères, l'archiviste conserve uniquement les archives qui répondent aux besoins des producteurs et des usagers, tandis que le reste est voué à l'élimination pour répondre à des impératifs de rationalisation de la production documentaire (Cook, 2013).

La modernisation des appareils administratifs et le développement de la réglementation sur l'accès équitable aux documents publics ont renforcé le rôle de l'archiviste comme expert. C'est ainsi qu'on voit naître des associations professionnelles, des programmes de formations universitaires et des certifications professionnelles visant à renforcer l'identité socioprofessionnelle de l'archiviste et aider celui-ci à nouer des relations de collaboration avec les autres experts en gestion de l'information (Cook, 2013). C'est ce qui reflète le paradigme de l'*identité* qui renvoie, selon l'auteur, aux efforts déployés par les spécialistes et leurs corps collectifs pour construire une représentation socioprofessionnelle de l'archiviste (Cook, 2013). Cela inclut le rôle de ce spécialiste comme expert maîtrisant l'art de la constitution des traces documentaires, leur traitement, leur agrégation à une mémoire, ainsi que leur diffusion à un large segment d'utilisateurs (Cook, 2013). Il agit également comme médiateur entre le passé et le présent, et fait de la gestion des archives une pratique répondant aux enjeux de la société. En clair, le paradigme de la mémoire incarne l'ouverture de la communauté archivistique vers la société, ce qui fait évoluer l'identité socioprofessionnelle de l'archiviste comme expert de la gestion des archives.

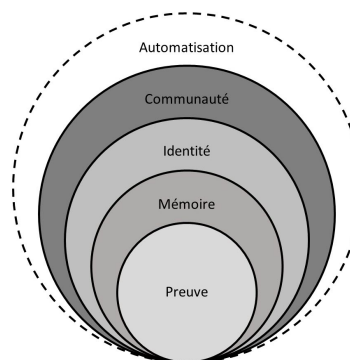
Cette identité a continué à évoluer au-delà des frontières institutionnelles, puisque l'archiviste a commencé à s'intéresser aux communautés socioculturelles et à leurs pratiques autonomes de la constitution des archives. C'est dans cette lignée de pensée que s'est développé le paradigme de la *communauté*, illustrant les efforts déployés par l'archiviste qui oriente désormais son rôle vers une voie plus sociétale, afin de réduire les silences et nouer des liens de collaboration avec ces communautés pour enrichir le patrimoine documentaire dont il est chargé de la gestion (Cook, 2013). Dans ce paradigme, l'archiviste est considéré comme un activiste dont la

voix s'allie à celle des communautés pour faire valoir leurs différentes identités socioculturelles à travers les archives.

Ces quatre paradigmes sont perçus, selon Cook (2013), comme des courants de pensée évoluant dans une logique incrémentielle, puisque le paradigme subséquent n'écrase pas ceux antérieurs, mais les complète et les inclut davantage dans la définition de sa portée (Cook, 2013). Autrement dit, le rôle de l'archiviste comme activiste, tel qu'il est perçu dans le paradigme de la communauté, n'exclut pas ses rôles dans la constitution de la preuve des activités, l'enrichissement de la mémoire institutionnelle et la construction de son identité socioprofessionnelle (Cook, 2013). Il les inclut davantage et les complète en y ajoutant une touche sociale qui offre à l'archiviste la possibilité de s'impliquer avec les diverses communautés qui existent dans une société donnée. Cela aide l'archiviste à participer au processus de constitution, de traitement, de diffusion et de préservation à long terme des archives pour garantir une meilleure fidélité des archives comme reflet des identités qui peuvent exister dans une société.

Figure 1

Les quatre paradigmes archivistiques de Cook (2013) et l'automatisation



Avec les développements technologiques actuels et l'engouement pour l'exploitation de l'IA à des fins d'optimisation des affaires organisationnelles, on commence à s'intéresser au rôle de l'IA dans l'amélioration des pratiques de gestion des documents d'archives (Colavizza et al., 2021; Darda, Carre et Cross, 2023; Han et al., 2022; Makhoul-Shabou et al., 2020; Pötzsch, 2018; Rolan et al., 2019; Terras, 2022). C'est ainsi qu'on s'interroge sur la place qu'occuperait l'automatisation dans la théorie et la pratique archivistiques. Pourrait-on parler de l'automatisation comme cinquième paradigme archivistique? Si oui, jusqu'à quel point? Nous postulons qu'il serait possible d'ajouter ce paradigme en l'inscrivant dans la continuité de l'évolution de la pensée archivistique contemporaine (voir figure 1). Ce paradigme aboutirait à la conceptualisation d'un cinquième profil de l'archiviste, soit celui de l'informaticien. Ainsi, nous allons

mettre en lumière la manière dont l'automatisation est mise en dialogue avec les paradigmes de la preuve, de la mémoire, de l'identité et de la communauté de Cook (2013). Nous soulignerons également les différentes intersections et complémentarités entre les profils l'archiviste comme juriste, historien, médiateur et activiste, ainsi que leurs relations avec celui de l'informaticien.

Les intersections entre les quatre paradigmes de Cook (2013) et l'automatisation

Cook (2013) considère que la logique de l'évolution des paradigmes archivistiques présuppose que chacun apporte un nouveau souffle au précédent, compte tenu des mutations technologiques et socio-politico-institutionnelles entourant le travail de l'archiviste. Le paradigme de la mémoire complète celui de la preuve, de la même manière que ces deux derniers sont enrichis par le discours sur l'identité socioprofessionnelle de l'archiviste et son rôle social en tant qu'activiste (Cook, 2013). Considérer l'automatisation comme cinquième paradigme archivistique exigerait l'exploration de ses intersections avec les quatre paradigmes de Cook (2013), notamment en mettant en lumière la manière dont il les redéfinirait et les repositionnerait dans le discours et la pratique actuels de la discipline et la profession archivistiques.

Preuve et automatisation

Dans le paradigme de la preuve, il est question de constituer les traces documentaires qui rendent compte des activités organisationnelles, que celles-ci soient stratégiques ou opérationnelles (Cook, 2013). Ces traces aident l'organisation à se défendre dans le cas des litiges, améliorer sa productivité et affirmer son identité collective (Cardin, 1995). Dans un environnement de travail marqué par la cohabitation des applications du télétravail avec les systèmes d'information organisationnels, plusieurs objets documentaires sont générés. Il est question des documents structurés et non structurés, des conversations *Teams*, des données massives et ouvertes, ainsi que des publications sur les médias sociaux. Tous ces types d'objets documentaires préfigurent des logiques différentes de capture et de traitement, pour être exploitables à titre de preuves d'activités organisationnelles courantes.

Par ailleurs, dans le paradigme de la preuve, l'emphasis est mise sur le processus de la constitution des archives comme traces documentaires fidèles, objectives et authentiques reflétant les activités de leur producteur (Cook, 2013). Il n'est dès lors pas étonnant de constater que la pierre angulaire de la théorie archivistique, à savoir le principe de la provenance, soit mise de l'avant dans la description de ce paradigme par Cook (2013). Alors que l'approche traditionnelle de la gestion des archives analogiques reconnaît l'importance de l'application d'un tel principe dans la contextualisation, le classement et la description des archives, il n'en va pas de même pour la gestion de ces traces dans les environnements numériques

(Duranti et Thibodeau, 2006; Lappin et al., 2021). En fait, avec le déploiement massif des plateformes collaboratives dans le cadre de la mise en place des environnements numériques du travail, les relations que les acteurs organisationnels entretiennent avec les documents d'archives sont diverses. Un même document peut être détenu à la fois par des agents affiliés à plusieurs unités administratives, mais les rapports que ces derniers entretiennent avec ce document sont différents. Cela se justifie par le fait qu'un document d'archives est toujours en devenir (Mckemmish, 1994), ce qui laisse entendre qu'il se recycle continuellement et s'agrège à d'autres objets documentaires facilement dans le contexte numérique (Yeo, 2012). Ces multiprovenances donnent des significations multiples à un même document, ce qui commande la nécessaire capture de ces agrégations dans un objectif d'arriver à une meilleure contextualisation du document d'archives. Duranti (1997) suggère le concept du lien organique (*archival bond*) illustrant les liens qu'un même document entretient avec les autres documents traitant de la même activité (Duranti, 1997). L'auteure pose qu'il est essentiel de capturer les données qui documentent ces agrégations, car elles font partie du contexte de création et d'utilisation de ce document (Duranti, 1997). Bak (2012) soutient cette posture et met de l'avant la description au niveau micro de ces documents, et ce, en cernant ses différentes logiques d'agrégation par les usagers en vue de veiller à leur contextualisation fidèle pour en améliorer l'intelligibilité.

L'IA pourrait être mise à profit à des fins d'optimisation des pratiques de gestion documentaire, grâce à l'usage des algorithmes développés selon la technique d'apprentissage supervisé. Cela pourrait soutenir l'analyse des corpus documentaires numériques et analogiques numérisés. Pour ces derniers, la mise à profit des techniques de reconnaissance optique des caractères pourrait être combinée aux fonctionnalités de l'IA liées à la reconnaissance du texte manuscrit (Duranti et al., 2021; Meng et Ghena, 2023; Terras, 2022). Celle-ci s'appuie sur les algorithmes de traitement automatique de la langue, notamment par le biais des techniques d'apprentissage supervisé (Murphy et al., 2023). Ce dernier est utilisé à des fins de traduction, de génération des résumés automatiques et d'indexation automatique (Smiraglia et Cai, 2017). Certains chercheurs se sont appuyés sur des études de cas pour évaluer le potentiel de l'IA à des fins de constitution et de gestion des archives comme preuves des activités (Colavizza et al., 2021; Meng et Ghena, 2023; Rolan et al., 2019; Terras, 2022). Sur la base de leurs résultats, ils avancent que l'identification des caractéristiques des corpus documentaires grâce à ces algorithmes aiderait à dégager, sous forme de mots-clés, les concepts les plus représentatifs des documents analysés (contenu et contexte) (Meng et Ghena, 2023; Terras, 2022). Une autre avenue qui pourrait être explorée est celle de la détection, par l'IA, des doublons enregistrés dans le système de gestion documentaire (Rolan et al., 2019). Cela

aiderait l'archiviste à les sélectionner et d'en identifier les similarités et les différences. Ainsi, dans le cas d'un même document appartenant à plus d'une seule unité administrative, l'IA rendrait plus aisée l'exploration des logiques d'agrégation de ce document à d'autres objets documentaires pour en assurer une meilleure contextualisation. La comparaison des métadonnées associées à ces documents favoriserait une meilleure capture des agrégations multiples d'un même document aux d'autres objets documentaires numériques. Ainsi, ces fonctionnalités de l'IA pourraient soutenir la traçabilité des activités en constituant des traces fidèles, contextualisées et exploitables.

