

The Application of Artificial Intelligence in Digital Archives: Trust, Collaboration, and Shared Professional Ethics

Lise, Jaillant¹, Arran Rees², Seyyed Ahmad Bagherzadeh Arjmandi³ 

1. Loughborough University, United Kingdom
2. University of Leeds, United Kingdom
3. PhD Candidate in Islamic History (Translator).

Article Info

Article type:

Research Article

Article history:

Received: 6 September 2025

Received in revised form: 29 October 2025

Accepted: 27 December 2025

Available online: 2 September 2026

Keywords:

Abstract

Introduction

The article adopts a bold approach, analyzing the “locked” state of digital archives not merely as a technical problem, but as a human challenge rooted in “systemic distrust.” This approach brings about a fundamental shift in the research paradigm surrounding digital archives and information science. This is because, unlike previous studies that have primarily focused on developing algorithms and technical solutions, this research demonstrates that even the most advanced technologies lose their effectiveness in the absence of “social capital” and trust among stakeholders. This approach highlights the importance of human trust and the quality of interactions among human actors as essential conditions for the successful use of digital archives, positioning the research as a pioneering and transformative contribution to the field. The article adopts an interdisciplinary approach to analyze this phenomenon, examining its various dimensions simultaneously. Archival and information sciences are employed to analyze operational and professional challenges; Digital Humanities to serve as a bridge between technology and content; technology ethics to examine ethical challenges associated with artificial intelligence; and organizational studies and the sociology of professions to understand the dynamics of trust and collaboration. This interdisciplinary approach was developed through semi-structured interviews with thirty specialists. This method enables a deeper understanding of concepts such as trust, facilitates more comprehensive and in-depth analysis, and provides access to the underlying attitudes, fears, and assumptions of the actors involved. In domestic research, digital archives and artificial intelligence have primarily been examined from two perspectives: technical and engineering, and library and archival studies. Studies with titles such as “Intelligent Algorithms for Information Retrieval in Digital Archives” and “Technical Solutions for the Preservation and Maintenance of Digital Resources” focus primarily on the instrumental and infrastructural dimensions of this phenomenon. By contrast, only a limited number of studies have addressed the human, cultural, and social dimensions of this field. For example, articles focusing on “Ethical Requirements in the Digital Environment” or “Access Policies in National Archives,” although they have addressed issues of trust and ethics, have rarely placed them at the center of their analysis. Moreover, they have not explored the systematic interrelationship between “trust” and “artificial intelligence” as a fundamental issue in unlocking digital archives. through its

Cite this article: Bagherzadeh Arjmandi, A. (2025). ite this article: Bagherzadeh Arjmandi, A. (2025). The Application of Artificial Intelligence in Digital Archives: Trust, Collaboration, and Shared Professional Ethics, 4(1), 94-120. *Journal of Documentary and Archival Studies*.



© The Author(s) retain the copyright.

Publisher: The Organization of Libraries, Museums and Documents Center of Astan-e Quds Razavi.

DOI: <https://doi.org/10.30481/psa.2025.486830.1096>

theoretical boldness, the present article not only addresses a gap in the international literature but also offers a conceptual framework for Persian-speaking researchers, encouraging them to move beyond the dominance of technical approaches and place the examination of “trust” and “human relationships” at the heart of future research on emerging technologies in the fields of archives and information. The translation and critique of this work can initiate interdisciplinary dialogue within the domestic academic community, which appears essential for moving beyond the stage of technological dominance toward the stage of its judicious integration with human values. The findings of the study indicate that distrust in digital archives has three structural dimensions that operate as interconnected barriers. First, interpersonal distrust, or “professional othering,” arising from misunderstandings and a lack of mutual understanding between archivists and researchers. A clear example is the expectation among researchers that archivists preserve “everything,” even though such an undertaking is not practically feasible. The second case is technological distrust, rooted in fears of the erosion of professional authority. Archivists whose professional identities are grounded in control over collections tend to resist against tools that could diminish their authority over those collections. Third, institutional distrust, whereby artificial intelligence is viewed not only as a tool that can reinforce systemic biases but also as a mechanism for legitimizing them. When an algorithm produces a decision, the result may appear objective and impartial, even though it can be shaped by underlying power structures and institutional biases. In addressing these challenges, the article shifts its focus away from technical solutions and emphasizes the need to foster dialogue and meaningful human interaction. The concept of the “archival loop,” as a powerful metaphor, brings all stakeholders—from document producers to final users—within a shared ecosystem of common interests. It suggests that the restoration of this fragmented loop can be achieved through dialogue and mutual understanding. Nevertheless, the study also has certain limitations, including a limited analysis of the distribution of power and privilege among stakeholders and the absence of an operational model for structuring such dialogue in a way that adequately accounts for differences in the institutional positions and levels of access among participants. Overall, this article offers a novel framework for analyzing and making sense of trust. Its scholarly contribution lies not in the development of a new algorithm, but lies in proposing an analytical framework that shifts the pathway toward opening up digital archives from technology to anthropology, and from hardware and software to social capital and human trust. The study’s central message is clear: before making any technical investment, it is essential to invest in trust and human relationships among stakeholders, as trust is what enables the meaningful use of technology and facilitates the development of a dynamic and ethically grounded ecosystem.

The research article entitled “Artificial Intelligence in Digital Archives: Trust, Collaboration, and Shared Professional Ethics,” authored by Lise Jaillant of Loughborough University and Arran Rees of the University of Leeds, was published in 2023 in volume 38, issue 2 of *Digital Scholarship in the Humanities*. The article is published on pages 571–585 and has the DOI 10.1093/lle/fqac073. It was initially made available online through Advance Access on 17 November 2022 and was subsequently peer-reviewed and accepted as a Full Paper. The article is published under the Creative Commons Attribution 4.0 International License, which permits its use, distribution, and reproduction in any medium, provided that appropriate credit is given to the original source. The article was published by Oxford University Press on behalf of the European Association for Digital Humanities (EADH). The version cited in this study was retrieved from the publisher’s website.

کاربرد هوش مصنوعی در آرشیوهای دیجیتال: اعتماد، همکاری و اخلاق حرفه‌ای مشترک

لیز ژایان^۱، آران ریس^۲، سید احمد باقرزاده ارجمندی^۳

۱. دانشگاه لافبرو، بریتانیا

۲. دانشگاه لیدز، بریتانیا

۳. دانشجوی دکتری تاریخ (مترجم)

اطلاعات مقاله

نوع مقاله:

پژوهشی

تاریخچه مقاله:

دریافت: ۱۴۰۴/۰۶/۱۵

بازنگری: ۱۴۰۴/۰۸/۰۷

پذیرش: ۱۴۰۵/۱۰/۰۶

انتشار: ۱۴۰۵/۰۶/۱۱

کلیدواژه‌ها:

اعتماد،

هوش مصنوعی،

اخلاق حرفه‌ای،

همکاری،

آرشیوهای دیجیتال

متولد شده.

اعتماد به مثابه کلید انسانی گشایش آرشیوهای دیجیتال *

مقاله حاضر با ارائه رویکردی جسورانه، مسئله «قفل بودن» آرشیوهای دیجیتال را نه به عنوان یک مشکل فنی صرف، بلکه به مثابه چالشی انسانی و ناشی از «بی‌اعتمادی سیستماتیک» تحلیل می‌کند. این نگاه، تحولی بنیادین در رویکرد پژوهش‌های مرتبط با آرشیوهای دیجیتال و علوم اطلاعات ایجاد می‌کند، زیرا برخلاف مطالعات پیشین که عمدتاً بر توسعه الگوریتم‌ها و راهکارهای فنی تمرکز داشته‌اند، این تحقیق نشان می‌دهد که حتی پیشرفته‌ترین فناوری‌ها در غیاب «سرمایه اجتماعی» و اعتماد میان ذی‌نفعان، کارکرد مؤثر خود را از دست خواهند داد. این رویکرد، اهمیت اعتماد انسانی و کیفیت تعامل میان کنشگران را به عنوان شرط لازم برای بهره‌برداری موفق از آرشیوهای

* مقاله پژوهشی با عنوان «کاربرد هوش مصنوعی در آرشیوهای دیجیتال: اعتماد، همکاری و اخلاق حرفه‌ای مشترک» توسط لیز ژایان (Lise Jaillant) از دانشگاه لافبرو و آران ریس (Arran Rees) از دانشگاه لیدز تألیف شده و در سال ۲۰۲۳م در جلد ۳۸، شماره ۲ ژورنال معتبر «مطالعات دیجیتال در علوم انسانی» (Digital Scholarship in the Humanities) منتشر شده است. این مقاله در صفحات ۵۷۱ تا ۵۸۵ این شماره جای گرفته و دارای شناسه دیجیتال (DOI: <https://doi.org/10.1093/llc/fqac073>) می‌باشد. مقاله به صورت پیش دسترسی (Advance Access) در ۱۷ نوامبر ۲۰۲۲م در دسترس قرار گرفته و به عنوان یک مقاله کامل پژوهشی (Full Paper) داوری و پذیرش نهایی شده است. این اثر تحت مجوز Creative Commons Attribution ۴.۰ International License منتشر شده که استفاده، توزیع و بازتولید آن را با ذکر منبع اصلی در هر رسانه‌ای آزاد می‌سازد. ناشر این مقاله، انتشارات دانشگاه آکسفورد به نمایندگی از انجمن اروپایی علوم انسانی دیجیتال (EADH) است. نسخه مورد استناد از طریق پایگاه ناشر در آدرس <https://academic.oup.com/dsh/article/38/2/571/6832097> بازایی شده است.

استناد: باقرزاده ارجمندی، سید احمد (۱۴۰۵). کاربرد هوش مصنوعی در آرشیوهای دیجیتال: اعتماد، همکاری و اخلاق حرفه‌ای مشترک. ۹۴ - ۱۲۰، (۱)۴.

پژوهشنامه مطالعات اسنادی و آرشیوی. <https://doi.org/10.30481/psa.2025.486830.1096>

ناشر: سازمان کتابخانه‌ها، موزه‌ها و مرکز اسناد آستان قدس رضوی © نویسندگان.



دیجیتال برجسته می‌سازد و پژوهش را در زمره مطالعات پیشگام و تحول‌آفرین قرار می‌دهد. برای تحلیل این پدیده، مقاله از رویکردی بین‌رشته‌ای بهره می‌برد که ابعاد مختلف مسئله را به صورت هم‌زمان مورد بررسی قرار می‌دهد. علوم اطلاعات و آرشیو، به منظور تحلیل چالش‌های عملیاتی و حرفه‌ای، علوم انسانی دیجیتال به عنوان پلی میان فناوری و محتوا، اخلاق فناوری برای واکاوی چالش‌های اخلاقی مرتبط با هوش مصنوعی و مطالعات سازمانی و جامعه‌شناسی حرفه‌ها برای فهم پویایی‌های اعتماد و همکاری مورد استفاده قرار گرفته‌اند. این رویکرد میان‌رشته‌ای با استفاده از صاحب‌های نیمه‌ساختاریافته با سی متخصص طراحی شده است؛ روشی که برای درک عمیق مفاهیمی همچون اعتماد، توانایی ارائه تحلیل‌هایی جامع‌تر و عمیق‌تر از هر روش کمی را دارد و امکان دسترسی به نگرش‌ها، ترس‌ها و پیش‌فرض‌های پنهان ذی‌نفعان را فراهم می‌آورد.

در بستر پژوهش‌های داخلی، موضوع آرشیوهای دیجیتال و هوش مصنوعی عمدتاً از دو منظر فنی-مهندسی و کتابداری-آرشیوی مورد توجه قرار گرفته است. مطالعاتی با عناوینی همچون «الگوریتم‌های هوشمند برای بازیابی اطلاعات در آرشیوهای دیجیتال» یا «راهکارهای فنی برای حفظ و نگهداری منابع دیجیتال» بر جنبه‌های ابزاری و زیرساختی این پدیده متمرکزند. در مقابل، محدود پژوهش‌هایی به تبیین ابعاد انسانی، فرهنگی و اجتماعی این حوزه پرداخته‌اند. برای نمونه، مقالاتی با محوریت «الزامات اخلاقی در محیط دیجیتال» یا «سیاست‌گذاری دسترسی در آرشیوهای ملی»، اگرچه به مسئله اعتماد و اخلاق اشاره داشته‌اند، اما کمتر آن را در کانون تحلیل خود قرار داده و به‌ویژه پیوند نظام‌مند «اعتماد» و «هوش مصنوعی» را به عنوان یک مسئله بنیادین در گشایش آرشیوهای دیجیتال مورد کاوش قرار نداده‌اند. از این منظر، مقاله حاضر با شجاعت نظری خود، نه تنها خلأیی در ادبیات جهانی را پر می‌کند، بلکه الگویی مفهومی برای پژوهشگران فارسی‌زبان نیز فراهم می‌آورد تا از انحصار رویکردهای فنی خارج شده و واکاوی «اعتماد» و «روابط انسانی» را در قلب مطالعات آینده خود درباره فناوری‌های نوین در حیطه آرشیو و اطلاعات قرار دهند. ترجمه و نقد این اثر می‌تواند زنجیره‌ای از گفت‌وگوهای میان‌رشته‌ای را در فضای آکادمیک داخلی کلید بزند که برای گذار از مرحله غلبه فناوری به مرحله تلفیق خردمندانه آن با ارزش‌های انسانی، ضروری به نظر می‌رسد.