Le paradigme de la preuve met de l'avant l'idée selon laquelle la traçabilité des activités répond à un besoin d'exploitation des archives par leur producteur (Cook, 2013). Or, pour que celles-ci soient pleinement exploitables, il importe d'en améliorer la repérabilité par des outils archivistiques performants, dont le plan de classification. Celui-ci se veut une structure hiérarchique de classes et de sous-classes, basée sur les activités organisationnelles et utilisée à des fins de classification des archives selon les activités qu'elles décrivent (Héon, 1995). Alors que traditionnellement, l'archiviste attribue manuellement aux documents d'archives une cote classificatoire en fonction des activités qu'ils décrivent, une telle opération pourrait être automatisée par le biais de l'usage des fonctionnalités de l'IA (Colavizza et al., 2021). Tout comme pour l'indexation automatique, les algorithmes de l'IA, reposant sur le traitement automatique du langage naturel et les techniques de forage des textes et des données, puisent dans les documents d'archives pour en faire ressortir des caractéristiques à mettre en lien avec les rubriques classificatoires du plan de classification. Cela améliorerait l'accessibilité de ces documents pour leur producteur afin de réaliser ou poursuivre ses activités.

Par ailleurs, la gestion documentaire soutient le devoir institutionnel de protection de la vie privée. Les organisations collectent des renseignements personnels sur leurs employés, leurs clients et leurs partenaires pour des besoins d'affaires. Les documents renfermant ces données doivent faire l'objet d'une gestion adéquate avec des mesures de protection de la vie privée. L'archiviste se soucie la gestion des documents nominatifs et confidentiels et leur catégorisation suivant leur degré de criticité. Alors qu'un tel processus se fait manuellement, en ajoutant une mention quant au caractère nominatif ou confidentiel du document (ex. dans les métadonnées), il pourrait être automatisé, en partie, par l'IA. Sur la base de leur exploration d'un ensemble d'articles théoriques en archivistique computationnelle et en informatique, Lemieux et Werner (2024) voient que l'IA pourrait être exploitée à des fins de renforcement de la gestion des risques associés à ce type de données. Ils mettent de l'avant le potentiel des technologies d'amélioration de la confidentialité (*Privacy Enhancing Technologies*), soutenues par l'IA, afin d'identifier les documents

nominatifs et confidentiels (Lemieux et Werner, 2024).

Néanmoins, si l'IA pourrait être exploitée pour la constitution des archives témoignant des activités du producteur, elle engendrerait un ensemble de risques. En fait, les algorithmes utilisés à des fins de reconnaissance du texte manuscrit n'ont pas encore atteint un degré élevé de précision, ce qui pourrait augmenter le taux du bruit dans les extrants archivistiques et engendrer des problèmes de pertinence (Meng et Ghena, 2023). Aussi, le recours total à l'IA à des fins de classification pourrait générer des rubriques classificatoires peu pertinentes par rapport à la portée des documents traités (Alaoui, 2023). Cela est principalement dû au fait que les algorithmes de classification automatique se basent sur l'analyse du contenu des documents et ne prennent pas en considération le contexte (Payne, 2022). Celui-ci, étant un élément central dans la pratique archivistique, est essentiel pour comprendre les conditions présidant la création des documents, et par conséquent, assure l'intelligibilité des documents d'un point de vue archivistique (Payne, 2022). Cela laisse entendre que la conception des algorithmes de l'IA à des fins de classification automatique gagnerait à intégrer davantage cet aspect contextuel pour améliorer le degré de précision au niveau des rubriques classificatoires proposées. Le même constat pourrait être fait pour l'indexation des documents en vue d'en améliorer le repérage. La suggestion des mots-clés par l'indexation automatique devrait être validée par l'archiviste en vue de s'assurer que les mots-clés sont pertinents et ne causent pas une surcharge d'information. La validation humaine aurait aussi pour finalité de vérifier que les extrants archivistiques ne renferment pas des données nominatives et confidentielles (Alaoui, 2023). Si l'identification de telles données pourrait être rendue aisée par les technologies de l'IA pour l'amélioration de la confidentialité (*Privacy enhancing technologies*), la performance actuelle de ces outils laisse à désirer (Lemieux et Werner, 2024). Le degré de précision des extrants archivistiques qui en sont générés remet en question leur validité (Lemieux et Werner, 2024). Enfin, la détection automatique des doublons des documents pourrait comporter un taux d'erreur (Rolan et al., 2019), et ce, par l'identification erronée de la similarité entre deux documents qui appartiennent à des unités administratives différentes. L'absence d'une évaluation d'un tel résultat par l'archiviste rendrait difficile l'exploration des logiques d'agrégation d'un même document par différents usagers, ce qui compromettrait la traçabilité des activités.

Toutes ces considérations laissent entendre que l'archiviste continuerait à exercer ses rôles de juriste par la validation continue des résultats générés par l'IA. Cependant, il serait appelé à acquérir des compétences d'informaticien pour participer au développement des algorithmes de l'IA à des fins de capture, de description, de classification et d'indexation des archives (Alaoui, 2023). Dans un objectif de favoriser une meilleure intégration de l'IA aux pratiques archivistiques, l'archiviste juriste-informaticien serait appelé à encourager

l'usage des schémas de métadonnées institutionnels pour assurer une meilleure capture du contexte des activités qui ont généré les documents (Alaoui, 2023). Ensuite, ces métadonnées pourraient être intégrées, avec les données massives extraites du contenu des corpus documentaires traités, aux données d'entraînement des algorithmes à des fins d'amélioration de la gestion des documents d'archives.

En résumé, l'automatisation soutiendrait la finalité de la preuve recherchée par l'archiviste dans le premier paradigme de la preuve, où les archives y sont perçues comme des traces documentaires neutres qui reflètent les activités du producteur. À l'ère de l'automatisation, elles feraient figure de données massives (Mordel, 2019) traitées par les algorithmes de l'IA. Les archives traitées par l'IA seraient mieux contextualisées, intelligibles et repérables, ce qui favoriserait leur exploitabilité. Cependant, compte tenu des limites de l'IA, l'archiviste continuerait à assumer son rôle de juriste mais développerait davantage ses compétences informatiques pour améliorer la performance des algorithmes de l'IA. Si le paradigme de la preuve met en lumière la nécessaire constitution des archives pour les besoins du producteur, le paradigme de la mémoire dépasse cette vision réductrice favorise une meilleure pluralisation des archives (Cook, 2013), soit leur diffusion à un segment plus large d'utilisateurs.

Mémoire et automatisation

Le développement du paradigme de la mémoire est engendré par les influences de modernisation de l'appareil administratif public et l'établissement des lois sur l'accès à l'information (Cook, 2013). Selon l'auteur, ces changements ont amené les organisations publiques à développer un devoir institutionnel en regard de la gestion de leurs archives, preuves de leurs activités. Toutefois, ces traces sont non seulement assimilées au niveau individuel, comme pièces, mais agrégées pour former ensemble une mémoire consignée (Cardin, 1995; Cook, 2013). Cela illustre le propre du paradigme de la *mémoire* : l'archiviste-historien y est amené à établir des critères d'évaluation des archives en se basant sur plusieurs aspects, dont les besoins exprimés par les acteurs organisationnels en regard de l'appropriation des documents (Cook, 2013). Or, la sélection des archives à intégrer à cette mémoire ne se limite pas aux besoins internes, c'est-à-dire ceux relatifs aux acteurs organisationnels, mais les dépasse pour considérer davantage les besoins des ayants droit, dont les citoyens, les journalistes, les chercheurs universitaires, les étudiants et les historiens (Cook, 2013). La rationalisation de la production documentaire en appliquant le principe de la gestion du cycle de vie documentaire, couplée au développement des activités de médiation culturelle et de diffusion des archives (ex. services de références des archives), constituent en elles-mêmes des manifestations de la réinvention du rôle de l'archiviste et sa relation avec ses usagers internes et externes (Cook, 2013). Ainsi, cet expert se soucie à la fois de la constitution

des traces des activités du producteur et leur diffusion aux acteurs organisationnels et aux ayants droit, tout en cernant leurs besoins en matière d'accès aux archives. Dans ce paradigme, l'archiviste prend non seulement en considération la dimension administrative des archives, mais aussi identitaire et culturelle (Cook, 2013), reflétant l'aspect historique et artéfactuel revêtu par ces traces.

Avec les développements technologiques, les moyens de création, de traitement et de diffusion de l'information se sont réinventés. La variété des dispositifs numériques soutenant la gestion des documents dans leurs différentes manifestations (ex. documents statiques, jeux de données, etc.) préfigure des logiques différentes du traitement de celle-ci. En guise d'exemple, les bases de données n'intègrent pas les mêmes logiques de gestion de ces traces que pour les systèmes de gestion des documents traditionnels (*Electronic Document and Records management systems*). Tel qu'avancé dans la discussion du paradigme de la preuve, l'automatisation soutiendrait l'archiviste dans sa démarche de constitution des preuves documentaires des activités organisationnelles, notamment en automatisant en partie certains processus archivistiques. Dans le paradigme de la mémoire, l'emphase est mise sur la rationalisation de la gestion des archives. Cela s'opère par l'application du principe de gestion du cycle de vie et des critères de l'évaluation de ces traces, des processus qui pourraient être automatisés grâce à l'IA. Grâce à l'extraction des mots-clés les plus significatifs dans les documents d'archives, il serait possible de les évaluer pour en déterminer ceux renfermant une valeur primaire ou secondaire pour l'organisation et ses usagers (Colavizza et al., 2021; Makhoul-Shabou et al., 2020). En prenant appui sur des résultats d'un ensemble d'études de cas, certains auteurs affirment que l'évaluation automatique rendrait plus aisé le processus de l'élaboration des règles de conservation des documents (Colavizza et al., 2021; Makhoul-Shabou et al., 2020). Dans un système de gestion documentaire, celles-ci sont associées aux rubriques classificatoires des documents, puisque les délais de conservation sont définis par type de documents (Alaoui, 2023). Ainsi, puisque l'IA soutiendrait la classification automatique des documents, il en irait de même pour la gestion de leur cycle de vie (Alaoui, 2023). Cependant, de tels processus automatisés nécessiteraient une validation humaine. L'archiviste-historien serait appelé à évaluer le degré de précision des résultats générés par l'IA. Cela aiderait aussi à minimiser les risques de destruction accidentelle des documents, laquelle pourrait compromettre le devoir de transparence et de reddition des comptes auquel les organisations actuelles sont soumises.