یافته‌های پژوهش نشان می‌دهد که بی‌اعتمادی در آرشیوهای دیجیتال سه وجه ساختاری دارد و به عنوان موانعی به هم پیوسته عمل می‌کند. نخست، بی‌اعتمادی بین‌افردی یا «غیرت‌سازی حرفه‌ای» که ناشی از سوء تفاهم‌ها و فقدان درک متقابل میان آرشیویست‌ها و پژوهشگران است. نمونه شاخص این مسئله، انتظار پژوهشگران از آرشیویست‌ها برای حفظ «همه چیز» است، درحالی‌که چنین امری از نظر عملی ممکن نیست. دوم، بی‌اعتمادی فناورانه که ریشه در ترس از فرسایش اقتدار حرفه‌ای دارد. آرشیویست‌هایی که هویت حرفه‌ای‌شان مبتنی بر کنترل مجموعه‌ها شکل گرفته، در برابر ابزارهایی که این کنترل را تضعیف می‌کنند، مقاومت نشان می‌دهند. سوم، بی‌اعتمادی نهادی که هوش مصنوعی را نه تنها به عنوان ابزاری برای تشدید سوگیری‌های سیستمیک، بلکه به عنوان مکانیسمی مشروعیت‌بخش آن‌ها معرفی می‌کند. هنگامی که الگوریتم حکمی صادر می‌کند، این حکمی به ظاهر عینی و بی‌طرف جلوه می‌کند، درحالی‌که واقعیت متأثر از ساختارهای قدرت و سوگیری‌های نهادی است.

در مواجهه با این چالش‌ها، مقاله به جای تمرکز بر راه‌حل‌های فنی، بر ضرورت خلق گفت‌وگو و تعامل انسانی تأکید می‌کند. مفهوم «حلقه آرشیوی» به عنوان استعاره‌ای قدرتمند، تمامی ذی‌نفعان از تولیدکننده سند تا کاربر نهایی را در یک اکوسیستم واحد با منافع مشترک جای می‌دهد و پیشنهاد می‌کند که بازسازی این حلقه گسسته از طریق گفت‌وگو و درک متقابل امکان‌پذیر شود. با این حال، پژوهش محدودیت‌هایی نیز دارد، از جمله تحلیل محدود از توزیع قدرت و امتیاز میان ذی‌نفعان و فقدان مدلی عملیاتی برای شکل‌دهی به این دیالوگ، به گونه‌ای که تفاوت موقعیت نهادی و سطح دسترسی میان مشارکت‌کنندگان به طور روشن مورد توجه قرار گیرد. در مجموع، این مقاله چارچوبی نوین برای تحلیل و معناسازی اعتماد ارائه می‌دهد. دستاورد علمی آن نه در توسعه الگوریتمی جدید، بلکه در ارائه چارچوبی تحلیلی است

که مسیر گشودن آرشیوهای دیجیتال را از فناوری به انسان‌شناسی و از سخت‌افزار و نرم‌افزار به سرمایه اجتماعی و اعتماد انسانی معطوف می‌سازد. پیام نهایی پژوهش روشن است: پیش از هر سرمایه‌گذاری فنی، سرمایه‌گذاری بر اعتماد و روابط انسانی میان ذی‌نفعان ضروری می‌باشد، زیرا همین اعتماد است که امکان بهره‌برداری واقعی از فناوری و شکل‌دهی اکوسیستمی پویا و اخلاق‌مدار را فراهم می‌آورد.

چکیده

سیاست‌گذاران به‌طور روزانه اسناد دیجیتال تولید می‌کنند و بخشی از این اسناد در مخازن آرشیوی نگهداری می‌شود. با این حال، دسترسی به این منابع آرشیوی به دلایل متعدد از جمله حفاظت داده‌ها، حساسیت اطلاعات، امنیت ملی و حقوق مالکیت معنوی بسیار دشوار است. هوش مصنوعی (AI) می‌تواند در جهت افزایش دسترسی به آرشیوها به‌کار گرفته شود، اما این کاربرد هنوز در مرحله آزمایشی قرار دارد. در کنار شکاف‌های مهارتی که موجب تأریک ماندن آرشیوها می‌شود، بررسی مسائل مربوط به بی‌اعتمادی و سوءتفاهم نیز ضروری است. این مقاله استدلال می‌کند که اگرچه کارمندان دولتی، آرشیویست‌ها و دانشگاهیان اصول حرفه‌ای مشابهی دارند که در قالب کدهای اخلاقی حرفه‌ای بیان شده‌اند، این اصول به‌ندرت میان آنان به اشتراک گذاشته می‌شود. این فقدان ارتباط منجر به شکل‌گیری احساس بی‌اعتمادی میان ذی‌نفعان می‌گردد. همچنین، بی‌اعتمادی نسبت به فناوری نیز یکی از موانع اجرای مؤثر ابزارهای هوش مصنوعی است. از این‌رو، نویسندگان پیشنهاد می‌کنند که آشکارسازی اخلاق حرفه‌ای مشترک میان ذی‌نفعان می‌تواند زمینه‌ساز همکاری‌های انسانی عمیق‌تر شود؛ همکاری‌هایی که به نوبه خود موجب شکل‌گیری اعتماد نسبت به سامانه‌ها و ابزارهای هوش مصنوعی خواهد شد. این پژوهش بر پایه مصاحبه‌های نیمه‌ساختاریافته با سی نفر از متخصصان دولتی، آرشیویست‌ها، مورخان، پژوهشگران علوم انسانی دیجیتال و دانشمندان رایانه انجام شده است. پژوهش‌های پیشین عمدتاً بر حفظ اسناد دیجیتال تمرکز داشته‌اند، نه بر دسترسی به آن‌ها و بیشتر به نقش آرشیویست‌ها پرداخته‌اند تا تولیدکنندگان اسناد مانند کارمندان دولتی. این مقاله نخستین تلاش برای بررسی کاربرد هوش مصنوعی در آرشیوهای دیجیتال از منظر اعتماد و همکاری میان تمامی حلقه‌ها است.

مقدمه

آرشیوهای دیجیتال متولدشده طی نزدیک به سی سال گذشته گردآوری و حفظ شده‌اند. با این حال، دسترسی به این اسناد همچنان چالشی پیچیده برای نهادهایی است که مسئول نگهداری و مدیریت آن‌ها هستند و نیز برای پژوهشگرانی که قصد استفاده از این منابع را دارند. موانع دسترسی متعدد و شامل مواردی همچون حفاظت از داده‌ها، حساسیت اطلاعات و حقوق مالکیت معنوی می‌شوند. در مورد مجموعه‌های آرشیوی دولتی، انتشار اطلاعاتی که بالقوه حساس یا خصوصی هستند می‌تواند امنیت ملی را به خطر اندازد و موجب شرمساری شرکای خارجی شود. با این حال، همه مجموعه‌ها چنین سطحی از خطر را ندارند و فراهم‌سازی دسترسی به اسناد دیجیتال برای پاسخ‌گویی دولت و امکان‌پذیر ساختن نگارش تاریخ امری ضروری است.

با توجه به این‌که اسناد دیجیتال متولدشده به‌طور فزاینده‌ای بخش عمده‌ای از دریافت‌های جدید آرشیوی را تشکیل

۱. آرشیوهای دیجیتال متولدشده (Born-digital archives) به منابع آرشیوی‌ای اطلاق می‌شود که از همان ابتدا در قالب دیجیتال تولید شده‌اند. برای نمونه: ایمیل‌ها، فایل‌های PDF، اسناد Word، فایل‌های صوتی و تصویری دیجیتال. این نوع اسناد با منابع «دیجیتال‌شده» (digitised materials) تفاوت دارند. منابع دیجیتال‌شده از طریق فرآیندهایی مانند اسکن یا عکاسی از نسخه‌های فیزیکی ایجاد می‌شوند.

می‌دهند، یافتن راه‌حلهایی برای رفع موانع دسترسی به این منابع به یک اولویت کلیدی تبدیل شده است. بازگشایی حجم عظیمی از داده‌های دیجیتال کاری نیست که بتوان به صورت دستی انجام داد. هوش مصنوعی^۱ این ظرفیت را دارد که آرشیوهای دیجیتال متولدشده را در دسترس‌تر و قابل استفاده‌تر سازد. اطلاعات حساس می‌توانند به صورت خودکار شناسایی شوند و این امر امکان انتشار داده‌های غیرحساس را فراهم می‌سازد. هوش مصنوعی همچنین در مواردی که جستجوی مبتنی بر کلیدواژه مؤثر نیست - برای مثال در آرشیوهای وب که شامل ترابایت داده هستند - قابل استفاده است. فرآیندهای پیشنهادی مبتنی بر هوش مصنوعی - مشابه عملکرد «کاربرانی که این کالا را خریدند، این کالا را نیز خریدند» - می‌توانند کشف منابع آرشیوی ای را ممکن سازند که پیشتر غیرقابل دسترس بوده‌اند. در سال‌های اخیر، پروژه‌های متعددی تلاش کرده‌اند تا چالش‌های کلیدی مرتبط با آرشیوهای دیجیتال متولدشده را شناسایی کرده و با بهره‌گیری از هوش مصنوعی و سایر فناوری‌های پیشرفته، راه‌حلهایی برای مشکل «قفل بودن» این منابع ارائه دهند.^۲ بااین حال، کاربرد هوش مصنوعی در حوزه آرشیو همچنان در مرحله آزمایشی قرار دارد و به ندرت فراتر از مجموعه‌های کوچک و داده‌هایی که برای آزمایش انتخاب شده‌اند، اجرا گردیده است. گسترش استفاده از هوش مصنوعی در حجم وسیعی از منابع آرشیوی با موانع متعددی مواجه است، از جمله دشواری در جذب متخصصان با مهارت‌های لازم و حقوق نسبتاً پایین در بخش آرشیو. این موانع را می‌توان با تأمین بودجه بیشتر برای جذب استعدادها و فراهم سازی فرصت‌های آموزشی برای کارکنان فعلی این حوزه برطرف کرد. بااین حال، چالش پیچیده‌تر، فقدان اعتماد و نبود اخلاق حرفه‌ای مشترک در سراسر چرخه آرشیوی است؛ از تولیدکنندگان اسناد گرفته تا آرشیویست‌ها و از آرشیویست‌ها تا پژوهشگران. فرهنگ نامۀ آکسفورد (OED) واژه «اعتماد» را چنین تعریف می‌کند: «باور راسخ به قابلیت اعتماد، حقیقت یا توانایی فرد یا چیزی» و «اطمینان یا ایمان به شخص یا چیز، یا به یکی از ویژگی‌های آن‌ها». این واژه از زبان‌های ژرمنی به ارث رسیده و اغلب با مفاهیمی چون محافظت، پناه و ایمنی پیوند دارد. ما زمانی اعتماد می‌کنیم که احساس امنیت داشته باشیم. در مقابل، در شرایطی که خطرناک یا ناامن به نظر برسد، اعتماد از میان می‌رود. ایجاد اعتماد برای بسیاری از حرفه‌ها ضروری است، زیرا باید نشان دهند فعالیتشان سودمند یا دست‌کم بی‌ضرر است. به همین دلیل، پژوهشگران دانشگاهی برای انجام برخی تحقیقات از جمله پژوهش‌هایی که شامل مشارکت انسانی هستند باید تأییدیه کمیته‌های اخلاق را دریافت کنند. بر اساس تعریف فرهنگ‌نامه آکسفورد، «اخلاق» به سیستم اصولی اشاره دارد که رفتار حرفه‌ای یا شخصی را هدایت می‌کند. اگرچه این چارچوب کلی تعریف شده است، اما تفسیر و اجرای آن در موقعیت‌های عینی همواره مورد مناقشه متخصصان بوده است. در حوزه فناوری‌های نوظهور مانند هوش مصنوعی، همین فقدان اجماع در کنار بی‌اعتمادی، می‌تواند به ایستایی کامل منجر شود. در چنین شرایطی، هوش مصنوعی نمی‌تواند کمکی به دسترسی بیشتر به آرشیوهای دیجیتال کند. گام نخست، پرداختن به فقدان اعتماد و نبود قواعد رفتاری مشترک است. زمانی که اعتماد انسانی بازسازی شود، هوش مصنوعی می‌تواند به عنوان گام بعدی وارد عمل شود و به گشودن آرشیوهای دیجیتال کمک کند.

۱. هوش مصنوعی (AI) به کاربرد فرآیندهای محاسباتی برای یادگیری، تصمیم‌گیری و حل مسئله اطلاق می‌شود. یادگیری ماشین (Machine Learning یا ML) که اغلب در مباحث مربوط به هوش مصنوعی مطرح می‌شود، یکی از کاربردهای آن به شمار می‌رود. یادگیری ماشین فرآیندی است که طی آن یک سامانه رایانه‌ای می‌تواند بر اساس تجربه‌های پیشین خود، به طور مستقل به یادگیری و بهبود عملکرد ادامه دهد. یادگیری ماشین به عنوان زیرمجموعه‌ای از هوش مصنوعی شناخته می‌شود.

۲. برای نمونه، پروژه‌هایی مانند AURA (آرشیوها در بریتانیا/ جمهوری ایرلند و هوش مصنوعی) AEOLIAN، aura-network.net (هوش مصنوعی برای سازمان‌های فرهنگی) aeolian-network.net و AI4LAM (هوش مصنوعی برای کتابخانه‌ها، آرشیوها و موزه‌ها) ai4lam در این زمینه فعالیت دارند. شایان ذکر است که هوش مصنوعی نه تنها در رابطه با آرشیوهای دیجیتال متولدشده، بلکه در مورد مجموعه‌های دیجیتال شده نیز با استقبال گسترده‌ای مواجه شده است.

در این مقاله، به بررسی مسائل مربوط به اعتماد و اخلاق حرفه‌ای مشترک پرداخته می‌شود، البته با استفاده از داده‌های کیفی حاصل از پروژه «گشودن گذشته دیجیتال ما» که تیمی از پژوهشگران علوم انسانی دیجیتال را با متخصصان دولتی و آرشیویست‌ها گرد هم آورد. این پروژه با همکاری دفتر کابینه بریتانیا سازمان‌دهی شد و اگرچه تمرکز اصلی آن بر بریتانیا بود، اما متخصصانی از آمریکای شمالی و اروپا نیز در آن مشارکت داشتند. هدف این پروژه، درک بهتر موانع اجرای ابزارهای مبتنی بر هوش مصنوعی برای دسترسی به مجموعه‌های آرشیوی دیجیتال و شناسایی دغدغه‌های کلیدی ذی‌نفعان در سراسر فرآیند آرشیوی است؛ از کارمندان دولتی و سیاست‌گذارانی که اسناد را تولید می‌کنند و در نهایت به آرشیو ملی می‌رسند تا آرشیویست‌ها و متخصصان حوزه GLAM (گالری‌ها، کتابخانه‌ها، آرشیوها و موزه‌ها) که وظیفه حفظ این مجموعه‌های دیجیتال را بر عهده دارند و پژوهشگران دانشگاهی که در تلاش برای دسترسی و استفاده از آن‌ها هستند.

این پژوهش بر پایه مصاحبه‌های نیمه‌ساختاریافته با ۳۰ نفر از متخصصان دولتی، آرشیویست‌ها، مورخان، پژوهشگران علوم انسانی دیجیتال و دانشمندان رایانه انجام شده است و دو پرسش اصلی را دنبال می‌کند: ۱. کاربردهای بالقوه هوش مصنوعی برای دسترسی و استفاده بهتر از آرشیوهای دیجیتال متولد شده چیست؟ ۲. موانع اجرای این فناوری‌ها کدام‌اند؟ یافته‌ها حاکی از آن است که دغدغه‌ها درباره کاربرد هوش مصنوعی میان ذی‌نفعان در خدمات دولتی، بخش GLAM و دانشگاه‌ها مشترک است. با این حال، گفت‌وگوها، رویدادها و پروژه‌های مشترک میان این سه گروه بسیار نادر است. این مقاله استدلال می‌کند که اگرچه کارمندان دولتی، متخصصان GLAM و دانشگاهیان اصول حرفه‌ای مشابهی دارند که در قالب کدهای اخلاقی بیان شده‌اند، این اصول به ندرت میان آنان به اشتراک گذاشته می‌شود. این فقدان ارتباط منجر به شکل‌گیری احساس بی‌اعتمادی میان ذی‌نفعان می‌گردد (برای مثال، زمانی که آرشیویست‌ها به پژوهشگران اعتماد ندارند که در مواجهه با اطلاعات حساس تصمیم درستی بگیرند). بی‌اعتمادی نسبت به فناوری نیز یکی از موانع اجرای مؤثر ابزارهای هوش مصنوعی است. بنابراین، نویسندگان پیشنهاد می‌کنند که آشکارسازی اخلاق حرفه‌ای مشترک میان ذی‌نفعان می‌تواند زمینه‌ساز همکاری‌های انسانی عمیق‌تر شود؛ همکاری‌هایی که به نوبه خود موجب شکل‌گیری اعتماد نسبت به سامانه‌ها و ابزارهای هوش مصنوعی خواهد شد. اگر تولیدکنندگان اسناد و آرشیویست‌ها به این باور برسند که پژوهشگران با داده‌ها به صورت اخلاقی برخورد خواهند کرد، تمایل بیشتری برای به اشتراک‌گذاری این داده‌ها برای پژوهش (از جمله پژوهش‌های محاسباتی) خواهند داشت. اعتماد به انسان‌ها می‌تواند زمینه‌ساز اعتماد به فناوری شود.

سه بخش اصلی مقاله حاضر این استدلال را توسعه می‌دهند که اخلاق حرفه‌ای مشترک می‌تواند به همکاری نزدیک‌تر میان تولیدکنندگان اسناد، آرشیویست‌ها و پژوهشگران منجر شود و در نهایت اعتماد به ابزارهای هوش مصنوعی را تقویت کند. بخش نخست به زمینه کلی و آثار مرتبط درباره موانع دسترسی به مجموعه‌های دیجیتال متولد شده و کاربردهای بالقوه هوش مصنوعی برای گشودن این مجموعه‌ها می‌پردازد و روش‌شناسی پژوهش را توضیح می‌دهد. بخش دوم، با تکیه بر مجموعه‌ای از مصاحبه‌ها، بر مسائل مربوط به بی‌اعتمادی میان ذی‌نفعان و بی‌اعتمادی نسبت به فناوری تمرکز دارد که به فقدان دسترسی به مجموعه‌های دیجیتال منجر می‌شود. بخش سوم، پیشنهاد ما برای ایجاد اعتماد از طریق همکاری و اخلاق حرفه‌ای مشترک را بررسی می‌کند تا زمینه‌ساز گسترش کاربرد هوش مصنوعی و فناوری‌های نوین در آرشیوهای دیجیتال شود. اعتماد محض به هوش مصنوعی نه مطلوب است و نه ممکن. با این حال، ایجاد اعتماد نسبت به ابزارهای هوش مصنوعی که به صورت جمعی توسعه یافته‌اند و قابل توضیح هستند، در میان خدمات دولتی، بخش GLAM و دانشگاه‌ها، می‌تواند به دسترسی بیشتر و استفاده مؤثرتر از آرشیوهای دیجیتال منجر شود.

زمینه، آثار مرتبط و روش‌شناسی

در طول حدود سی سال گذشته، تلاش‌های قابل توجهی برای ایجاد رویه‌های مؤثر در حفظ منابع دیجیتال متولد شده در بخش GLAM (گالری‌ها، کتابخانه‌ها، آرشیوها و موزه‌ها) صورت گرفته است. (Deegan & Tanner, 2006; Delve & Anderson, 2014; Corrado & Sandy, 2017; Owens, 2018; Harvey & Weatherburn, 2018) نهادهای پشتیبان تأثیرگذاری مانند «ائتلاف حفظ دیجیتال» نقش مهمی در نهادینه‌سازی و پایداری فعالیت‌های متخصصان حفظ دیجیتال در طیف وسیعی از سازمان‌های میراث فرهنگی ایفا کرده‌اند. با وجود آن‌که فرآیندهای حفظ دیجیتال به‌طور فزاینده‌ای در هسته فعالیت‌های نهادهای GLAM ادغام شده‌اند، فرآیندهای مربوط به دسترسی به همان مجموعه‌های دیجیتال متولد شده با سرعت کمتری توسعه یافته‌اند. (Jaillant, 2019) دلایل متعددی برای این امر وجود دارد، از جمله مسائل حقوقی مرتبط با محتوا مانند نگرانی‌های مربوط به حریم خصوصی و حقوق مالکیت معنوی و همچنین چالش‌های فناورانه که ارائه دسترسی به منابع در قالب‌های مناسب را دشوار می‌سازند.

نگرانی‌های حقوقی عمدتاً به مقررات عمومی حفاظت از داده‌ها (GDPR) در اتحادیه اروپا، قانون حفاظت از داده‌ها ۲۰۱۸م در بریتانیا و قوانین کپی‌رایت مربوط می‌شوند. GDPR و قوانین حفاظت از داده‌ها، نهادها و سازمان‌هایی را که با داده‌های شخصی سروکار دارند، موظف می‌سازند تا «حریم خصوصی از طراحی» را در فرآیندهای خود لحاظ کنند. با این حال، این مقررات در مواردی که داده‌های شخصی به منظور منافع عمومی آرشیو می‌شوند، معافیت‌هایی نیز در نظر گرفته‌اند. (Lomas, 2019) راهنمای آرشیو ملی بریتانیا درباره آرشیو داده‌های شخصی (۲۰۱۸م) تصریح می‌کند که قوانین جدید حفاظت از داده‌ها مانع آرشیو کردن نیستند، اما بر نحوه بیان اهداف آرشیو داده‌های شخصی و چگونگی ارائه دسترسی به مجموعه‌هایی که شامل این داده‌ها هستند، تأثیر می‌گذارند. (The National Archives, 2018) اسنادی که حاوی داده‌های شخصی تشخیص داده می‌شوند، با دقت مدیریت شده و معمولاً تا پایان عمر فرد مربوطه از دسترس عمومی خارج می‌مانند (آرشیو ملی پیشنهاد می‌کند که طول عمر ۱۰۰ ساله برای این افراد فرض شود، مگر خلاف آن اثبات گردد).^۲ با این حال، در برخی شرایط خاص، دسترسی به این اسناد ممکن است مجاز شود و تصمیم‌گیری در این زمینه باید با رعایت تدابیر حفاظتی تعریف شده در قوانین حفاظت از داده‌ها و مستندسازی کامل توسط آرشیو انجام گیرد. (The National Archives, 2018)

نگرانی‌های اضافی مربوط به قوانین کپی‌رایت نیز بر نحوه پردازش و دسترس‌پذیری منابع آرشیوی تأثیرگذار هستند. اجرای قوانین کپی‌رایت در مورد منابع دیجیتال متولد شده در نهادهای GLAM موضوعی پیچیده و در حال تحول است. بسیاری از این مجموعه‌ها شامل محتوایی هستند که متعلق به چندین صاحب حق کپی‌رایت است، با سطوح مختلفی

1. Digital Preservation Coalition

۲. آرشیو ملی بریتانیا (The National Archives) فرضیاتی درباره طول عمر افراد دارد که رویکرد آن‌ها را در مواجهه با داده‌های شخصی، زمانی که فرد مورد نظر به‌طور مستقیم شناخته شده نیست، تعیین می‌کند. در این سند آمده است: «با توجه به تعداد زیاد افرادی که معمولاً در مجموعه‌های آرشیوی حضور دارند، خدمات آرشیوی قادر نخواهند بود تشخیص دهند که آیا این افراد هنوز زنده هستند یا خیر. در صورتی که مشخص نباشد فرد موضوع داده زنده است یا در گذشته، می‌توان از فرضیات کاری زیر استفاده کرد: فرض شود که طول عمر فرد ۱۰۰ سال است؛ اگر سن فرد بزرگسال مشخص نباشد، فرض شود که در زمان ثبت سند ۱۶ ساله بوده است؛ اگر سن کودک مشخص نباشد، فرض شود که در زمان ثبت سند کمتر از ۱ سال داشته است. اگر مشخص شود که فرد بیش از ۱۰۰ سال سن دارد و هنوز زنده است، رعایت قوانین حفاظت از داده‌ها همچنان الزامی است. چنین فردی حق دارد درخواست دسترسی به داده‌های مربوط به خود را ارائه دهد یا از سایر حقوق قانونی خود استفاده کند.» (آرشیو ملی بریتانیا، ۲۰۱۸، صفحه ۳۴).