Un autre exemple de l'usage de l'IA à des fins de constitution de la mémoire serait la reconnaissance du texte manuscrit (Colavizza et al., 2021; Terras, 2022). Tel qu'avancé dans le paradigme de la preuve, cela rendrait plus exploitables les manuscrits grâce à leur transcription (Colavizza et al., 2021; Terras, 2022). La reconnaissance du texte manuscrit sou-

tiendrait dès lors le travail de l'archiviste-historien, soucieux de la constitution et la valorisation des anciens manuscrits comme objets artéfactuels chargés de significations culturelles et identitaires. Plus précisément, elle favoriserait une meilleure agrégation de ces archives, caractérisées par leur aspect d'unicité et d'exemplarité, à une mémoire historique qui pourrait être diffusée auprès des usagers externes, dont les historiens et les citoyens amateurs. Or, le problème de précision des résultats générés par l'IA est soulevé par certains auteurs qui insistent sur l'importance de l'intervention humaine pour la vérification de la qualité des extrants (Rolan et al., 2018).

Par ailleurs, dans le paradigme de la mémoire, l'archiviste se soucie de la constitution, la diffusion et la valorisation de la mémoire pour les besoins des usagers internes et externes. Avec la tendance des institutions culturelles à diffuser en ligne une partie du patrimoine documentaire pour rejoindre des segments plus larges d'usagers, les archivistes ont d'ores et déjà revisité leurs rôles dans la diffusion des archives et la médiation documentaire avec les usagers internes et externes. Grâce à l'organisation des expositions virtuelles et à la mise en ligne des archives sur les médias sociaux, ou encore leur promotion sur les sites web officiels des institutions culturelles, les usagers ont accès à un large éventail des documents qui peuvent répondre différemment à leurs préoccupations. Or, considérant le fait que ces usagers ont des profils hétérogènes qui influent sur leurs besoins d'information et la manière dont ils consomment les archives, l'archiviste serait appelé à adopter des stratégies sophistiquées pour passer du rôle réactif au rôle proactif dans l'identification et l'analyse des besoins exprimés par ces usagers. Un moyen d'y arriver serait de transposer des stratégies utilisées en marketing pour réaliser une segmentation des usagers en fonction de leurs profils et besoins. Cela aiderait aussi cet expert à ajuster les critères d'évaluation et de sélection des archives à conserver d'une manière permanente pour ainsi enrichir la mémoire.

En marketing, une des stratégies les plus populaires mises à profit à des fins de conception de nouveaux produits et services est celle de l'utilisation des *personas*. Ces derniers sont des personnages fictifs avec des caractéristiques typiques et communes, telles que les propriétés sociodémographiques (ex. âge, sexe, position sociale), les traits comportementaux (ex. interaction avec les produits, besoins) et les compétences (Miaskiewicz et Kozar, 2011). Ces *personas* aident à établir une segmentation des cibles en fonction de ces caractéristiques, afin de mieux y adapter l'offre des produits et services en conséquence. Il s'agit dès lors d'une stratégie qui favorise non seulement une meilleure catégorisation des clientèles cibles, mais aussi une anticipation de leurs besoins. En sciences de l'information, le milieu des bibliothèques a déjà opté pour la transposition de ces stratégies de marketing pour segmenter les usagers, comprendre leurs besoins et adapter l'offre culturelle en conséquence (Zini, 2020). Cela n'est pas

surprenant, puisque les bibliothécaires ont une longue tradition en matière d'étude des usagers et de leurs attentes, ainsi que le développement des stratégies de médiation documentaire avec eux, et ce, en utilisant des techniques et stratégies propres au domaine du marketing (Garoufailou et al., 2013; Gupta, 2015; Gupta et Savard, 2011). Le marketing est également abordé par certains auteurs pour mieux connaître les usagers et les segmenter en fonction de leurs besoins et habitudes (Cardin et Desilets, 2017; Yeo, 2005). Cependant, la littérature sur ce sujet s'avère moins abondante qu'en bibliothéconomie.

Un des outils que les spécialistes du marketing utilisent dans le cadre de l'étude des besoins des usagers et pour développer ces *personas* est le système de traitement des données d'interaction avec l'information en ligne. En guise d'exemple, *Google Analytics* se veut un dispositif technologique qui offre un tableau de bord avec des données traitées qui offrent une vue d'ensemble de l'interaction des internautes avec l'information diffusée en ligne, par exemple sur les sites web institutionnels, ou encore sur les médias sociaux (ex. *Facebook*, *Twitter*, *Instagram*). Pour ce faire, ce dispositif se fonde sur la collecte des données de navigation des usagers (ex. localisation géographique, tranche d'âge, sexe, temps passé sur une page web, type de mots-clés utilisés, type de contenu avec lequel on interagit le plus, etc.). Ces données de navigation, ou encore de traçage, documentent le processus d'interaction des usagers avec les informations diffusées en ligne, ce qui aide les spécialistes en marketing à se construire une idée plus claire et globale des intérêts des clientèles cibles et les catégoriser en fonction des *personas*. Les systèmes de traitement de telles données, comme c'est le cas pour *Google Analytics*, se veulent des outils d'aide à la décision fondée sur une données probantes, puisqu'il offre des tableaux de bord avec des indicateurs de performance, permettant de générer des rapports d'audience et segmenter celle-ci en fonction des paramètres initialement définis.

L'usage de ces données de navigation en ligne à des fins de segmentation des usagers pourrait soutenir le rôle des archivistes dans l'évaluation et la sélection des archives à portée historique et patrimoniale pour enrichir la mémoire et, par conséquent, diffuser ces traces documentaires aux usagers qui s'y intéressent. Le nombre de visites des pages web et celui de clics sur les archives numérisées ou nativement numériques mises en ligne pourraient être révélateurs du taux d'engagement des usagers avec les objets documentaires diffusés sur les plateformes numériques. Ces données seraient encore plus pertinentes lorsqu'elles sont croisées avec d'autres variables telles que l'âge, le sexe, la localisation géographique et les centres d'intérêts respectifs de chaque usager. Un tel croisement permettrait d'en tirer des significations faisant en sorte que les archivistes soient en mesure de mieux connaître les attentes des usagers, voire les anticiper, lors de l'établissement des valeurs renfermées par les archives et leur diffusion aux

clientèles cibles.

À l'ère de l'automatisation, les dispositifs de l'IA pourraient soutenir les archivistes dans la constitution et la diffusion de la mémoire. D'une part, ces données de navigation en ligne, reflétant les comportements des usagers en matière de consommation des archives numériques, pourraient être mobilisées à des fins d'automatisation des processus d'évaluation de ces traces documentaires par l'archiviste. Certes, l'archiviste se base sur plusieurs critères lors de la sélection des archives à conserver d'une manière permanente, ainsi que sur son intuition et sa subjectivité (Cook, 2011). L'exploitation de ces données des usagers favoriserait une meilleure inclusion des perspectives de ces derniers dans le processus d'enrichissement de la mémoire par des archives renfermant des valeurs en elles-mêmes. Aussi, ces données de navigation pourraient être traitées par des outils de l'IA pour assurer une diffusion personnalisée des archives, et ce, en fonction des besoins de chaque usager. Si l'on revient à l'univers du marketing, les professionnels de la publicité numérique ont d'ores et déjà opté pour l'intégration de l'IA à des fins de pratique du marketing relationnel qui renvoie à la diffusion du contenu promotionnel selon les intérêts de chaque usager, et ce, en fonction de ses données de navigation en ligne (Ray et Sabadie, 2016). Similairement, dans le milieu des bibliothèques, l'intégration des systèmes de recommandation aux catalogues favorise une meilleure découvrabilité des ressources documentaires (Gupta et Pandey, 2019), vu qu'il est question d'une diffusion personnalisée de l'information sur la base des données de traçage traitées par ces systèmes. L'usage de l'IA à des fins de recommandation des archives en fonction de ces données de navigation favoriserait une meilleure proactivité des archivistes dans la médiation documentaire auprès des usagers, puisque ces derniers auraient accès à un large éventail d'archives qui s'insèrent mieux dans leurs schèmes de pratiques respectifs. L'usage de l'IA pour prédire les types des archives auxquelles les usagers s'intéresseraient n'est toutefois pas abordé dans la littérature scientifique examinée. Le potentiel décrit ici demeure théorique et reste à explorer en pratique.

Cependant, la mise à profit de l'IA dans cette optique comporterait certaines limites. Si l'archiviste, en mobilisant ses aptitudes d'informaticien, pourrait s'impliquer dans le développement des algorithmes de recommandation des archives, une telle conception se base sur des données d'entraînement faisant l'objet de choix subjectifs. Qui plus est, considérant la variété des besoins des usagers, leurs compétences, attentes, valeurs personnelles et cadres de référence culturels, il serait difficile de prédire, dans l'absolu, les besoins de ces clientèles en matière d'accès aux archives. Les chercheurs universitaires sont issus de différentes disciplines, ce qui influencerait la portée des archives qu'ils consomment. Les journalistes couvrent, dans leur quotidien, plusieurs facettes de la vie sociale, ce qui multiplierait leurs besoins. Ainsi,

la médiation des archivistes et l'utilisation des services de référence des archives seraient encore essentielles pour offrir aux usagers une meilleure expérience avec les archives auxquelles ils s'intéressent. Un autre enjeu qui émergerait dans ce contexte serait de nature éthique et aurait trait à la collecte des données de navigation en ligne des usagers à des fins de marketing des archives. Ces données font figure de données personnelles, dont la collecte doit faire l'objet d'un consentement préalable de chaque usager, et ce, après l'avoir informé de la manière dont ses données sont collectées et des fins pour lesquelles elles seront utilisées.