از توصیف و اغلب بدون مدرک رضایت. برای مثال، یک وبسایت آرشیو شده ممکن است شامل اطلاعات شخصی، تصاویر، نوشته‌های خلاقانه یا ویدیوهایی باشد که در آن گنجانده شده‌اند. نمی‌توان فرض کرد که تمام محتوا متعلق به فرد یا سازمان تولیدکننده آن است و اغلب امکان شناسایی صاحب حق کپی‌رایت وجود ندارد. بسیاری از نهادهای GLAM که دارای آرشیوهای وب هستند، با تصمیم‌گیری‌های دشواری مواجه‌اند درباره اینکه چه میزان از مجموعه‌هایشان را می‌توانند در دسترس قرار دهند، درحالی‌که باید خطرات احتمالی ناشی از این دسترسی را نیز مدیریت کنند. (Hockx - Vlassenroot et al., 2019; Yu, 2014) قوانین کپی‌رایت در بریتانیا و اتحادیه اروپا، دیجیتال سازی منابع فیزیکی را به عنوان نوعی حفظ در نهادهای GLAM به رسمیت شناخته‌اند و معافیت‌هایی برای حمایت از این فرآیند ایجاد کرده‌اند. بالین حال، مدیریت منابع دیجیتال متولد شده در زمینه حقوقی تا حد زیادی نادیده گرفته شده است. ایجاد نسخه‌های حفظ و دسترسی از محتوای born-digital همچنان یک حوزه خاکستری قانونی محسوب می‌شود. (Koscik & Myska, 2019)

با توجه به بسته بودن بسیاری از مجموعه‌های دیجیتال به دلیل مسائل مربوط به حفاظت داده‌ها و کپی‌رایت، هوش مصنوعی به عنوان راه‌حلی بالقوه برای مشکل آرشیوهای قفل شده مطرح شده است. (Jaillant & Caputo, 2022; Jaillant 2022a, 2022b) در این مقاله، هوش مصنوعی به صورت گسترده تعریف می‌شود، به جای تمرکز بر زیرشاخه‌های خاص مانند یادگیری ماشین، بینایی رایانه‌ای یا پردازش زبان طبیعی، AI به عنوان مجموعه‌ای از رویکردهای محاسباتی برای تصمیم‌گیری و انجام وظایف در نظر گرفته شده است. اگرچه اغلب در قالب فناوری‌های نوظهور مورد بحث قرار می‌گیرد، هوش مصنوعی سابقه‌ای طولانی دارد، با دوره‌هایی از هیجان و دوره‌هایی از بی‌توجهی. (Agar, 2020; Jackson, 2019) دوره کنونی با توسعه پایدار مشخص می‌شود و فناوری‌های AI به طور فزاینده‌ای در تمامی جنبه‌های زندگی روزمره ما حضور دارند. نقش فناوری‌های هوش مصنوعی در دسترس پذیرتر و قابل استفاده تر ساختن آرشیوهای دیجیتال متولد شده به طور فزاینده‌ای به عنوان یک مسیر مهم برای تولیدکنندگان و مدیران اسناد (از جمله کارمندان دولتی مسئول مدیریت دانش، اطلاعات و اسناد)، متخصصان GLAM و پژوهشگرانی که علاقه‌مند به استفاده از منابع آرشیوی هستند، شناخته می‌شود.

نمونه‌هایی از ابزارهای هوش مصنوعی که برای افزایش دسترسی و قابلیت استفاده از آرشیوهای دیجیتال به کار گرفته شده‌اند عبارت‌اند از: طبقه‌بندی خودکار حساسیت اسناد دولتی پیش از انتشار از طریق درخواست‌های آزادی اطلاعات یا انتقال به آرشیوهای ملی. (Baron et al., 2020; McDonald et al., 2020a; The National Archives UK, 2016) آزمایش‌هایی با بینایی رایانه‌ای برای بهبود کشف مجموعه‌های تصویری آرشیوی. (Angelova et al., 2020) ابزارهای جستجوی الگوریتمی. (Nix & Decker, 2021) علاوه بر این، شبکه‌های پژوهشی مانند AEOLIAN، AURA و AI4LAM، ذی‌نفعان مختلفی را گرد هم می‌آورند که به کاربرد هوش مصنوعی در مجموعه‌های آرشیوی علاقه‌مند هستند.

با افزایش آگاهی عمومی نسبت به فناوری‌های AI، نظارت عمومی نیز افزایش یافته و پرسش‌های اخلاقی از جمله شفافیت و سوگیری وارد گفتمان‌های عمومی شده‌اند. (Kelley et al., 2021; Cave et al., 2019; Fast & Horvitz, 2017) درخواست‌ها برای توسعه استانداردهای گسترده‌تر و فرآیندهای اجرایی برای الگوریتم‌هایی که پایه بسیاری از فناوری‌های AI هستند، رو به افزایش است. پروژه‌هایی مانند Assembling Accountability تلاش داشته‌اند تا چارچوب‌هایی برای ارزیابی تأثیرات الگوریتمی هنگام استفاده از ابزارهای مبتنی بر AI توسعه دهند. این پروژه استدلال می‌کند که ارزیابی تأثیرات، «مدلی امیدوارکننده برای حکمرانی الگوریتمی است، زیرا گزارشی از آسیب‌های بالقوه و واقعی یک سامانه را با سازوکاری برای شناسایی مسئولیت جبران آن‌ها ترکیب می‌کند.» (Moss et al., 2021, p. 1)

کارمندان دولتی، متخصصان بخش GLAM (گالری‌ها، کتابخانه‌ها، آرشیوها و موزه‌ها) و پژوهشگران دانشگاهی به طور فزاینده‌ای با مسائل اخلاقی مرتبط با سامانه‌های هوش مصنوعی درگیر هستند. در سال ۲۰۱۸م، دولت بریتانیا «مرکز اخلاق

داده و نوآوری^۱ را تأسیس کرد تا مشاوره‌ای مستقل درباره تدابیر لازم برای تضمین استفاده‌ای ایمن، اخلاقی و نوآورانه از هوش مصنوعی ارائه دهد. یکی از مأموریت‌های اصلی این مرکز، کمک به نهادهای بخش عمومی برای استفاده از داده‌ها و هوش مصنوعی به گونه‌ای است که اعتماد عمومی را جلب کرده و حفظ کند.^۲ در اکتبر ۲۰۲۰م، دفتر علوم و فناوری پارلمان بریتانیا یادداشتی درباره «یادگیری ماشین قابل تفسیر» منتشر کرد تا اطمینان حاصل شود که سامانه‌های یادگیری ماشین به صورت اخلاقی و مسئولانه طراحی و اجرا می‌شوند. (Christie, 2020)

در بخش GLAM، مجموعه‌ای نوظهور از پژوهش‌ها به موضوعاتی همچون شفافیت، پاسخ‌گویی و سوگیری در کتابخانه‌ها و آرشیوها (Padilla, 2019; Cordell, 2020) و همچنین در موزه‌ها (Murphy & Villaespesa, 2020) پرداخته‌اند. این آثار به بررسی گفتمان عمومی پیرامون کاربردهای اخلاقی هوش مصنوعی می‌پردازند. با این حال، چنین تلاش‌هایی به اندازه کافی با ملاحظات اخلاق حرفه‌ای در میان کارمندان دولتی، متخصصان بخش GLAM و پژوهشگران دانشگاهی پیوند نخورده‌اند. پرسش‌های اخلاقی مرتبط با فناوری‌های نوین به تدریج با کدهای اخلاقی‌ای تلاقی پیدا می‌کنند که بر رفتار حرفه‌ای در حوزه‌های مختلف نظارت دارند. کارمندان دولتی، متخصصان بخش GLAM (گالری‌ها، کتابخانه‌ها، آرشیوها و موزه‌ها) و پژوهشگران دانشگاهی همگی به نوعی از رویه‌های اخلاقی پایبند هستند، خواه از طریق کدهای اخلاقی تثبیت شده مانند کد خدمات دولتی^۳ یا کد اخلاقی شورای بین‌المللی آرشیوها^۴ یا از طریق فرآیندهای تأیید اخلاقی که در پژوهش‌های دانشگاهی رایج هستند. علاوه بر رویه‌های اخلاقی نهادینه شده در خدمات دولتی، بخش GLAM و حوزه‌های دانشگاهی، کاربرد فناوری‌های هوش مصنوعی نیز مشمول مجموعه‌ای از کدها و دستورالعمل‌های اخلاقی است، اگرچه این دستورالعمل‌ها کمتر تثبیت شده‌اند و به طور رسمی میان شرکت‌های خصوصی، مؤسسات پژوهشی و نهادهای عمومی توافق نشده‌اند (Jobin et al., 2019; Mittelstadt, 2019; Hagendorff, 2020)

یک مطالعه نشان داد که تاکنون ۸۴ مجموعه اصول یا دستورالعمل اخلاقی برای هوش مصنوعی در سطح جهانی منتشر شده‌اند. (Jobin et al., 2019). با این حال، این چارچوب‌های موجود فاقد اقدامات مشخص و هماهنگی میان بخش‌های حرفه‌ای مختلف هستند. اگرچه مجموعه‌ای واحد و مورد توافق جهانی از اصول اخلاقی در رابطه با هوش مصنوعی وجود ندارد، مفاهیمی همچون شفافیت، عدالت و انصاف، عدم آسیب‌رسانی، مسئولیت‌پذیری و حفظ حریم خصوصی از جمله موضوعاتی هستند که بیشترین بحث‌ها پیرامون آن‌ها صورت گرفته است. (Jobin et al., 2019: 391)^۵ این پنج اصل همچنین در کدهای اخلاقی و دستورالعمل‌های حرفه‌ای که توسط مصاحبه‌شوندگان ما دنبال می‌شوند نیز قابل مشاهده‌اند. ارزیابی میزان اثربخشی این کدها و دستورالعمل‌های اخلاقی خارج از حوزه این مقاله است. با این حال، همان‌طور که گابریل اشاره می‌کند، ما باید «با دقت بیشتری به اصولی توجه کنیم که از اجماع جهانی و هم‌پوشان دیدگاه‌ها

1. Centre for Data Ethics and Innovation

۲. <https://www.gov.uk/government/organisations/centre-for-data-ethics-and-innovation/about> (دسترسی در تاریخ ۲۳ آوریل ۲۰۲۲).

۳. کد خدمات دولتی را می‌توان در نشانی زیر یافت:

<https://www.gov.uk/government/publications/civil-service-code/the-civil-servicecode> (دسترسی در تاریخ ۲۳ آوریل ۲۰۲۲)

۴. کد اخلاقی شورای بین‌المللی آرشیوها (ICA) به طور رسمی در مجمع عمومی این شورا در ۱۹۹۶م تصویب شد. این کد در دسترس است از طریق: <https://www.ica.org/en/ica-code-ethics> (دسترسی در تاریخ ۲۳ آوریل ۲۰۲۲)

۵. دستورالعمل‌های مرتبط شامل اصول FAIR در زمینه قابلیت یافتن، دسترسی‌پذیری، تعامل‌پذیری و استفاده مجدد هستند که با هدف تسهیل استفاده مجدد از داده‌ها تدوین شده‌اند (Wilkinson et al., 2016) و همچنین اصول CARE شامل منفعت جمعی، اختیار در کنترل، مسئولیت‌پذیری و اخلاق که برای حاکمیت داده‌های بومی طراحی شده‌اند: <https://www.gida-global.org/care> (دسترسی در تاریخ ۲ سپتامبر ۲۰۲۲)

پشتیبانی می‌کنند» تا بتوانیم از هوش مصنوعی به درستی بهره ببریم. (Gabriel, 2020: 433)

کار کردن بر اساس اصول مشترک که در قالب کدهای اخلاقی بیان شده‌اند، نقطه شروعی مؤثر برای ایجاد اعتماد در کاربرد هوش مصنوعی در آرشیوهای دیجیتال متولد شده به شمار می‌رود. مفاهیم اخلاقی رایج مانند شفافیت، عدالت و انصاف، عدم آسیب‌رسانی، مسئولیت‌پذیری و حفظ حریم خصوصی که توسط جابین و همکاران مطرح شده‌اند، ذاتاً با اصول بیان شده در کدهای اخلاقی حرفه‌ای و رویه‌های موجود در خدمات دولتی، بخش GLAM و پژوهشگران دانشگاهی در تضاد نیستند. کد خدمات دولتی ارزش‌هایی همچون صداقت، راستی، عینیت و بی‌طرفی را به عنوان اصول کلیدی برای کارمندان دولتی مشخص کرده است و این اصول در تمامی جنبه‌های نحوه عملکرد حرفه‌ای آن‌ها نفوذ دارد. تعدادی از این اصول از طریق قوانین و مقرراتی مانند قانون آزادی اطلاعات و قانون اسناد عمومی پشتیبانی می‌شوند؛ قوانینی که فرهنگ پاسخگویی را در اسنادی که کارمندان دولتی و دارندگان مناصب عمومی روزانه تولید می‌کنند، نهادینه می‌سازند. علاوه بر این، کدهای اخلاقی و رویه‌های حرفه‌ای نقش مهمی در حرفه‌ای‌سازی بخش GLAM ایفا کرده‌اند. در ۱۹۸۶م، شورای بین‌المللی موزه‌ها (ICOM) کد اخلاقی خود را منتشر کرد.^۱ ده سال بعد، شورای بین‌المللی آرشیوها (ICA) چارچوب اخلاقی خود را معرفی کرد. همچنین، فدراسیون بین‌المللی انجمن‌های کتابداری (IFLA) در ۲۰۱۲م یک کد اخلاقی را تصویب نمود.^۲ این کدهای اخلاقی اغلب با کدهای خاص هر کشور تکمیل می‌شوند و بریتانیا نیز دارای کدهای اخلاقی از سوی انجمن موزه‌های بریتانیا و CILIP (انجمن کتابداری و اطلاعات بریتانیا) است.^۳ صداقت، بی‌طرفی، عینیت، حفظ حریم خصوصی و همکاری حرفه‌ای از جمله اصولی هستند که آرشیویست‌ها در کار خود به آن‌ها پایبند هستند. علاوه بر کدهای اخلاقی که برای کارمندان دولتی و متخصصان بخش GLAM تدوین شده‌اند، پژوهشگران دانشگاهی نیز در چارچوب‌های اخلاقی مشخصی فعالیت می‌کنند. اخلاق در پژوهش‌های دانشگاهی تحت مجموعه‌ای از اصول اخلاقی اداره می‌شود که در راهبردهای هر دانشگاه و فرایندهای تأیید پژوهش ادغام شده‌اند. دفتر یکپارچگی پژوهشی بریتانیا^۴، مسئول تدوین و نظارت بر کد رفتار پژوهشگران در بریتانیا است.^۵ دانشگاه لافبرو که میزبان پروژه Unlocking our Digital Past است، دارای چارچوب سیاست اخلاقی‌ای است که بر اساس کد رفتار UKRIO تدوین شده و با هدف ادغام ملاحظات اخلاقی در شیوه‌های پژوهشی طراحی شده است. این مقاله ادعا نمی‌کند که تمامی متخصصان فعال در خدمات دولتی، بخش GLAM و حوزه دانشگاهی، پیش از اتخاذ تصمیمات جدید یا انجام وظایف تازه، حتماً به کدهای اخلاقی یا آیین‌نامه‌های رفتاری خود مراجعه می‌کنند، بلکه بیان می‌کند که این کدها به‌طور کلی بر اصول و حوزه‌های مورد توجه در این بخش‌ها تأثیرگذار هستند.

مفاهیمی همچون اعتماد، شفافیت و اخلاق حرفه‌ای به عنوان موضوعات تکرارشونده در سراسر پروژه Unlocking our Digital Past مطرح شدند. این پروژه بخشی از همکاری مستمر میان دانشگاه لافبرو و دفتر کابینه دولت بریتانیا بود که

۱. کد اخلاقی ICOM از طریق نشانی زیر در دسترس است: <https://icom.museum/en/resources/standards-guidelines/code-of-ethics> (دسترسی در تاریخ ۲۳ آوریل ۲۰۲۲)

۲. کد اخلاقی IFLA از طریق نشانی زیر در دسترس است: <https://www.ifla.org/faife/professional-codes-of-ethics-for-librarians> (دسترسی در تاریخ ۲۵ آوریل ۲۰۲۲)

۳. کد اخلاقی انجمن موزه‌ها از طریق نشانی زیر در دسترس است: <https://www.museumsassociation.org/campaigns/ethics/> <https://www.cilip.org.uk/page/ethics> (دسترسی در تاریخ ۲۳ آوریل ۲۰۲۲) و کد اخلاقی CILIP نیز از طریق نشانی زیر قابل مشاهده است: <https://www.cilip.org.uk/page/ethics> (دسترسی در تاریخ ۲۵ آوریل ۲۰۲۲)

۴. UK Research Integrity Office UKRIO

۵. کد رفتار پژوهشگران UKRIO از طریق نشانی زیر در دسترس است: <https://ukrio.org/publications/code-of-practice-for-research-introduction/> (دسترسی در تاریخ ۲۵ آوریل ۲۰۲۲)

بر کاربرد هوش مصنوعی در آرشیوهای دیجیتال متولدشده تمرکز داشت. پروژه مجموعه‌ای از کارگاه‌ها را برگزار کرد که هدف آن افزایش گفت‌وگوی میان‌رشته‌ای میان کارمندان دولتی، متخصصان بخش GLAM و دانشگاهیان بود. پس از دریافت تأییدیه از کمیته اخلاقی مؤسسه، تیم ما از متخصصان علوم انسانی دیجیتال، سی مصاحبه با اندیشمندان و فعالان کلیدی انجام داد که در دسترسی‌پذیری و قابلیت استفاده از آرشیوهای دیجیتال متولدشده، اکنون و در آینده، نقش دارند.

زمینه تخصصی ما در تاریخ فرهنگی، مطالعات ادبی و مطالعات موزه است، نه علوم رایانه و همین امر بر ماهیت پرسش‌های ما تأثیر گذاشت. از این رو، به جای تمرکز بر جنبه‌های فنی هوش مصنوعی، پرسش‌های ما بر موانع اصلی در دسترس‌سازی مجموعه‌های دیجیتال و دیجیتال‌شده و راه‌حل‌های ممکن برای این مسائل متمرکز بود. ما به شیوه‌های روزمره حرفه‌ای‌ها علاقه‌مند بودیم، از تولیدکنندگان اسناد که باید پیش از انتقال به آرشیو ملی و سایر مخازن، اسناد آرشیوی را آماده کنند تا آرشیویست‌هایی که باید این اسناد را حفظ و قابل دسترس نمایند. همچنین با تاریخ‌نگارانی مصاحبه کردیم که برای پژوهش‌هایشان به آرشیوها (از جمله آرشیوهای دیجیتال متولدشده) نیاز دارند. رشته مشترک این مصاحبه‌ها، کمبود اعتماد به سایر ذی‌نفعان و به فناوری بود. این بی‌اعتمادی به نوبه خود بر دسترس‌پذیری و قابلیت استفاده از آرشیوهای دیجیتال تأثیر می‌گذارد.

بی‌اعتمادی به سایر ذی‌نفعان و به فناوری

گفت‌وگوهای مشترک میان متخصصان دولت، بخش GLAM و دانشگاه معمولاً استثنا هستند، نه قاعده. این وضعیت بر نحوه نگاه حرفه‌ای‌ها به یکدیگر تأثیر می‌گذارد و اغلب منجر به سوء تفاهم می‌شود. جیمز بیکر که در رشته تاریخ تحصیل کرده و پیش از ورود به دانشگاه به عنوان متصدی در کتابخانه بریتانیا فعالیت داشته، مثالی از نبود گفت‌وگو میان آرشیویست‌ها و پژوهشگران ارائه می‌دهد. درحالی‌که آرشیویست‌ها ارزیابی و گزینش اسناد را بخش مرکزی نقش خود می‌دانند، پژوهشگران اغلب انتظار دارند که آرشیوها کامل باشند. به عبارت دیگر، دانشگاهیان آرشیویست‌ها را صرفاً نگهدارنده اسناد می‌دانند، نه متخصصانی که باید اسناد فاقد ارزش ماندگار را حذف کنند. این وظیفه ارزیابی به‌ویژه در مواجهه با حجم عظیم آرشیوهای دیجیتال متولدشده بسیار پیچیده‌تر شده است. بیکر می‌گوید: «همیشه با دوستان آرشیویستم همدردی می‌کنم، هر بار که تاریخ‌نگاران می‌گویند: خب، شما باید همه چیز را جمع‌آوری کنید، چرا تلاش نمی‌کنید؟ ما به همه این چیزها نیاز داریم و قرار است وارد یک عصر تاریک دیجیتال شویم و آن‌ها می‌گویند: نه، قرار نیست عصر تاریک دیجیتال باشد، ما همیشه گزینش می‌کنیم، فقط مسئله این است که موضوع گزینش حالا در عصر دیجیتال متولدشده بسیار پیچیده‌تر شده است.»^۱

میان مصاحبه‌شوندگانی که به‌طور نزدیک با اسناد دیجیتال متولدشده سروکار دارند، اجماعی وجود داشت: نگهداری و توصیف همه چیز گزینه‌ای عملی نیست. کلیفورد لینچ، مدیر اجرایی ائتلاف اطلاعات شبکه‌ای، اشاره کرد که در روزهای آغازین اینترنت، تلاش‌هایی برای اعمال شیوه‌های فهرست‌نویسی کتابخانه‌ای بر منابع اینترنتی صورت گرفت، اما این تلاش‌ها «خیلی زود کنار گذاشته شدند، زیرا مقیاس آن‌ها به‌هیچ‌وجه قابل مدیریت نبود».^۲ رویکرد محاسباتی به سبک گوگل در جست‌وجوی اسناد، به هنجار رایج برای جست‌وجو در مقیاس وسیع تبدیل شده است. مسئله مقیاس به‌ویژه در مورد اسناد دیجیتال دولتی بسیار حاد است. جیسون بارون که به مدت سیزده سال مدیر بخش دعاوی در اداره ملی آرشیوها و اسناد ایالات متحده (NARA) بود و سپس به بخش خصوصی و دانشگاهی منتقل

۱. مصاحبه با جیمز بیکر، ۲۲ ژوئن ۲۰۲۲.

۲. مصاحبه با کلیفورد لینچ، ۱۷ ژوئن ۲۰۲۱.

شد، مثالی از حجم عظیم ایمیل‌های رؤسای جمهور آمریکا ارائه می‌دهد: «آرشیو ملی ایالات متحده در حال حاضر دارای ۵۰۰ میلیون ایمیل از دوران ریاست جمهوری ریگان تا اوپاما است. ایمیل‌های دوران ترامپ هنوز شمارش نشده‌اند یا به‌طور کامل پردازش نشده‌اند، اما این مجموعه شامل ۵۰۰ میلیون ایمیل مجزاست و بیش از یک میلیارد صفحه نیز وجود دارد، زیرا بسیاری از این ایمیل‌ها دارای پیوست هستند.»^۱ بارون در این سمت از رویکردی حمایت کرد که هدف آن کاهش تعداد ایمیل‌های دولتی منتقل شده به مخازن آرشیوی بود. این رویکرد با عنوان CAPSTONE در ۲۰۱۳ م اتخاذ شد و منجر به سیاستی در زمینه حفظ ایمیل‌ها شد که بر اساس میزان ارشدیت تولیدکننده سند تنظیم شده بود.^۲ افرادی که در رأس یا نزدیک به رأس یک نهاد قرار دارند، تمام ایمیل‌هایشان به‌طور دائمی حفظ می‌شود، درحالی‌که ایمیل‌های سایر افراد پس از مدت مشخصی (معمولاً هفت سال) حذف می‌گردد. CAPSTONE به مسئله افزایش نمایی حجم ایمیل‌ها پاسخ داد، اما ایمیل‌ها البته تنها نوع اسناد دیجیتالی متولد شده نیستند که توسط دولت ایالات متحده تولید می‌شوند.

لِزلی جانستون (مدیر بخش حفظ دیجیتال در NARA) اشاره می‌کند که دولت ترامپ حدود ۵۰۰ ترابایت سند آرشیوی تولید نموده، درحالی‌که این رقم برای دو دوره ریاست جمهوری اوپاما حدود ۲۰۰ ترابایت بوده است.^۳ کتابخانه‌ها و مؤسسات آرشیوی با کمبود نیروی انسانی مواجه هستند و توانایی لازم برای مدیریت حجم عظیم داده‌ها را ندارند. جیسون بارون مثالی از این وضعیت ارائه می‌دهد: کتابخانه ریاست جمهوری کلینتون تنها شش آرشیویست دارد، درحالی‌که با صافی شامل حدود ۱۰ میلیون سند برای پاسخ‌گویی به درخواست‌های قانون آزادی اطلاعات (Freedom of Information Act) روبه‌روست.^۴ این وضعیت در کوتاه مدت منجر به فقدان پاسخ‌گویی می‌شود و در بلندمدت خطر تأثیرگذاری منفی بر حافظه فرهنگی را به همراه دارد.

جیسون بارون و دیگران استدلال کرده‌اند که اتومات سازی یک ضرورت است، نه یک انتخاب. بارون می‌پرسد: «چطور می‌خواهید یک میلیارد شیء را جست‌وجو کنید؟» و ادامه می‌دهد: «خب، شما این کار را به صورت دستی، به عنوان یک آرشیویست انجام نمی‌دهید و با جست‌وجوی کلیدواژه‌ای هم انجام نمی‌دهید - اگرچه می‌توانید امتحان کنید - اما آن‌ها به شدت ناکارآمد هستند.»^۵ درواقع، رویکردی مبتنی بر جست‌وجوی کلیدواژه‌ای در مجموعه‌هایی که با حجم بسیار بالا و فقدان فراداده مشخص می‌شوند، عملکرد مناسبی ندارد. برای نمونه، جست‌وجوی واژه‌ی «Brexit» در آرشیو وب بریتانیا حدود ۶۳,۹۶۸,۸۸۵ نتیجه به دست می‌دهد. این حجم عظیم داده‌ها بدون ساختار مناسب، نه تنها بازایی اطلاعات را دشوار می‌سازد، بلکه نشان‌دهنده محدودیت‌های روش‌های سنتی جست‌وجو در مواجهه با آرشیوهای دیجیتال متولد شده است.^۶ همان‌طور که لئونتین تالبوم (پژوهشگر دکتری در آرشیو ملی بریتانیا و کالج دانشگاهی لندن) به ما گفت، تحلیل دستی هر سند امکان‌پذیر نیست.^۷ در عوض، تنها راه برای فهم این حجم عظیم نتایج، رویکرد محاسباتی است. جیسون بارون می‌پرسد: «چطور می‌توانند در گشودن بخش قابل توجهی از تاریخ قرن بیست و یکم میلادی و قبل از

۱. مصاحبه با جیسون بارون، ۲۲ ژوئن ۲۰۲۱.

۲. این منبع از طریق نشانی زیر در دسترس است: <https://www.archives.gov/records-mgmt/bulletins>؛ ۲۰۱۳/۲۰۱۳-۰۲.html (دسترسی در تاریخ ۲۳ آوریل ۲۰۲۲)

۳. مصاحبه با لِزلی جانستون، ۱۸ ژوئن ۲۰۲۱.

۴. مصاحبه با جیسون بارون، ۲۲ ژوئن ۲۰۲۱.

۵. همان.

۶. <https://www.webarchive.org.uk> (دسترسی در تاریخ ۲۳ آوریل ۲۰۲۲)

۷. مصاحبه با لئونتین تالبوم، ۲۲ ژوئن ۲۰۲۱.