Toutes ces considérations laissent entendre que l'IA ne pourrait pas remplacer le rôle de l'archiviste, mais elle soutiendrait plutôt ce spécialiste dans l'exécution de ses tâches. L'archiviste serait amené à incarner, à la fois, le profil d'un historien et d'un informaticien pour, d'une part, identifier les requis pour entraîner les algorithmes en vue d'améliorer la précision des extrants. D'autre part, la conformité aux lois et aux règlements serait une obligation pour conserver les documents essentiels et respecter la vie privée des usagers qui interagissent avec les plateformes en ligne. Qui plus est, l'intervention humaine serait encore nécessaire pour comprendre les besoins des usagers et y adapter l'offre culturelle, car l'évaluation devrait tenir compte à la fois des attentes des usagers internes et externes en matière de diffusion en ligne des archives.

Somme toute, les paradigmes de la mémoire et de la preuve, soutenus par l'IA, redéfiniraient les rôles de l'archiviste dans la constitution, le traitement, l'évaluation, la diffusion et la préservation à long terme des archives. Cela réinventerait son image en tant qu'expert de la gestion des archives, autrement dit, son identité socioprofessionnelle.

Identité et automatisation

Dans le troisième paradigme, il est question d'une quête d'identité socioprofessionnelle de l'archiviste en tant qu'expert de gestion documentaire et gardien de confiance de la mémoire. Il se veut un médiateur entre le passé et le présent, un spécialiste qui fait de la gestion des archives un moyen pour suggérer des solutions aux différents enjeux de la société (Cook, 2013). Comme le souligne Cook (2013), ce paradigme reflète l'ensemble des efforts déployés par les archivistes qui s'allient dans des communautés de pratique, de regroupements de pairs, ainsi que des associations professionnelles et académiques, afin de formaliser les programmes de formation et en recherche en archivistique. À cela s'ajoute la création des revues scientifiques en archivistique (Cook, 2013), lesquelles font figure d'espace de médiation scientifique et d'échange de savoirs théoriques et pratiques entre les communautés de chercheurs, d'étudiants et de professionnels en archivistique. Qui plus est, ce paradigme traduit le désir de la communauté archivistique de tisser des liens de collaboration avec d'autres experts des domaines connexes. Il est question de celui de la

gestion de l'information et des archives historiques, respectivement, les experts en TI, les muséologues, les historiens et les généalogistes (Cook, 2013). On évoque également le développement des politiques, des normes internationales et des guides pratiques établis sur une base consensuelle pour harmoniser les pratiques et valoriser l'apport de la gestion des archives à la société (Cook, 2013). Enfin, Cook (2013) souligne la tension liée à la technologie et met en lumière le rôle de l'archiviste sur la scène numérique. Somme toute, il est question d'une quête d'identité socioprofessionnelle qui permet aux archivistes de s'affirmer en tant qu'experts, de fédérer leur expertise avec d'autres spécialistes, toutefois en défendant leur territoire d'autorité que la logique des rapports entretenus avec les documents d'archives leur impose.

Dans un contexte documentaire marqué par l'introduction de l'automatisation à des fins d'optimisation de la gestion des archives, les archivistes verraient leur identité se redéfinir. Compte tenu des particularités des fonctionnalités de l'IA et de la manière dont elle soutiendrait les missions de ces spécialistes, ces experts seraient amenés à réinventer leur identité socioprofessionnelle. La collaboration avec d'autres spécialistes des domaines connexes, tels que la science des données, l'informatique, le droit, la sociologie et la paléographie serait également envisagée. Cette collaboration, impliquant l'archiviste en tant que médiateur, intégrerait aussi ses rôles comme juriste, historien et informaticien. Elle aurait pour finalité d'encadrer l'IA et d'améliorer l'usage responsable, éthique et efficace des dispositifs de l'IA à des fins de gestion des archives. Un autre rôle qui redéfinirait l'identité socioprofessionnelle de l'archiviste serait celui de la formation, c'est-à-dire former une relève de professionnels capables de faire un bon usage de l'IA. À cela s'ajouterait celui de la contribution à la redéfinition des chantiers de la recherche scientifique.

Dans le paradigme de l'identité, l'archiviste assumerait un rôle d'informaticien, mais couplé à celui du juriste et d'historien. Dans un objectif de favoriser une meilleure intégration de l'automatisation aux pratiques archivistiques, des compétences en matière de l'usage des dispositifs de l'IA seraient requises pour ce spécialiste. La compréhension des logiques du fonctionnement de l'apprentissage machine (*machine learning*) et la distinction des caractéristiques de celui-ci (supervisé, semi-supervisé et non supervisé) seraient des compétences à acquérir par l'archiviste. Il serait aussi question de celles liées au développement et à l'usage des algorithmes de l'IA en se basant sur des principes de la probabilité appliquée. L'archiviste, jouissant d'aptitudes transversales, jouerait un rôle d'évaluation de la conformité des extraits archivistiques en regard des principes reconnus par la communauté scientifique et professionnelle (Alaoui, 2023; Kaldeli et al., 2021; Rolan et al., 2018). Il est question du principe de provenance, du cycle de vie et de classification archivistique en maintenant les liens organiques. À cela s'ajouteraient d'autres interven-

tions qui ne sont pas soulignées dans la littérature examinée. La validation des extraits générés par l'IA devrait être documentée pour capitaliser l'expertise et la transférer à ses pairs. Plus précisément, la tenue d'un registre des opérations assistées par l'IA serait nécessaire pour justifier, par exemple, pourquoi certains documents ont été détruits après l'application des étiquettes de rétention d'une manière automatique. Elle aiderait aussi à responsabiliser, dans une optique d'imputabilité, les acteurs en jeu (ex. dirigeants et employés utilisant un système de gestion documentaire) afin d'évaluer leur degré d'intervention sur les processus archivistiques automatisés. Autrement dit, il serait question de s'assurer que les extraits générés font l'objet d'une vérification humaine, ce qui aiderait à baliser l'agentivité de la machine dans cette optique. Cependant, l'archiviste ne pourrait pas travailler seul, mais il solliciterait la collaboration d'un ensemble d'experts en provenance des domaines connexes pour pallier les enjeux informationnels, légaux et technologiques engendrés par l'IA.

Certains chercheurs reconnaissent la nécessaire collaboration de l'archiviste avec d'autres spécialistes dans l'optique de l'automatisation. Dans le cadre de la gestion des archives historiques, la collaboration avec les paléographes serait bénéfique pour valider la qualité de la transcription faite des manuscrits par la reconnaissance du texte manuscrit (Hodel, 2021). Il serait aussi question de la définition des jeux de données à inclure dans l'entraînement des algorithmes à des fins d'amélioration de leur degré de précision pour la reconnaissance du texte manuscrit (Colavizza et al., 2021; Hodel, 2021). Sur le plan technologique, dans sa conceptualisation de la gestion documentaire assistée par l'IA, Alaoui (2023) voit que l'archiviste serait appelé à collaborer avec un ensemble d'acteurs organisationnels. Il serait question de la fédération de l'expertise archivistique avec celle des développeurs des algorithmes. Une première piste à considérer serait celle du développement des compétences numériques. En fait, il serait faux de tenir pour acquis que les aptitudes de l'archiviste favoriseraient toujours la conception, la mise en œuvre et l'appropriation des outils de l'IA. Les développeurs pourraient jouer un rôle de formateurs auprès des archivistes pour les accompagner dans leur démarche de compréhension et d'usage de ces outils. Ensuite, les archivistes, se dotant des aptitudes d'informaticiens, participeraient, de concert avec les développeurs, à la définition du périmètre des données à inclure dans les algorithmes à des fins de traitement automatique des archives (Alaoui, 2023; Colavizza et al., 2021; Hodel, 2021). Considérant le temps requis pour l'entraînement des algorithmes, leur test sur un échantillon des corpus documentaires et l'évaluation de leur performance, une définition d'un plan pour l'amélioration continue des algorithmes serait une piste à explorer. L'usage des tableaux de bord avec des indicateurs de performance, par l'archiviste et les développeurs des algorithmes, serait également bénéfique afin d'instaurer une culture d'amélioration continue de la qualité

des algorithmes. Cela aiderait, ainsi, à augmenter le degré de précision des extraits archivistiques générés (ex. mots-clés, rubriques classificatoires, transcription, identification des documents nominatifs et confidentiels, la définition de la valeur archivistique des documents et la gestion du cycle de vie).

Sur le plan éthique et légal, la fédération de l'expertise de l'archiviste et des spécialistes en droit s'avèrerait une piste de solution face à l'inquiétude exprimée par plusieurs chercheurs en archivistique quant à l'usage des dispositifs de l'IA. Il est plus spécifiquement question des enjeux qu'il présente sur le plan de la protection de la vie privée (Colavizza et al., 2021; Jaillant et Caputo, 2022; Tzouganatou, 2021). Tel que précédemment évoqué, les fonctionnalités de l'IA pourraient être bénéfiques pour suggérer des archives en fonction des préoccupations des usagers, et ce, comme c'est le cas pour les systèmes de recommandation dans les bibliothèques et les pratiques du marketing relationnel chez les communicateurs. Or, le fonctionnement de ces algorithmes exigerait le traitement des données de navigation en ligne des usagers, renvoyant, entre autres, à leurs profils sociodémographiques et leur localisation. Par conséquent, la collaboration avec les spécialistes du droit serait bénéfique, puisque ces derniers se veulent les mieux placés pour se prononcer sur tels sujets éthiques. Cela aiderait à mieux prouver la légitimité en regard des lois en vigueur, dont la *Loi sur la protection des renseignements personnels* (L.R.C., ch. P-21) et la *Loi sur l'intelligence artificielle et les données* (LIAD).