آن موفق باشند، اگر از تکنیک‌های یادگیری ماشین برای انجام جست‌وجوها استفاده نکنند؟^۱

باید توجه داشت که جست‌وجوهای مبتنی بر هوش مصنوعی راه‌حل‌هایی کامل نیستند، زیرا ممکن است نتایج مرتبط را نیز از قلم بیندازند.^۲ کاربرد هوش مصنوعی و یادگیری ماشین در اسناد آرشیوی موضوع تازه‌ای نیست. در اوایل دهه ۲۰۰۰م، جیسون بارون عضوی از تیم حقوقی دولت بود که در پرونده بزرگی علیه شرکت فیلیپ موریس و سایر شرکت‌های دخانیات مشارکت داشت. اطلاعات کلیدی در آرشیو ایمیل‌های کلینتون پراکنده بود؛ آرشیوی که شامل ۳۰ میلیون ایمیل می‌شد. بارون تصمیم گرفت با متخصصان علوم اطلاعات و علوم رایانه مشورت کند تا راهی بهتر برای جست‌وجوی وکلای حقوقی بیابد؛ روشی غیر از جست‌وجوی کلیدواژه‌ای.^۳

در همان زمان، شرکت‌های تجاری شروع به عرضه نرم‌افزارهای eDiscovery کردند که با بهره‌گیری از یادگیری ماشین، اطلاعات را در حجم عظیمی از داده‌ها جست‌وجو و استخراج می‌کردند. بخش آرشیو در دو سوی اقیانوس اطلس نیز به بررسی ظرفیت‌های eDiscovery برای حل چالش‌های اسناد دیجیتال متولد شده پرداخت. در ۲۰۱۶م، گزارشی از آرشیو ملی بریتانیا (TNA) نتیجه‌گیری کرد که: «بازبینی با کمک فناوری، با استفاده از نرم‌افزار eDiscovery، می‌تواند از ادارات دولتی در فرآیند ارزیابی، گزینش و بررسی حساسیت، به عنوان بخشی از انتقال اسناد دیجیتال متولد شده به TNA، پشتیبانی کند.» (The National Archives UK, 2016, p. 5)

پژوهشگرانی مانند گراهام مک‌دونالد (مدرس بازیابی اطلاعات در دانشگاه گلاسگو) نشان داده‌اند که هوش مصنوعی و یادگیری ماشین واقعاً می‌توانند به گشودن آرشیوهای دولتی کمک کنند و این کار را با استفاده از رویکرد اصلی انجام می‌دهند: «ابتدا محافظت، سپس جست‌وجو» یا «ابتدا جست‌وجو، سپس محافظت».^۴ کار مک‌دونالد عمدتاً بر پایه رویکرد نخست است. ابزارهایی که تیم او توسعه داده‌اند، می‌توانند به انسان‌ها کمک کنند تا حساسیت‌های موجود در یک مجموعه خاص را شناسایی کرده و سپس پیش از انتقال اسناد به مؤسسات آرشیوی، تصمیمات آگاهانه‌ای اتخاذ کنند. برای نمونه، تولیدکنندگان اسناد می‌توانند اطلاعات حساس را پیش از انتقال، سانسور یا حذف کنند. این فرآیند سطح ریسک در دسترسی به آن مجموعه را کاهش می‌دهد.

رویکرد دوم شامل دسترسی کامل به مجموعه است، همراه با چارچوب جست‌وجویی که هنگام مواجهه با اطلاعات حساس، آن‌ها را پنهان می‌سازد. در این روش، ماشین با استفاده از زیرمجموعه‌ای از داده‌ها که به عنوان حساس یا محرمانه شناسایی شده‌اند آموزش می‌بیند و سپس انتظار می‌رود که بتواند حساسیت را در مجموعه‌های داده بزرگ‌تر تشخیص دهد. مشکل اینجاست که تعریف جهانی‌ای برای «حساسیت» وجود ندارد. این مفهوم تا حد زیادی به زمینه و بافت بستگی دارد و ماشین همیشه قادر نیست اطلاعات زمینه‌ای را به درستی تشخیص دهد؛ اطلاعاتی که ممکن است یک سند را حساس کند.^۵ درحالی‌که گراهام مک‌دونالد اذعان دارد که فرآیند بررسی حساسیت هنوز کامل و بی‌نقص نیست، از دیدگاه جیسون بارون، ریسک دسترسی به اسناد آرشیوی ارزش پذیرفتن را دارد. هوش مصنوعی «بازدهی فوق‌العاده‌ای خواهد داشت»، حتی اگر با خطر سوگیری همراه باشد.^۶

مهم‌ترین نکته این است که به پژوهشگران، روزنامه‌نگاران، مورخان و سایر کاربران اجازه داده شود تا کشف‌هایی

۱. مصاحبه با جیسون بارون، ۲۲ ژوئن ۲۰۲۱.

۲. برای اطلاعات بیشتر درباره جست‌وجوی مبتنی بر هوش مصنوعی، بنگرید به Prescott و Winters (۲۰۱۹).

۳. همان.

۴. مصاحبه با گراهام مک‌دونالد، ۲۳ ژوئن ۲۰۲۱.

۵. برای اطلاعات بیشتر درباره بررسی حساسیت، بنگرید به McDonald (۲۰۲۰a، ۲۰۲۰b).

۶. مصاحبه با جیسون بارون، ۲۲ ژوئن ۲۰۲۱.

انجام دهند که در خدمت منافع عمومی باشد. سطح ریسک در این حوزه پایین‌تر از سیستم عدالت کیفری است؛ جایی که هوش مصنوعی در گذشته منجر به نتایج فاجعه‌بار شده است. برای مثال، الگوریتم‌هایی که در ایالات متحده برای ارزیابی احتمال ارتکاب مجدد جرم توسط زندانیان استفاده می‌شدند، دارای سوگیری نژادی بودند. (Wadsworth et al., 2018) با وجود خطرات همراه با هوش مصنوعی، پلیس و سیستم قضایی در دو سوی اقیانوس اطلس، فناوری‌های پیشرفته را ابزاری ضروری در مبارزه با جرم و جنایت می‌دانند. در ۲۰۱۴م، وزارت کشور بریتانیا گزارشی درباره استفاده از eDiscovery در تحقیقات پزشکی قانونی دیجیتال منتشر کرد. (Lawton et al., 2014) چهار سال بعد، توسعه فناوری جدیدی برای شناسایی خودکار محتوای تروریستی در پلتفرم‌های آنلاین را اعلام کرد.^۱ اگر هوش مصنوعی به‌طور متداول توسط افسران پلیس و قضات برای تحلیل مجموعه‌های بزرگ داده به کار گرفته می‌شود، چرا در آرشیوهای دولتی به‌ندرت استفاده می‌شود؟ هنوز هم استفاده از هوش مصنوعی برای بررسی حساسیت یا یافتن نتایج در مجموعه‌های دیجیتال عظیم، استثنایی تلقی می‌شود.

برای آموزش سیستم‌های هوش مصنوعی، دسترسی به داده‌ها ضروری است و همین مفهوم دسترسی - حتی اگر محدود به مجموعه‌های آموزشی باشد - برای بسیاری از تولیدکنندگان اسناد و آرشیویست‌ها مسئله‌ساز است. وقتی به کاربران اعتماد نمی‌شود که تصمیمات درستی بگیرند، منطقی به نظر می‌رسد که کل مجموعه‌ها بسته شوند یا دسترسی فقط به گروه‌هایی داده شود که «قابل اعتماد» تلقی می‌شوند. برای نمونه، آرشیوهای وب معمولاً به دلیل مسائل مربوط به حق نشر، کاربران را مجبور می‌کنند برای دسترسی به محتوا به کتابخانه مراجعه کنند. محتوا فقط از طریق مرورگرهای موجود در محل قابل مشاهده است، نه از طریق دانلود خارج از محل. جین وینترز، استاد علوم انسانی دیجیتال در دانشگاه لندن، اشاره می‌کند که دانمارک استثنایی برای پژوهشگران در این سیاست محدودکننده دارد. پژوهشگرانی که وابسته به دانشگاه‌های دانمارکی هستند و همچنین همکاران خارجی می‌توانند از هر نقطه‌ای در جهان به داده‌های آرشیو وب به‌صورت دورکاری دسترسی داشته باشند. وینترز می‌گوید: «باید تعهدنامه‌هایی امضا کنید و غیره که داده‌ها را به‌درستی مدیریت خواهید کرد.» این را می‌توان گام نخست در جهت «آغاز تفکر درباره گشودن دسترسی‌ها» دانست.^۲

خطر انتشار داده‌های بالقوه حساس باید در برابر خطر پنهان نگه داشتن و غیرقابل دسترس بودن اسناد سنجیده شود. درواقع، دسترسی در مرکز اصول اخلاقی بخش GLAM (موزه‌ها، کتابخانه‌ها، آرشیوها و مراکز میراث فرهنگی) قرار دارد. انجمن موزه‌ها در کد اخلاقی خود، «تعامل عمومی و منفعت عمومی» را به‌عنوان اصل نخست معرفی می‌کند. به‌طور مشخص، این کد بیان می‌کند که موزه‌ها و دیگر سازمان‌های میراث فرهنگی باید «دسترسی عمومی و تعامل معنادار با موزه‌ها، مجموعه‌ها و اطلاعات مربوط به آن‌ها را بدون تبعیض فراهم کنند». به همین ترتیب، انجمن کتابخانه‌ها و اطلاعات بریتانیا (CILIP) اعلام می‌کند که «حفظ و تداوم دسترسی به دانش» در مرکز مأموریت کتابخانه‌ها قرار دارد. همچنین، شورای بین‌المللی آرشیوها (ICA) تأکید می‌کند که «آرشیویست‌ها باید گسترده‌ترین سطح ممکن از دسترسی به مواد آرشیوی را ترویج دهند و خدماتی بی‌طرفانه به همه کاربران ارائه کنند». البته، دسترسی تنها جنبه‌ای نیست که متخصصان GLAM باید در نظر بگیرند. کدهای اخلاقی شامل ارجاعاتی به حریم خصوصی و حق نشر به‌عنوان محدودیت‌هایی برای دسترسی هستند. برای ICA، «آرشیویست‌ها باید هم به دسترسی و هم به حریم خصوصی احترام بگذارند». بااین حال، این تعادل اغلب به سمت حفظ حریم خصوصی و حفاظت از داده‌ها سنگینی می‌کند. درحالی‌که آرشیویست‌ها باید به چارچوب‌های قانونی احترام بگذارند و از داده‌های خصوصی افراد محافظت کنند، آن‌ها می‌توانند

۱. <https://www.gov.uk/government/news/new-technology-revealed-to-help-fight-terrorist-content-online> (دسترسی در

۲۳ آوریل ۲۰۲۲).

۲. مصاحبه با جین وینترز، ۲۴ ژوئن ۲۰۲۱.

ریسک‌های محدودی را بپذیرند و برخی مجموعه‌ها را بدون بررسی کامل همه مواد، در دسترس قرار دهند. نیاز به اعتماد و کدهای مشترک رفتاری و حرفه‌ای در مصاحبه‌ها به‌وضوح مطرح شد. جین وینترز (دانشگاه لندن) مثالی از آرشیوهای شخصی ارائه داد، نه آرشیوهای وب یا مجموعه‌های بزرگ مقیاس. او گفت: «روابط مبتنی بر اعتماد میان نویسندگان سپرده‌گذار، آرشیویست‌ها و پژوهشگران» بسیار مهم است. با این حال، چنین روابطی «زمان‌بر» و «طردکننده» هستند، زیرا «اگر کتابداران را بشناسید، صحبت با آن‌ها بسیار آسان‌تر است تا زمانی که از بیرون وارد شوید». به عبارت دیگر، دسترسی اغلب بر اساس گفت‌وگوهای اولیه با آرشیویست‌ها و به صورت فرد به فرد تصمیم‌گیری می‌شود. داشتن این مذاکرات و ساختن روابط مبتنی بر اعتماد پیرامون استفاده از این مواد می‌تواند مجموعه‌هایی را که بیشتر بسته بودند، در دسترس قرار دهد.^۱

اظهارات وینترز بازتابی از تجربه پژوهشگران ادبی است که به مجموعه‌هایی دست‌یافته‌اند که پیشتر بسته بودند. لیز ژایان درباره تجربه خود در دسترسی به آرشیو ایمیل نویسنده بریتانیایی ایان مک‌یون در مرکز هری رنسوم در تگزاس نوشته است؛ آرشیوی که به طور رسمی برای پژوهشگران غیرقابل دسترس است. دسترسی از طریق ارتباطات شخصی و به دلیل مشارکت آرشیویست مرکز HRC در همان پروژه پژوهشی فراهم شد. (Jaillant, 2019) ارائه دسترسی موردی به کاربران شناخته شده، راه حل بلندمدتی نیست، زیرا گروه‌های کاملی را کنار می‌گذارد، همچون پژوهشگران مستقل یا تاریخ‌نگاران خانوادگی که وابستگی دانشگاهی یا شبکه‌های حرفه‌ای ندارند. همان‌طور که ذکر شد، دسترسی بدون تبعیض در مرکز اصول اخلاقی بخش GLAM قرار دارد. آدام نیکس، مدرس کسب و کار مسئولانه در دانشگاه بیرمنگام، طرفدار سیاستی نظام‌مند برای دسترسی است که بر پایه احترام به مجموعه‌ای از رفتارهای اخلاقی بنا شده باشد. او می‌گوید: «دوست دارم تمرکز بیشتری بر تقسیم مسئولیت مراقبت میان آرشیوها و کاربران بینم. فکر می‌کنم کاربران باید نقش فعال‌تر و آگاهانه‌تری در حفظ یکپارچگی فرآیند کشف آرشیوی ایفا کنند و فکر می‌کنم این کار با داشتن کدهای رفتاری توسعه‌یافته و محترم‌شده‌ای ممکن است که کاربر از پیش با آن‌ها موافقت آگاهانه داشته باشد به این معنا که حتی اگر اطلاعات حساسی را ببیند، آن اطلاعات از آن فرد پژوهشگر فراتر نرود.»^۲