Un autre point à considérer serait la nécessaire minimisation des biais engendrés par le traitement des archives par des dispositifs d'IA. Le développement et l'entraînement des algorithmes de l'IA sont faits par les développeurs en fonction de leurs propres convictions, ce qui peut générer des biais (Jaillant et Caputo, 2022; Tzouganatou, 2021). À ce sujet, Tzouganatou (2021) voit que les données traitées par la machine renferment des biais qui influencent la performance de celle-ci, puisqu'elle est entraînée par des êtres humains qui possèdent leurs propres convictions et schèmes sociaux (Tzouganatou, 2021). Par conséquent, il est important d'évaluer les risques des biais qui pourraient être engendrés par le recours de l'IA à des fins de traitement des archives. Il est question, entre autres, des photographies qui représentent les différentes classes sociales, les groupes ethniques et les genres des personnes. La collaboration des archivistes avec les experts en droit, voire avec les sociologues et les représentants des usagers, se révélerait vitale afin d'aboutir à des solutions qui favorisent une utilisation rationnelle de l'IA dans le traitement des archives, et ce, tout en respectant les valeurs liées à l'équité, à la diversité et à l'inclusion.

Par ailleurs, certains chercheurs soulignent le nécessaire encadrement de la manière dont l'IA s'intègre aux processus archivistiques (Colavizza et al., 2021; Jaillant et Caputo, 2022). Les propos avancés par ces auteurs sont articulés dans des discours qui encouragent une telle intervention mais,

en réalité, les dispositifs procéduraux et les outils normatifs sont manquants. Il n'existe pas de consensus sur la manière dont l'archiviste pourrait intervenir pour développer des outils d'encadrement de l'IA. Sur le plan générique, plusieurs cadres régissant l'utilisation de l'IA ont déjà commencé à se développer, tels que la *Loi sur l'intelligence artificielle et les données au Canada*, ou encore la *Déclaration de Montréal* pour une IA responsable. La portée de ces cadres converge vers un point commun, soit les enjeux éthiques et juridiques entourant l'IA, la question du remplacement de l'humain par la machine, la précision des extraits et les retombées sociales de l'IA sur les relations humaines.

Aussi, sur le plan normatif, la publication de la dernière norme internationale de l'*ISO 42000 :2023* sur la conception et le déploiement des systèmes de l'IA a pour but de normaliser les pratiques dans cette optique. Elle répond aussi à des préoccupations concernant le respect de la vie privée, la nécessaire supervision humaine, l'amélioration continue et la gestion des risques. Une telle norme est teintée d'une dimension organisationnelle, puisqu'elle responsabilise les acteurs organisationnels désireux de mettre à profit les dispositifs de l'IA pour réinventer leurs modes de travail. Dans cette lignée de pensée, Alaoui (2023) montre, dans sa réflexion théorique, comment la norme *ISO 24143* sur la gouvernance de l'information pourrait être utilisée comme outil de pilotage des pratiques de gestion documentaire assistée par l'IA. Alors que ses pistes semblent pertinentes, elles gagneraient à faire l'objet d'expérimentation pour en évaluer la faisabilité.

En clair, la communauté scientifique et professionnelle ne s'entend pas encore sur l'appropriation de ces outils d'encadrement en archivistique. La disposition de mécanismes de régulation et d'encadrement propres à la discipline et à la pratique archivistiques serait nécessaire pour comprendre l'usage de l'IA dans les pratiques de gestion des archives. La portée des outils à développer en archivistique pourrait, ensuite, faire l'objet d'une opérationnalisation sous forme de politiques et des procédures organisationnelles adaptées à la réalité de chaque organisation.

Dans le paradigme de l'identité, le rôle de l'archiviste dépasse celui de la fédération de l'expertise avec d'autres spécialistes pour inclure également celui de la formation (Cook, 2013). Il serait appelé à agir en proactivité pour développer des programmes de formation continue, de certifications reconnues par les associations professionnelles, des référentiels de compétences requises à l'ère de l'automatisation, comme on le voit d'ailleurs chez les associations professionnelles en archivistique. La formation d'une relève qui serait capable, elle-même, de former de futurs archivistes qui jouissent des aptitudes recherchées à l'ère de l'automatisation serait aussi d'importance égale. Il est digne de mentionner qu'il ne s'agit pas de nouveaux rôles pour l'archiviste dans la formation, que cela soit dans un contexte universitaire que professionnel, puisque la littérature scientifique est riche en articles

qui mettent de l'avant les rôles de formation, d'éducation joués par l'archiviste dans l'optique du transfert de son savoir théorique et pratique (Carini, 2009; Vassilakaki et Moniarou-Papaconstantinou, 2017).

Puisque l'IA apporte son lot d'enjeux pour la pratique archivistique, que cela soit sur le plan éthique, normatif ou encore social, force est d'admettre que les programmes de formation et de perfectionnement gagneraient à être révisés. Cela aurait pour objectif de développer une conscience collective à l'égard de ces enjeux, et de suggérer des pistes de solutions pertinentes pour les pallier. Proactivité, collaboration, gestion du changement, communication, médiation culturelle, seraient toutes des qualités dont l'archiviste jouirait pour réaffirmer son rôle en les intégrant dans ses programmes de formation. Cela se justifie par le fait que ceux-ci constituent le miroir de l'évolution de l'identité archivistique au fil du temps (Ngoepe, Jacobs et Mojapelo, 2022; White et Gilliland, 2011). Les compétences que l'archiviste chercherait à développer chez les futures générations seraient transversales. Elles devraient porter sur la manière dont l'IA pourrait être mise à profit dans le cadre des pratiques archivistiques pour soutenir les finalités probatoires et mémorielles de la gestion des archives. Autrement dit, les programmes de formation ne devraient pas exclure les rôles traditionnels de l'archiviste, mais les inclure davantage dans le nouveau profil de l'archiviste, celui de l'informaticien. Comme le dénote Marciano et ses collaborateurs (2018), le développement des programmes de formation sur l'intégration de l'IA aux pratiques archivistiques nécessiterait une collaboration multidisciplinaire entre plusieurs chercheurs. Celle-ci interpellerait les intervenants en sciences des données, en informatique, en sciences historiques, en sociologie, en droit et celles de la gestion.

Enfin, l'archiviste serait amené à contribuer à la définition des chantiers de recherche scientifique en archivistique. En fait, une des particularités de la recherche en archivistique réside dans le fait qu'elle met en dialogue la théorie et la pratique. La recherche fondamentale et appliquée se nourrissent mutuellement et se développent au gré des avancées technologiques et des tendances socioculturelles. Compte tenu de cette particularité, l'archiviste contribuerait, de concert avec les chercheurs universitaires, à la définition des chantiers de recherche scientifique en lien avec l'automatisation et l'intégration des mécanismes de l'IA à la pratique archivistique. À ce sujet, les conférences et les communications présentées dans le cadre des congrès, des colloques, des symposiums et des journées d'études font figure d'une arène de médiation des savoirs entre les théoriciens et les praticiens en archivistique. Tous ces moyens constitueraient une occasion pour suggérer des pistes de recherche scientifique futures, notamment par les professionnels qui font usage des dispositifs de l'IA à des fins d'automatisation du traitement des archives. Cela ferait en sorte que les chercheurs mobilisent le savoir théorique et la rigueur méthodologique à des fins de contribution à

l'avancement des connaissances en archivistique, lesquelles revêtraient une valeur ajoutée à la fois théorique et pratique.

En clair, le paradigme de l'identité, incarnant le profil de l'archiviste comme médiateur, se redéfinirait avec l'usage de plus en plus prépondérant des dispositifs de l'IA à des fins de traitement et de diffusion des archives. En tant que spécialiste de la gestion de l'information, l'archiviste serait amené à exercer plusieurs rôles et à développer son profil socioprofessionnel pour jouir d'une reconnaissance auprès des experts qui seraient, eux aussi, impliqués dans la conception et l'évaluation des dispositifs de l'IA dans plusieurs contextes. L'archiviste s'impliquerait en outre auprès de ses usagers pour évaluer leur perception de l'utilité de l'IA, les informer sur les bénéfices de l'automatisation et les sensibiliser à son usage responsable, éthique et rationnel. Or, notre exposé montre que la redéfinition de l'identité socioprofessionnelle de l'archiviste à l'ère de l'automatisation est encore en cours. Certains aspects ne font pas encore l'objet d'un consensus scientifique et professionnel, comme la portée des politiques, des règlements et des normes qui encadreraient l'intégration de l'IA à la gestion des documents administratifs et des archives historiques.

L'automatisation améliorerait la socialisation de l'archiviste avec plusieurs parties prenantes, puisqu'il chercherait constamment à développer des stratégies de dialogue et de collaboration avec elles. Celle-ci dépasserait les frontières institutionnelles pour inclure d'autres acteurs qui forment des groupes sociaux partageant des significations communes, incarnant ainsi le paradigme de la communauté.

Communauté et automatisation

Le dernier paradigme suggéré par Cook (2013) dans son article sur l'évolution des paradigmes archivistiques est celui de la communauté, renvoyant au rôle de l'archiviste comme activiste. Sous l'influence du postmodernisme (Huvila, 2008), il marque une évolution dans le rôle de l'archiviste en l'orientant vers plus de socialisation, puisque ce dernier s'implique de plus en plus dans des activités et des initiatives qui dépassent les frontières institutionnelles (Alaoui, 2021). Cette implication, teintée d'une touche plus sociale, consiste à prendre part active au sein des communautés socioculturelles qui existent dans une société donnée, dans un objectif de réduire les « silences » (Cook, 2013). Elle soutient en outre la représentativité de ces communautés à travers les archives détenues par les institutions culturelles. Elle s'inscrit dans l'optique des archives communautaires (*community archives*), incarnant des pratiques autonomes de la création et la gestion des archives des communautés (Flinn et Sexton, 2019). Autrement dit, les regroupements socioculturels constituent, par eux-mêmes, leur mémoire culturelle, et ce, par l'agrégat des traces documentaires qui reflètent leurs activités (Flinn et Sexton, 2019; Piggott, 2018; Welland et Cossham, 2019). Ces communautés estiment qu'il n'est pas nécessaire de solliciter l'archiviste pour former leur mémoire, puisqu'elles se voient

les mieux placées pour constituer, décrire et valoriser leurs traces documentaires qui reflètent leur vécu. L'implication de l'archiviste auprès de ces communautés s'inscrit dans un souci d'enrichissement du patrimoine archivistique national, en veillant à la représentation archivistique de l'ensemble des identités qui peuvent exister au sein d'une même société.

Par ailleurs, le paradigme de la communauté encourage la mise à profit des fonctionnalités de socialisation offertes par les technologies du web 2.0 à des fins de rapprochement entre ces communautés et les archivistes (Cook, 2013). Cela renvoie au concept des archives participatives (*participatory archives*) (Benoit et Eveleigh, 2019; Eveleigh, 2017; Huvila, 2008). Qu'il s'agisse de l'indexation sociale des archives, leur transcription, leur traduction ou encore leur numérisation, ces formes de production participative (*crowdsourcing*) misent sur la sollicitation de l'expertise de ces communautés dans la valorisation du patrimoine archivistique relatant leurs expériences vécues (Flinn et Sexton, 2019; Wright de Hernandez, s.d.). La diversité des points de vue des membres de ces communautés favorise le développement d'une multitude de perspectives et par conséquent, une variété des significations tirées des archives, et qui peuvent être mises à profit pour la valorisation de celles-ci. Cela peut illustrer la stimulation de la valeur artéfactuelle, émotive, historique ou encore patrimoniale de ces archives (Poole, 2019). Les pratiques des archives participatives concernent non seulement la sollicitation de ces communautés, mais les dépassent pour impliquer davantage d'utilisateurs à profils hétérogènes, tels que les historiens, les artistes, les généalogistes et les citoyens amateurs. Ces utilisateurs s'impliquent dans la stimulation des différentes valeurs renfermées par les archives qui les intéressent. En clair, ces pratiques induisent une délégation d'une partie des tâches communément exécutées par l'archiviste à une variété d'utilisateurs dans une optique de favoriser la pluralisation des archives et leur valorisation collaboratives, et ce, sous la supervision de l'archiviste. Les extraits générés par les utilisateurs dans ce contexte sont revus par cet expert pour en vérifier la conformité en regard des exigences en vigueur entourant la pratique archivistique (ex. qualité des métadonnées, description, transcription, qualité de la numérisation, etc.).

Comment l'automatisation redéfinirait-elle le déploiement de ces pratiques participatives dans le paradigme de la communauté? Comment le rôle de l'archiviste se réinventerait-il dans ce contexte? Les membres des communautés seraient-ils amenés à exécuter les mêmes tâches à l'ère de l'automatisation? Si oui, jusqu'à quel point? Quel serait le rôle de l'archiviste dans ce paradigme? Poursuivrait-il son rôle d'activiste? Serait-il plutôt remplacé par celui de l'archiviste-informaticien? Nous avançons que le rôle de l'archiviste, incarnant à la fois celui de l'informaticien et de l'activiste, consisterait à distinguer ce qui serait automatisable de ce qui ne le serait pas. Cela impliquerait une réingénierie des logiques de conduite des projets des archives participatives

afin de déléguer les tâches non automatisables aux bénévoles. Pour ce faire, un ensemble d'aspects archivistiques pourrait être défini afin d'établir une telle distinction. Il est question de (1) la *contextualisation*, (2) l'*émotion*, (3) les *liens identifiants*, (4) la *valeur patrimoniale*, (5) la *valeur historique*, (6) l'*intelligibilité*.

Un des bénéfices des archives participatives, notamment de l'indexation sociale, est celui de l'enrichissement de la description archivistique par des repères contextuels (Poole, 2019). Il peut s'agir d'étiquettes renvoyant à des aspects historiques, temporels, géographiques, thématiques, ou encore identitaires. La diversité des étiquettes attribuées aux documents d'archives mis en ligne sur les plateformes participatives reflète les multiples points de vue des utilisateurs et leurs perceptions. En fait, à l'ère du numérique, un même document d'archives peut renfermer de multiples significations en fonction de sa logique d'agrégation aux autres objets documentaires. La capture de ces étiquettes et leur association aux archives numériques se veulent ainsi une trace de la manière dont celles-ci sont agrégées les unes avec les autres. Cela améliore ainsi la découvrabilité de ces archives et leur potentiel d'exploitabilité par un segment plus large des utilisateurs, ce qui stimule la visibilité du patrimoine documentaire détenu par les Archives sur la scène numérique.

Afin de tirer le plein profit de ces bénéfices, l'intervention humaine serait nécessaire. Certes, l'IA pourrait traiter le contenu des corpus documentaires, que ceux-ci se manifestent sous forme de textes, d'images ou de sons (Kaldeli et al., 2021). Les fonctionnalités de l'IA sont évolutives, puisqu'elles permettraient l'extraction des données à partir de corpus documentaires textuels, ainsi que la reconnaissance des personnes, des endroits, des manifestations et des époques historiques représentées à travers les photographies. Cela constituerait un avantage qui aiderait les Archives à optimiser leurs ressources et consacrer leurs efforts à des tâches plus complexes. Cependant, comme le montre l'étude de Kaldeli et ses collaborateurs (2021), ces fonctionnalités se focalisent sur le contenu du document d'archives lui-même (Kaldeli et al., 2021). Or, en archivistique, un document d'archives résulte d'un contenu, d'un contenant et d'un contexte (Cardin, 2013-2014). Le contexte de création se veut essentiel, puisqu'il traduit les circonstances historiques, organisationnelles, sociales et personnelles ayant motivé la création des archives. Une IA ne serait pas capable de documenter les conditions de création des archives, puisque cela nécessiterait des témoins qui étaient impliqués dans les événements relatés par ces archives, ou encore qui jouissent d'une expertise en la matière (ex. recherche rétrospective, transmission de bouche à oreille entre les membres de la famille, les communautés, etc.). Aussi, cela commanderait le recours aux instruments de recherche et aux inventaires qui fournissent une description archivistique de ces traces, entre autres, les circonstances ayant présidé à celles-ci.

Dans la même lignée de pensée, la capture des émotions par les algorithmes de l'IA serait difficile. En fait, les archives sont des traces chargées d'affects, de symboles et de traits identitaires qui sont différemment perçus par les individus. Cela traduit les valeurs archivistes subjectives telles que la valeur émotive (Lee, 2016; Lemay et Klein, 2015; Palmer, 2010) et la valeur patrimoniale (Lessard, 2004). La définition de ces valeurs serait tributaire de la personnalité de chaque usager et de ses rapports avec les archives et les événements qu'elles relatent. La relation de médiation que chaque individu développe avec les archives est singulière, car elle retrace des événements qu'il a vécus, des liens personnels qu'il entretient avec sa famille, ses proches et ses amis. Elle incarne les souvenirs, beaux ou mauvais, qui se dégagent dans l'esprit individuel, et influence la manière dont les archives sont valorisées et consommées. La mise en valeur de ces aspects identitaires et émotionnels nécessiterait l'intervention des usagers, ce qui laisse entendre que les fonctionnalités de l'IA ne pourraient soutenir que partiellement l'indexation sociale des archives dans le cadre des pratiques participatives.

Malgré ces limites de l'IA, force est de constater qu'elle renfermerait un potentiel d'apprentissage automatique qui permettrait de détecter, en partie, les erreurs dans les extraits générés par les usagers dans le cadre des projets des archives participatives (Benoit et Eveleigh, 2019). La littérature scientifique rend compte des enjeux que les contenus générés par les usagers dans le contexte participatif représentent sur le plan de la qualité archiviste et de la pertinence des étiquettes attribuées aux archives en ligne (Poole, 2019). Les étiquettes générées peuvent nuire à la découvrabilité et la réparabilité des archives, ce qui pourrait aller à l'encontre des finalités recherchées par l'implication des usagers bénévoles. Par exemple, on signale le problème de la pertinence des étiquettes par rapport aux activités décrites par les archives historiques et les erreurs morphologiques (Poole, 2019). À cela s'ajoute la qualité linguistique de la transcription et de la traduction générées par les bénévoles à partir des contenus des manuscrits à lisibilité réduite. L'évaluation de la conformité de ces extraits et leur sélection en fonction de leur pertinence par les archivistes pourraient s'avérer coûteuses et laborieuses (Benoit et Eveleigh, 2019). En fait, cette opération de curation consiste à passer en revue l'ensemble des étiquettes et d'évaluer leur pertinence, et ce, avant de rendre les résultats de la description collaborative disponibles en ligne au grand public. Dans un objectif d'optimiser les ressources et réduire la complexité de telles tâches, le recours à des fonctionnalités de l'IA pourrait s'avérer une avenue intéressante à explorer pour détecter automatiquement les étiquettes non conformes à des critères préalablement définis par les archivistes (Benoit et Eveleigh, 2019). Plus précisément, les algorithmes basés sur l'apprentissage machine supervisé favoriseraient une meilleure détection des erreurs morphologiques pour améliorer la qualité des extraits archivistes générés par

les usagers bénévoles. Il serait question de développer des algorithmes avec des critères linguistiques qui permettraient d'identifier les erreurs syntaxiques et linguistiques pour soutenir les archivistes dans leur processus de curation de ces extraits.

Enfin, un autre avantage de mobiliser l'IA dans ce contexte participatif résiderait dans la compréhension des besoins des membres des communautés à l'égard de la portée des archives mises en ligne. Cela ferait en sorte que l'offre culturelle soit adaptée en conséquence, ce qui pourrait stimuler la motivation de ces usagers pour prendre part plus active dans ces projets. En fait, plusieurs auteurs soulignent dans leurs écrits scientifiques que la motivation constitue l'un des enjeux majeurs entourant ce genre de projets, de même que la fidélisation des bénévoles et leur engagement sur le long terme (Benoit et Eveleigh, 2019; Poole, 2019). Le recours à l'IA pourrait être bénéfique pour stimuler cette motivation, notamment en optant pour des techniques de prédiction et de suggestion des thématiques des archives qui susciteraient l'intérêt des usagers bénévoles. Ce type de fonctionnalités de l'IA, connue sous le libellé de « l'IA prédictive » qui repose sur l'apprentissage supervisé de la machine, se fonde sur le traitement d'un ensemble particulier des données afin d'établir des prédictions, suggérer des conseils ou des pistes d'action à des fins de prise de décision (Freire et al., 2019). Plusieurs auteurs en provenance d'une variété de disciplines scientifiques ont mis de l'avant les bénéfices de l'IA prédictive dans le soutien au processus de prise de décision, comme dans le domaine des sciences de la gestion (Jakubik et al., 2022). Il semblerait pertinent pour l'archiviste d'envisager le recours à ce type d'IA dans le cadre des projets des archives participatives. Cela constituerait une adoption des recommandations des auteurs qui posent que les archivistes devraient opter pour des démarches d'amélioration continue pour mieux fidéliser les usagers et stimuler leur motivation à participer aux projets (Benoit et Eveleigh, 2019; Poole, 2019). En se basant sur les données de navigation en ligne sur la plateforme participative et la portée des archives avec lesquelles chaque usager interagit, ainsi que les indicateurs comme le nombre de clics, les algorithmes de l'IA prédictive pourraient faire des prédictions sur la portée des archives qui susciteraient le plus les intérêts respectifs des usagers. Cela aiderait les archivistes à établir certaines priorisations et des cartographies des archives auxquelles les usagers bénévoles pourraient s'intéresser le plus.