این فرآیند پیشنهادی برای گشودن آرشیوها به روی کاربران اخلاق‌مدار، بر پایه اعتماد است، اما بر پایه امکان اعمال مجازات‌ها (از جمله مجازات‌های قانونی) برای کسانی که از کدهای رفتاری تخطی می‌کنند. سیستمی مشابه از اعتماد و امکان مجازات در پژوهش‌های فرانسیس هریس و فرگوس لیون درباره همکاری میان فرهنگ‌های حرفه‌ای توصیف شده است. یافته‌های اصلی آن‌ها عبارت‌اند از: اول، اعتماد عنصر حیاتی در همکاری میان رشته‌ها و بخش‌هاست. دوم، حفظ همکاری میان فرهنگ‌های حرفه‌ای، چالش‌های خاصی برای ایجاد اعتماد به همراه دارد. سوم، اعتماد در تیم‌های میان‌رشته‌ای و فرارشته‌ای بر پایه هنجارها، اطلاعات، مجازات‌ها و کنترل‌ها بنا شده است. (Harris and Lyon, 2013)

جان شریدن، مدیر دیجیتال آرشیو ملی بریتانیا (TNA)، طرفدار کمی‌سازی خطر (ریسک) و در نتیجه اعتماد است. او می‌گوید: «وقتی به فرآیندهایمان در زمینه دیجیتال‌سازی نگاه می‌کنیم، همه بیشتر بر مدیریت ریسک از طریق دانش تخصصی تکیه دارند تا بر سیستم‌ها.»^۳ شریدن از رویکردی محاسباتی با استفاده از مدل‌های آماری برای ارزیابی ریسک حمایت می‌کند. در همین راستا، TNA به بررسی شبکه‌های بیزی پرداخته است؛ مدلی گرافیکی که مجموعه‌ای از متغیرها و شرایط را نمایش می‌دهد و اغلب برای تحلیل احتمالات به کار می‌رود. (Barons et al., 2021) از دیدگاه شریدن، مؤسسات آرشیوی باید تعادلی مناسب میان ریسک و دسترسی به مواد بالقوه حساس برقرار کنند. این موضوع صرفاً

۱. مصاحبه با جین وینترز، ۲۴ ژوئن ۲۰۲۱.

۲. مصاحبه با آدام نیکس، ۲۲ ژوئن ۲۰۲۱.

۳. مصاحبه با جان شریدن، ۲۹ ژوئن ۲۰۲۱.

«شفافیت به هر قیمتی نیست»، زیرا «آرشیوها ویکی لیکس نیستند و ما در کار ویکی لیکس نیستیم [...]». این کار مسئولانه نیست، نه در برابر سوژه‌های داده، نه در برابر حقوق مالکیت فکری دیگران و نه از نظر قانونی. بنابراین، باید تکنیک‌هایی برای ارائه دسترسی به صورت مسئولانه توسعه دهیم.»

تا زمانی که مدل‌های نظام‌مند ارزیابی ریسک به‌طور ذاتی در رویه‌های دیجیتال آرشیوها جای نگرفته‌اند، جان شریدن پیشنهاد می‌کند از رویکردی استفاده شود که آن را «درجه‌بندی دسترسی» می‌نامد. این فرآیند که هدف آن «حداکثرسازی استفاده و درعین حال مدیریت ریسک‌های انتشار» است، می‌تواند به سیستمی منعطف‌تر برای انتشار اسناد بالقوه حساس منجر شود - بدون نقض قوانین حفاظت از داده‌ها - از طریق شناسایی سطوح مختلف ریسک و تعیین استثنای لازم^۱. با افزایش تعداد اسناد دیجیتال متولد شده و محدودیت ظرفیت ارزیابی تخصصی، فناوری‌های هوش مصنوعی می‌توانند به مدیریت و کاهش ریسک‌ها کمک کنند. چندین مصاحبه‌شونده اظهار داشتند که اگرچه ابزارهای هوش مصنوعی می‌توانند از دسترسی و استفاده از آرشیوهای دیجیتال پشتیبانی کنند، اما هنوز موانعی برای گسترش آزمایش‌ها و پذیرش گسترده این فناوری‌ها وجود دارد. این نگرانی‌ها را می‌توان به‌طور کلی در سه دسته طبقه‌بندی کرد: پاسخگویی، کنترل و سوگیری. پیشنهاد می‌شود که این نگرانی‌ها از طریق همکاری نزدیک‌تر میان کارمندان دولتی، متخصصان GLAM و دانشگاهیان - در حوزه‌های علوم انسانی و علوم کامپیوتر - برطرف شوند.

پاسخگویی

نیاز به پاسخگویی و شفافیت درباره فعالیت‌های دولت، بخش اساسی از کدهای رفتاری و اصول اخلاق حرفه‌ای است که کارمندان دولتی بر اساس آن‌ها فعالیت می‌کنند. پاسخگویی و شفافیت همچنین در قانون اسناد عمومی نهادینه شده‌اند. ابزارهایی که برای تصمیم‌گیری درباره انتقال اسناد دولتی به آرشیو ملی استفاده می‌شوند، مشمول همین اصل پاسخگویی هستند.^۲ جیمز لاپین، مسئول دوران دیجیتال در دفتر کابینه^۳، توضیح می‌دهد که اگر تصمیمی برای استفاده از ابزار هوش مصنوعی گرفته شود، آنگاه «افرادی که آن را به کار می‌گیرند، مسئول نتایج هستند» و «شما باید از الگوریتم خود دفاع کنید [...] از فرآیند آموزش الگوریتم و فرآیند نظارت بر آن».^۴ تردید نسبت به اجرای فناوری‌های هوش مصنوعی اغلب با ترس از نتایج غیرمنتظره یا ناخواسته و مسئولیت‌پذیری در قبال آن‌ها مرتبط است. کلیفورد لینچ (ائتلاف اطلاعات شبکه‌ای) اشاره می‌کند که اغلب الگوریتم به اشتباه مقصر شناخته می‌شود: «من از الگوریتم عصبانی نمی‌شوم، از کسی که آن را نوشته هم نه، من از کسانی ناراضی‌ام که آن را به کار می‌گیرند، درحالی‌که باور دارند دارد کار درست را انجام می‌دهد، بدون اینکه بفهمند واقعاً با چه چیزی سروکار دارند».^۵ اظهارات لینچ بر نیاز به درک و مسئولیت‌پذیری کسانی که ابزارهای هوش مصنوعی را اجرا می‌کنند تأکید دارد و همچنین بر تصور اشتباهی که الگوریتم را مقصر می‌داند.

اندرو بلیک، رئیس دپارتمان اقتصاد سیاسی در کینگز کالج لندن، یادآوری می‌کند که جنجال پیرامون تخصیص

۱. همان.

۲. KIM اشاره دارد به مدیریت دانش، اطلاعات و اسناد.

۳. KIRM

۴. مصاحبه با جیمز لاپین، ۲۹ ژوئن ۲۰۲۱.

۵. مصاحبه با کلیفورد لینچ، ۱۷ ژوئن ۲۰۲۱.

نمرات A-Level [۱] با استفاده از هوش مصنوعی در ۲۰۲۰م ابتدا به الگوریتم نسبت داده شد، اما گفتمان عمومی این دیدگاه را رد کرد و خواستار پاسخگویی در قبال اجرای الگوریتم شد.^۲ این نیاز به پاسخگویی در سایر مصاحبه‌ها نیز از طریق تأکید بر ضرورت هوش مصنوعی قابل توضیح^۳ مطرح شد. هوش مصنوعی قابل توضیح توسط هفت نفر از مصاحبه‌شوندگان به عنوان راهی برای تضمین پاسخگویی و شفافیت در ابزارهایی که توسط خدمات دولتی، مؤسسات نگهدارنده آرشیوهای دیجیتال و پژوهشگرانی که در تلاش برای دسترسی به آن‌ها هستند، استفاده می‌شود، مورد بحث قرار گرفت. جین وینترز (دانشگاه لندن) صراحتاً بیان کرد که هوش مصنوعی قابل توضیح بخشی از ایجاد روابط مبتنی بر اعتماد است و آرشیوها معمولاً در آگاهی از مسائلی مانند اعتماد و شفافیت عملکرد خوبی دارند.^۴

این موضوع بار دیگر به کدهای اخلاقی حرفه‌ای و چارچوب‌هایی که متخصصان GLAM در آن فعالیت می‌کنند، مرتبط می‌شود. کد اخلاقی شورای بین‌المللی آرشیوها (ICA) به طور خاص به نیاز به ثبت و توجیه اقدامات، ارائه خدمات بی‌طرفانه به کاربران و «اعتماد ویژه‌ای که به آن‌ها داده شده» اشاره می‌کند. (International Council on Archives, 1996) لورا میلر نیز توضیح می‌دهد که آرشیویست‌ها در چارچوب این کدهای اخلاقی فعالیت می‌کنند و بی‌طرفی و شفافیت از اهمیت ویژه‌ای در کار آرشیوی اخلاق مدار برخوردارند. (Millar, 2017, p. 95)

در این میان، تنش واقعی میان اصول اخلاقی حرفه‌ای آرشیویست‌ها و کاربرد هوش مصنوعی‌هایی که امکان توضیح عملکردشان وجود ندارد، پدیدار می‌شود. این نوع از هوش مصنوعی که قابل توضیح نیست، اغلب به دلیل شفافیت‌ناپذیری آن به عنوان «جعبه سیاه» شناخته می‌شود. ماهیت غیرقابل توضیح هوش مصنوعی جعبه سیاه، برای مورخان مصاحبه‌شونده نیز مشکل‌ساز بود. آدام نیکس (دانشگاه بیرمنگام) توضیح داد که در سیستم‌های جعبه سیاه: «ما با چیزی مواجهیم که نمی‌توانیم به عنوان پژوهشگر به طور کامل نقدش کنیم و نمی‌توانیم به طور کامل نتایجی را که به دست آورده‌ایم یا یافته‌هایی را که تولید کرده‌ایم، درک کنیم؛ این تبدیل به بخشی از فرآیند می‌شود که نمی‌توانیم آن را بفهمیم».^۵ دیدگاه نیکس با الزامات پاسخگویی مطرح شده در کد رفتار پژوهشگران UKRIO نیز هم‌راستا است.^۶ اگرچه ماهیت برخی ابزارهای هوش مصنوعی با اصول اخلاقی کلیدی در تضاد است، اما دانشمندان رایانه و

۱. [مدرک (A-Level) Advanced Level] یک مدرک تحصیلی سطح پیشرفته در نظام آموزشی بریتانیا است که معمولاً در پایان دوره دبیرستان (Sixth Form) و پیش از ورود به دانشگاه اخذ می‌شود. این مدرک نقش کلیدی در پذیرش دانشگاهی دارد و در بسیاری از کشورهای مشترک‌المنافع نیز با ساختار مشابهی ارائه می‌شود.

۲. در ۲۰۲۰م به دلیل همه‌گیری کووید-۱۹، آزمون‌های A-Level در بریتانیا لغو شد و نمرات توسط نهاد ناظر آزمون‌ها (Ofqual) اختصاص یافت. آن‌ها از داده‌های گذشته و پیش‌بینی‌های معلمان استفاده کردند تا نمرات را به صورت محاسباتی تولید کنند که منجر به کاهش ۴۰٪ از نمرات نسبت به پیش‌بینی معلمان شد. این موضوع با اعتراض عمومی مواجه شد و دولت مجبور شد نتایج را لغو کرده و فقط از پیش‌بینی‌های معلمان استفاده کند. برای جزئیات بیشتر بنگرید به: <https://inews.co.uk/news/technology/a-level-results-2020-trust-algorithms-exams-scandal-635222> (دسترسی در ۲۵ آوریل ۲۰۲۲). استفاده از واژه «الگوریتم» در ذهن عموم و رسانه‌ها این تصور را ایجاد کرد که این نمونه‌ای از هوش مصنوعی «جعبه سیاه» است، در حالی که در واقع مبتنی بر تحلیل آماری انسانی و کاملاً مستند بود.

۳. مصاحبه با اندرو بلیک، ۱۶ ژوئیه ۲۰۲۱.

۴. Explainable AI

۵. مصاحبه با جین وینترز، ۲۴ ژوئن ۲۰۲۱.

۶. Black Box

۷. مصاحبه با آدام نیکس، ۲۲ ژوئن ۲۰۲۱.

۸. اصل ۲.۵ از کد رفتار UKRIO برای پژوهشگران درباره پاسخگویی. <https://ukrio.org/publications/code-of-practice-for-principles--2/research> (دسترسی در ۲۵ آوریل ۲۰۲۲).