En bref, à la mi-chemin entre le paradigme de la communauté et celui de l'automatisation, l'archiviste incarnerait à la fois le rôle de l'informaticien et celui de l'activiste. D'une part, il serait impliqué dans le développement des algorithmes à des fins de production participative. D'autre part, en tant qu'activiste, l'intervention de cet expert consisterait à valider la qualité des extraits archivistes traités par l'IA, notamment en regard de leur capacité à faire valoir les archives qui

témoignent des différentes identités qui existent dans la société. Bien que ces conclusions soient théoriques et manquent d'un véritable consensus de la communauté archivistique à cet égard, nous estimons que ces deux rôles de l'archiviste comme informaticien et activiste l'aideraient à déterminer ce qui serait automatisable de ce qui ne le serait pas. Il serait aussi question de l'amélioration continue de la performance des outils de l'IA, et ce, par l'identification des erreurs et la définition de la portée des données d'entraînement pour améliorer la performance des algorithmes.

Conclusion

Peut-on parler de l'automatisation comme cinquième paradigme archivistique? Il serait difficile d'avancer une réponse définitive, puisque plusieurs aspects sont en cours de développement. Nous avons abordé la manière dont l'automatisation, cristallisant un nouveau profil de l'archiviste, celui de l'archiviste-informaticien, et comment il complèterait les profils antérieurs cartographiés par Cook (2013). Cependant, nous ne pouvons pas emprunter une position totale sur un tel paradigme, car certaines dimensions théoriques et pratiques ne font pas encore l'objet d'un consensus. Il est question de l'encadrement par des lois, des normes et des politiques propres à l'archivistique, et qui gagneraient à faire l'objet d'un accord consensuel de la part de la communauté archivistique scientifique et professionnelle. L'approche marketing et son apport à la prédiction des archives demeure aussi inexplorée. En outre, l'usage de l'IA dans les projets de production participative des archives n'est pas suffisamment abordé. Les propos avancés par les chercheurs nous amènent enfin à conclure que l'IA ne peut pas, du moins actuellement, se substituer à l'intervention humaine, étant donné les enjeux informationnels, éthiques, légaux et sociaux qu'elle poserait pour l'archiviste.

L'automatisation fait naître de nouvelles représentations de l'objet documentaire et des approches de sa création, son traitement, sa diffusion et sa préservation à long terme. En fait, les dispositifs de l'IA traitent les archives comme intrants faisant figure de données massives, en vue de générer des extrants archivistiques décrits, classés, indexés et surtout découvrables et exploitables. Il reste à voir jusqu'à quel point l'archiviste réussira à jouer ses nouveaux rôles dans le paradigme de l'automatisation, notamment en instaurant un changement de culture et en révisant les principes et les outils d'encadrement à cette fin.

À propos de l'auteur

Siham Alaoui est candidate au doctorat en archivistique et communication publique à l'Université Laval, Québec (Canada). Elle est titulaire d'une maîtrise en sciences de l'information (obtenue en 2015 à l'Université de Montréal) et d'un baccalauréat en sciences de l'information (obtenu en 2013 à l'École des sciences de l'information, Maroc). Elle s'intéresse

à l'étude de la manière dont l'automatisation pourrait soutenir la gestion de l'information documentaire. Elle explore également les mécanismes de gouvernance d'information à déployer pour développer des avenues de collaboration entre les professionnels de l'information, notamment dans le contexte actuel de la transformation numérique des universités. Elle est auteure de plusieurs articles scientifiques et professionnels publiés dans des revues spécialisées en sciences de l'information (e.g. Archives, Canadian Journal of Information and Library Science, Documentation et Bibliothèques, Comma). Elle a également donné des communications dans le cadre des conférences et des colloques.

Références

- Alaoui, S. (2021). L'archive participative, les archivistes et les usagers: quels défis? Quelles pistes de solutions?/Participatory archiving, archivists and users: What are the challenges? What are the possible solutions?. *Canadian Journal of Information and Library Science*, 43(3), 217-244. <https://muse.jhu.edu/article/781386>
- Alaoui, S. (2023). L'intelligence artificielle et la gestion documentaire : quels apports? Quels enjeux? *The Canadian Journal of Information and Library Science*, 46(2). <https://doi.org/10.5206/cjils-rsib.v46i2.16695>
- Bak, G. (2012). Continuous Classification: Capturing Dynamic Relationships among Information Resources. *Archival Science*, 12(3), 287-318. <https://doi.org/10.1007/s10502-012-9171-8>
- Balasubramanian, P. (2023). Automation in Data Science, Software, and Information Services. In S. Y. Nof (Ed.), *Springer Handbook of Automation* (pp. 989-1014). Springer International Publishing. https://doi.org/10.1007/978-3-030-96729-1_46
- Bastian, J. A., & Flinn, A. (2018). Community Archives, Community Spaces: Heritage, Memory and Identity. 1re éd. Facet. <https://doi.org/10.29085/9781783303526>
- Benoit, E. I., & Eveleigh, A. (2019). Challenges, Opportunities and Future Directions of Participatory Archives. Chapter. In *Participatory Archives: Theory and Practice* (p. 211-218). Facet.
- Cardin, M. (1995). *Archivistique: information, organisation, mémoire: l'exemple du Mouvement coopératif Desjardins*. Éditions Septentrion.
- Cardin, M. (2013). Penser l'exploitation des archives en tant que système complexe. *Archives*, 45(1), 135-146.
- Cardin, M., & Desjardins, C. (2017). *Conservation et valorisation collaboratives dans une perspective marketing : vers un modèle d'archivistique ouverte. IFLA: Les nouveaux enjeux de la conservation du patrimoine documentaire / New challenges for the preservation of documentary heritage: Vol. 19*.
- Carini, P. (2009). Archivists as Educators: Integrating Primary Sources into the Curriculum. *Journal of Archi-*

- val Organization, 7(1 2), 41-50. <https://doi.org/10.1080/15332740902892619>
- Colavizza, G., Blanke, T., Jeurgens, C., & Noordegraaf, J. (2022). Archives and AI: An Overview of Current Debates and Future Perspectives. *Journal on Computing and Cultural Heritage*, 15(1), 1-15. <https://doi.org/10.1145/3479010>
- Cook, T. (2011). We Are What We Keep; We Keep What We Are: Archival Appraisal Past, Present and Future. *Journal of the Society of Archivists*, 32(2), 173-189. <https://doi.org/10.1080/00379816.2011.619688>
- Cook, T. (2013). Evidence, Memory, Identity, and Community: Four Shifting Archival Paradigms. *Archival Science*, 13(2-3), 95-120. <https://doi.org/10.1007/s10502-012-9180-7>
- Darda, K., Carre, M., & Cross, E. (2023). Value Attributed to Text-Based Archives Generated by Artificial Intelligence. *Royal Society Open Science*, 10(2), 220915. <https://doi.org/10.1098/rsos.220915>
- Duchemin, M. (1977). Le respect des fonds en archivistique: principes théoriques et problèmes pratiques. *Gazette des archives*, 97(1), 71-96.
- Duranti, L. (1997). The Archival Bond. *Archives and Museum Informatics*, 11(3 4), 213-218. <https://doi.org/10.1023/A:1009025127463>
- Duranti, L., Abdul-Mageed, M., Hofman, D., & Sullivan, P. (2021). I Trust AI, the latest InterPARES research project. *Anuario Escuela de Archivología*, 13, 36-55. <https://revistas.unc.edu.ar/index.php/anuario/article/view/37898>
- Duranti, L., & Rogers, C. (2012). Trust in Digital Records: An Increasingly Cloudy Legal Area. *Computer Law & Security Review*, 28(5), 522-531. <https://doi.org/10.1016/j.clsr.2012.07.009>
- Duranti, L., & Thibodeau, K. (2006). The Concept of Record in Interactive, Experiential and Dynamic Environments: The View of InterPARES*. *Archival Science*, 6(1), 13-68. <https://doi.org/10.1007/s10502-006-9021-7>
- Eveleigh, A. (2017). Participatory archives. In *Currents of archival thinking* (p. 299-325). Libraries Unlimited.
- Flinn, A., & Sexton, A. (2019). Activist Participatory Communities in Archival Contexts: Theoretical Perspectives. In *Participatory Archives: Theory and Practice* (p. 173-190). Facet.
- Freire, I. T., Moulin-Frier, C., Sanchez-Fibla, M., Arsiwalla, X. D., & Verschure, P. (s. d.). *Modeling the Formation of Social Conventions from Embodied Real-Time Interactions*. <http://arxiv.org/abs/1802.06108>
- Garoufallou, E., Siatri, R., Zafeiriou, G., & Balampanidou, E. (2013). The Use of Marketing Concepts in Library Services: A Literature Review. *Library Review*, 62(4/5), 312-334. <https://doi.org/10.1108/LR-06-2012-0061>
- Gouvernement du Canada. (1985). *La Loi sur la protection des renseignements personnels*. <https://laws-lois.justice.gc.ca/fra/lois/p-21/index.html>
- Gouvernement du Canada. (2022). *La Loi sur l'intelligence artificielle et les données (LIAD) - document complémentaire*. <https://ised-isde.canada.ca/site/innover-meilleur-canada/fr/loi-lintelligence-artificielle-donnees-liad-document-complementaire>.
- Gupta, D. K. (2015). Role of IFLA in Marketing Initiatives in Library and Information Services. In *Marketing and Consumer Behavior: Concepts, Methodologies, Tools, and Applications*, edited by Information Resources Management Association (p. 380-388). IGI Global. <https://doi.org/10.4018/978-1-4666-7357-1.ch017>
- Gupta, D., & Savard, R. (2011). *Marketing Libraries in a Web 2.0*. De Gruyter Saur. <https://doi.org/10.1515/9783110263534>
- Gupta, V. et P., & Shriram. (2019). Recommender Systems for Digital Libraries: A review of concepts and concerns. *Library philosophy and practice*, 2417. <https://core.ac.uk/download/pdf/215162001.pdf>
- Hedges, M., Marciano, R., & Goudarouli, E. (2022). Introduction to the Special Issue on Computational Archival Science. *Journal on Computing and Cultural Heritage*, 15(1), 1-2. <https://doi.org/10.1145/3495004>
- Héon, G. (1995). Les plans de classification en archivistique. *Archives*, 27(1), 73-90.
- Hodel, T. (2022). Chapter 6: Supervised and Unsupervised: Approaches to Machine Learning for Textual Entities. In *Archives, Access and Artificial Intelligence*, Édité par Lise Jaillant (p. 157-178). Bielefeld University Press. <https://doi.org/10.1515/9783839455845-007>
- Huvila, I. (2008). Participatory Archive: Towards Decentralised Curation, Radical User Orientation, and Broader Contextualisation of Records Management. *Archival Science*, 8(1), 15-36. <https://doi.org/10.1007/s10502-008-9071-0>
- Jaillant, L. (2022). *Archives, Access and Artificial Intelligence: Working with Born-Digital and Digitized Archival Collections*. Bielefeld University Press. <https://doi.org/10.1515/9783839455845>
- Jaillant, L., & Caputo, A. (2022). Unlocking Digital Archives: Cross-Disciplinary Perspectives on AI and Born-Digital Data. *AI & SOCIETY*, 37(3), 823-835. <https://doi.org/10.1007/s00146-021-01367-x>
- Jakubik, J., Schöffner, J., Hoge, V., Vössing, M., & Kühl, N. (2023). An Empirical Evaluation of Predicted Outcomes as Explanations in Human-AI Decision-Making. *Machine Learning and Principles and Practice of Knowledge Discovery in Databases*, 1752, 353-368. https://doi.org/10.1007/978-3-031-23618-1_24
- Jin, J., Yu, K., Kua, J., Zhang, N., Pang, Z., & Han, Q.-L. (2023). Cloud-Fog Automation: Vision, Enabling Technologies, and Future Research Directions. *IEEE Tran-*

- sactions on *Industrial Informatics*, 20(2), 1039-1054. <https://doi.org/10.1109/TII.2023.3272696>
- Kaldeli, E., Menis-Mastromichalakis, O., Bekiaris, S., Ralli, M., Tzouvaras, V., & Stamou, G. (2021). CrowdHeritage: Crowdsourcing for Improving the Quality of Cultural Heritage Metadata. *Information*, 12(2), 64. <https://doi.org/10.3390/info12020064>
- Lappin, J., Jackson, T., Matthews, G., & Ravenwood, C. (2021). Rival Records Management Models in an Era of Partial Automation. *Archival Science*, 21(3), 243-266. <https://doi.org/10.1007/s10502-020-09354-9>
- Lee, J. A. (2016). Be/Longing in the Archival Body: Eros and the Endearing Value of Material Lives. *Archival Science*, 16(1), 33-51. <https://doi.org/10.1007/s10502-016-9264-x>
- Lemay, Y., & Klein, A. (2015). Archives et émotions. *Documentation et bibliothèques*, 58(1), 5-16. <https://doi.org/10.7202/1028930ar>
- Lemieux, V. L., & Werner, J. (2023). Protecting Privacy in Digital Records: The Potential of Privacy-Enhancing Technologies. *Journal on Computing and Cultural Heritage*, 16(4), 1-18. <https://doi.org/10.1145/3633477>
- Lessard, R. (2005). Les archives publiques au Québec : La difficile construction d'une mémoire collective. *Archives*, 36(2), 173-198.
- Marciano, R., Lemieux, V., Hedges, M., Esteva, M., Underwood, W., Kurtz, M., & Conrad, M. (2018). Chapter 9: Archival Records and Training in the Age of Big Data. *Advances in Librarianship*, 44, 179-199. <https://doi.org/10.1108/S0065-28302018000044B010>
- McKemmish, S. (1994). Are records ever actual? In S. McKemmish & M. Piggott (Eds.), *The Records Continuum: Ian Maclean and Australian Archives First Fifty Years* (pp. 187-203). Monash University. https://bridges.monash.edu/articles/chapter/Are_records_ever_actual_/3823350
- Meng, F., & Ghena, B. (2023). Research on Text Recognition Methods Based on Artificial Intelligence and Machine Learning. *Advances in Computer and Communication*, 4(5), 340-344. <https://doi.org/10.26855/acc.2023.10.014>
- Miaskiewicz, T., & Kozar, K. A. (2011). Personas and User-Centered Design: How Can Personas Benefit Product Design Processes? *Design Studies*, 32(5), 417-430. <https://doi.org/10.1016/j.destud.2011.03.003>
- Murphy, R. M., Klopotoska, J. E., Keizer, N. F., Jager, K. J., Leopold, J. H., Dongelmans, D. A., Abu-Hanna, A., & Schut, M. C. (2023). Adverse Drug Event Detection Using Natural Language Processing: A Scoping Review of Supervised Learning Methods. *PLOS ONE*, 18(1), 0279842. <https://doi.org/10.1371/journal.pone.0279842>
- Ngoepe, M., Jacobs, L., & Mojapelo, M. (2022). Inclusion of Digital Records in the Archives and Records Management Curricula in a Comprehensive Open Distance E-Learning Environment. *Information Development*, 40(2), 190-201. <https://doi.org/10.1177/02666669221081812>
- Organisation internationale de normalisation. (2023). ISO/IEC 42001:2023 - Technologies de l'information : Intelligence artificielle - Système de management. (2023). ISO. Genève, Suisse. <https://www.iso.org/fr/standard/81230.html>
- Palmer, D. (2010). Emotional Archives: Online Photo Sharing and the Cultivation of the Self. *Photographies*, 3(2), 155-171. <https://doi.org/10.1080/17540763.2010.499623>
- Piggott, M., Bastian, J. A., & Flinn, A. (2018). Community Archives and the Records Continuum. In *Community Archives, Community Spaces: Heritage, Memory and Identity* (p. 41-60).
- Poole, A. H. (2019). Social Tagging and Commenting in Participatory Archives: A Critical Literature Review. In *Chapter. In Participatory Archives: Theory and Practice* (p. 15-32). Facet.
- Pötzsch, H. (2018). Archives and Identity in the Context of Social Media and Algorithmic Analytics: Towards an Understanding of iArchive and Predictive Retention. *New Media & Society*, 20(9), 3304-3022. <https://doi.org/10.1177/1461444817748483>
- Ray, D. et S., & William. (2016). *Marketing relationnel. Rentabiliser les politiques de satisfaction, fidélité, réclamation*. Dunod.
- Recanati, A. M. (2023). Job Disruption. In A. M. Recanati, AI Battle Royale (pp. 183-294). Springer International Publishing. https://doi.org/10.1007/978-3-031-19278-4_6
- Rolan, G., Humphries, G., Jeffrey, L., Samaras, E., Antsoukova, T., & Stuart, K. (2019). More Human than Human? Artificial Intelligence in the Archive. *Archives and Manuscripts*, 47(2), 179-203. <https://doi.org/10.1080/01576895.2018.1502088>
- Shabou, M., Basma, J. T., Knafo, J., & Gaudinat, A. (2020). Algorithmic Methods to Explore the Automation of the Appraisal of Structured and Unstructured Digital Data. *Records Management Journal*, 30(2), 175-200. <https://doi.org/10.1108/RMJ-09-2019-0049>
- Silva, S. C., Corbo, L., Vlačić, B., & Fernandes, M. (2023). Marketing Accountability and Marketing Automation: Evidence from Portugal. *EuroMed Journal of Business*, 18(1), 145-164. <https://doi.org/10.1108/EMJB-11-2020-0117>
- Smiraglia, R. P., & Cai, X. (2017). Tracking the evolution of clustering, machine learning, automatic indexing and automatic classification in knowledge organization. *Knowledge organisation*, 44(3), 215-233.
- Terras, M. (2022). Chapter 7: Inviting AI into the Archives: The Reception of Handwritten Recognition Technology into Historical Manuscript Transcription. In *Archives, Access and Artificial Intelligence, édité par Lise Jaillant* (p. 179-204). Bielefeld University Press. <https://doi.org/10.1515/9783839455845-008>

- Tzouganatou, A. (2022). Openness and Privacy in Born-Digital Archives: Reflecting the Role of AI Development. *AI & Society*, 37(3), 991-999. <https://doi.org/10.1007/s00146-021-01361-3>
- Université de Montréal (2024). *La Déclaration de Montréal sur l'IA responsable*. <https://declarationmontreal-iaresponsable.com/>
- Vassilakaki, E., & Moniarou-Papaconstantinou, V. (2017). Beyond Preservation: Investigating the Roles of Archivist. *Library Review*, 66(3), 110-126. <https://doi.org/10.1108/LR-09-2016-0077>
- Welland, S., & Cossham, A. (2019). Defining the Un-definable: An Analysis of Definitions of Community Archives. *Global Knowledge, Memory and Communication*, 68(8/9), 617-634. <https://doi.org/10.1108/GKMC-04-2019-0049>
- White, K. L., & Gilliland, A. J. (2010). Promoting Reflexivity and Inclusivity in Archival Education, Research, and Practice. *The Library Quarterly*, 80(3), 231-248. <https://doi.org/10.1086/652874>
- Williams, L. (2019). What Computational Archival Science Can Learn from Art History and Material Culture Studies. *2019 IEEE International Conference on Big Data (Big Data)*, 3153-3155. <https://doi.org/10.1109/BigData47090.2019.9006527>
- Yeo, G. (2005). Understanding users and use: a market segmentation approach. *Journal of the Society of Archivists*, 26(1), 25-53.
- Yeo, G. (2012). Bringing things together: Aggregate records in a digital age. *Archivaria*, 43-91.
- Zini, É. (2020). L'utilisation des personas à Bibliothèque et Archives nationales du Québec: Une rencontre avec notre public. *Documentation et bibliothèques*, 66(3), 16-28. <https://doi.org/10.7202/1071198ar>