برنامه‌نویسان در حال توسعه فناوری‌هایی هستند که شفاف‌تر و قابل توضیح‌تر باشند. گراهام مک‌دونالد (دانشگاه گلاسگو) توضیح داد که اگرچه بسیاری از ابزارها در حال حاضر جعبه سیاه هستند، اما کارهایی در زمینه «شفافیت، قابلیت توضیح و تضمین انصاف در پیش‌بینی‌ها» در حال انجام است.^۱ مک‌دونالد همچنین تأکید کرد که شفافیت یکی از مهم‌ترین عوامل در ایجاد اعتماد کاربران به سیستم‌های هوش مصنوعی خواهد بود؛ اعتمادی که از درک عملکرد سیستم ناشی می‌شود.^۲ نمونه‌ای از این رویکرد را می‌توان در پروژه بررسی حساسیت دیجیتال^۳ مشاهده کرد که به صورت مشترک میان دفتر امور خارجه، کشورهای مشترک‌المنافع و توسعه (FCDO)، شرکت مشاوره‌ای SVGC و دانشگاه گلاسگو در حال اجراست.^۴

اندرو دیکسون، مدیرعامل SVGC، توضیح داد که هنگام استفاده از ابزارهای هوش مصنوعی برای تصمیم‌گیری درباره اسناد عمومی که قرار است آرشیو شوند، عموم مردم حق دارند بدانند این انتخاب‌ها چگونه انجام می‌شوند. این الزام، برگرفته از کدهای حرفه‌ای خدمات دولتی در زمینه پاسخ‌گویی و شفافیت است. دیکسون توضیح داد که این اصل بر انتخاب ابزارهای مورد استفاده در پروژه تأثیر گذاشته است: «ما از شبکه‌های عصبی استفاده نمی‌کنیم چون قابل توضیح نیستند [...] ما از فناوری‌هایی که نتایج تصادفی تولید می‌کنند، اجتناب کرده‌ایم» و در عوض، از فناوری‌هایی استفاده کرده‌اند که «قابل پیش‌بینی و قابل تکرار هستند [...] تا بتوان نتیجه را تعیین کرد».^۵ در اینجا می‌بینیم که چگونه اصول اخلاقی حرفه‌ای که بر خدمات دولتی، بخش GLAM و پژوهشگران دانشگاهی حاکم است، بر فناوری‌های مورد استفاده تأثیر می‌گذارند.

مک‌دونالد که در این پروژه نیز فعالیت دارد، تأکید کرد که توضیح شهودی از عملکرد ابزار هوش مصنوعی برای ایجاد اعتماد بسیار حیاتی است، به‌ویژه زمانی که کاربران این ابزارها، خود دانشمندان رایانه‌ای نیستند که الگوریتم‌ها را نوشته، آزمایش و ارزیابی کرده‌اند. پروژه‌های موفق هوش مصنوعی مانند این نمونه، حاصل همکاری میان کارمندان دولتی که فرآیند بررسی حساسیت را مدیریت می‌کنند، شرکای آرشیوی، پژوهشگران دانشگاهی و دانشمندان رایانه‌ای هستند که تخصص فنی را ارائه می‌دهند. در این همکاری‌ها، ذی‌نفعان اصول حرفه‌ای و کدهای اخلاقی را به خوبی منتقل کردند و از این طریق، ابزارهای هوش مصنوعی‌ای به کار گرفته شدند که با مفاهیم پاسخ‌گویی و شفافیت هم‌خوانی دارند.

کنترل

دومین تنش مطرح‌شده در مصاحبه‌ها، مربوط به فقدان کنترل و اعتماد به ابزارهای هوش مصنوعی بود. یازده نفر از مصاحبه‌شوندگان به موضوع کنترل در زمینه کاربرد هوش مصنوعی در آرشیوهای دیجیتال اشاره داشتند. اگرچه نمی‌خواهیم همکاران در خدمات دولتی، بخش GLAM و دانشگاه را «وسواسی در کنترل» معرفی کنیم، اما باید اذعان

۱. مصاحبه با گراهام مک‌دونالد، ۲۳ ژوئن ۲۰۲۱.

۲. همان.

۳. Digital Sensitivity Review

۴. بررسی حساسیت دیجیتال، بررسی فایل‌های دیجیتال برای تشخیص وجود اطلاعات حساس است. این پروژه در FCDO از سال ۲۰۱۸ آغاز شده و همچنان ادامه دارد. اطلاعات بیشتر در: <https://www.svgc.co.uk/fcd-services-consultancy> و <https://www.fcdoservices.gov.uk/what-we-offer/digital-records-sensitivity-review> (دسترسی در ۲۵ آوریل ۲۰۲۲).

۵. مصاحبه با اندرو دیکسون، ۱ ژوئیه ۲۰۲۱. همان‌طور که پیش‌تر در مقاله توضیح داده شد، ما از تعریف گسترده‌ای از هوش مصنوعی استفاده می‌کنیم - به عنوان مجموعه‌ای از رویکردهای محاسباتی، نه صرفاً تمرکز بر یادگیری عمیق و شبکه‌های عصبی.

۶. مصاحبه با گراهام مک‌دونالد، ۲۳ ژوئن ۲۰۲۱.

کرد که در این حرفه‌ها، کنترل بر آنچه تولید، منتشر، آرشیو، توصیف، گنج‌انده یا حذف می‌شود، اهمیت زیادی دارد. برای کارمندان دولتی و متخصصان GLAM، مسئولیت حرفه‌ای در زمینه حفظ حریم خصوصی و مدیریت اطلاعات حساس وجود دارد. جیمز لاپین (دفتر کابینه) توضیح داد که پس از تصمیم‌گیری برای انتشار یک سند، دیگر راه بازگشتی نیست: «اگر از هوش مصنوعی برای تعیین مجوزهای دسترسی استفاده کنید، برای باز کردن دسترسی [...] این تصمیم غیرقابل بازگشت است، چون وقتی دسترسی را به من دادید، دیگر فایده‌ای ندارد که آن را پس بگیرید، مثل این است که اسب از اصطبل بیرون رفته باشد»^۱. به طور مشابه، لئونیتین تالبوم (آرشیو ملی بریتانیا/ دانشگاه کالج لندن) توضیح داد که آرشیویست‌ها معمولاً ترجیح می‌دهند مواد فقط در اتاق‌های مطالعه در دسترس باشند، جایی که می‌توان کنترل بیشتری بر نوع دسترسی داشت.^۲ همچنین، اندرو رایللی، آرشیویست ارشد در مرکز آرشیو چرچیل، بیان کرد که آرشیویست‌ها برای کنترل ارزش قائل هستند و از دست دادن این کنترل در مواجهه با داده‌های حساس از طریق هوش مصنوعی یکی از نگرانی‌های اصلی اوست.^۳

این مسئولیت‌ها در زمینه حفاظت از حریم خصوصی و اطلاعات حساس با کدهای رفتاری و اصول اخلاق حرفه‌ای حاکم بر این حوزه‌ها ارتباط مستقیم دارد. اصل صداقت در خدمات دولتی به‌ویژه در این زمینه اهمیت دارد، زیرا به تعهد انجام وظایف به صورت مسئولانه، حرفه‌ای و با در نظر گرفتن استانداردهای اخلاقی سایر حرفه‌ها اشاره می‌کند.^۴ کد اخلاقی شورای بین‌المللی آرشیوها (ICA) به طور صریح‌تر به مسئولیت آرشیویست‌ها در قبال حریم خصوصی اشاره دارد و تصریح می‌کند که آرشیویست‌ها باید «مراقب باشند که حریم خصوصی سازمانی و شخصی، همچنین امنیت ملی، بدون از بین بردن اطلاعات محافظت شود، به‌ویژه در مورد اسناد الکترونیکی که به‌روزرسانی و حذف در آن‌ها رایج است» (International Council on Archives, 1996)

ترس از دست دادن کنترل به‌گونه‌ای متفاوت توسط پژوهشگران علوم انسانی در مصاحبه‌ها بیان شد. در اغلب موارد، این فقدان کنترل به ندانستن آنچه از دست رفته مربوط می‌شد. مورخان بر اهمیت یافته‌های تصادفی و غیرمنتظره هنگام دسترسی به جعبه‌ای از اسناد فیزیکی آرشیوی تأکید کردند. لینی آکی، پژوهشگر در دانشگاه وست مینستر، توضیح داد که در مواجهه با جست‌وجوهای مبتنی بر هوش مصنوعی، احساس «ناراحتی فوری» داشت، زیرا: «نمی‌خواهم چیزی در این فرآیند باشد که باعث شود ۱۰٪ از اطلاعات حذف شود، فقط به این دلیل که بی‌ربط تلقی شده یا با معیارهای جست‌وجو مطابقت نداشته، درحالی‌که از نظر من مرتبط بوده است».^۵ این دیدگاه توسط امیلی رابینسون، مدرس ارشد علوم سیاسی در دانشگاه ساسکس، نیز تأیید شد. او اشاره کرد که استفاده از ابزارهای جست‌وجوی مبتنی بر هوش مصنوعی ممکن است مانعی برای یافتن چیزهایی باشد که نمی‌دانستید به دنبالشان هستید.^۶ هلن مک‌کارتی، استاد تاریخ مدرن و معاصر بریتانیا در دانشگاه کمبریج، احساس خود نسبت به کاربرد هوش مصنوعی در مجموعه‌های آرشیوی دیجیتال را دوگانه توصیف کرد: از یک سو، پذیرای ظرفیت‌های بالقوه آن بود و از سوی دیگر، نگران از دست رفتن کنترل

۱. مصاحبه با جیمز لاپین، ۲۹ ژوئن ۲۰۲۱.

۲. مصاحبه با لئونیتین تالبوم، ۲۲ ژوئن ۲۰۲۱.

۳. مصاحبه با اندرو رایللی و کریستوفر نولز، ۷ ژوئیه ۲۰۲۱.

۴. «صداقت»، کد خدمات دولتی. در دسترس از:

<https://www.gov.uk/government/publications/civil-service-code/the-civil-service-code> (دسترسی در ۲۵ آوریل ۲۰۲۲).

۵. مصاحبه با لینی آکی، ۱ ژوئیه ۲۰۲۱.

۶. مصاحبه با امیلی رابینسون، ۱۵ ژوئیه ۲۰۲۱.

توسط کارمندان دولتی و آرشیویست‌ها و همچنین احتمال از دست رفتن اطلاعات مهم.^۱ یکی از راه‌هایی که دانشمندان رایانه برای پاسخ به این نگرانی‌ها در پیش گرفته‌اند، تغییر روایت درباره کاربرد هوش مصنوعی است. اصطلاح «انسان در حلقه»^۲ اغلب برای اشاره به شیوه‌ای به کار می‌رود که در آن ابزارهای هوش مصنوعی با تعامل انسانی آموزش داده می‌شوند. گراهام مک‌دونالد توضیح داد: «اگر این روایت را مطرح کنید که انسان در حلقه است، این انتظار ایجاد می‌شود که هوش مصنوعی قرار است تصمیم‌گیری کند و کار را انجام دهد، اما این درست نیست و در عمل، چنین انتظاری وجود ندارد».^۳ در عوض، مک‌دونالد پیشنهاد کرد که باید بیشتر بر فرآیندهایی تمرکز کنیم که در آن رایانه در حلقه است. یعنی هوش مصنوعی در خدمت کمک و پشتیبانی از انسان در فرآیند تصمیم‌گیری باشد.^۴ او افزود که «این تغییرات کوچک در روایت می‌تواند به تغییر دیدگاه مردم نسبت به این فناوری کمک کند».^۵ تغییر چنین دیدگاه‌هایی می‌تواند تا حدی از نگرانی‌هایی بکاهد که کارمندان مسئول دولتی، آرشیویست‌ها، یا پژوهشگران درباره از دست دادن کنترل بر حوزه کاری خود دارند.

سوگیری

سومین تنش مطرح شده در مصاحبه‌ها، مربوط به سوگیری الگوریتمی بود. دوازده نفر از مصاحبه‌شوندگان تأکید کردند که سوگیری در الگوریتم‌ها، نگرانی آنان یا اطرافیانشان بوده است. مفهوم سوگیری به‌ویژه در ارتباط با کدهای رفتاری و اصول اخلاق حرفه‌ای در خدمات دولتی، بخش GLAM و حوزه دانشگاهی اهمیت دارد. بی‌طرفی یکی از ویژگی‌های کلیدی در کد اخلاقی خدمات دولتی است و به‌طور خاص بیان می‌کند که کارمند دولت نباید «به‌گونه‌ای عمل کند که به‌طور غیرموجه به افراد یا منافع خاصی امتیاز دهد یا تبعیض قائل شود».^۶ این دیدگاه در کد اخلاقی شورای بین‌المللی آرشیوها (ICA) نیز بازتاب یافته است که تصریح می‌کند آرشیویست‌ها باید «خدماتی بی‌طرفانه ارائه دهند» (Interna-tional Council on Archives, 1996: 2) همچنین، کد رفتار UKRIO بر اهمیت شناخت کرامت، حقوق و ایمنی افراد در پژوهش تأکید دارد.^۷

اونا مورفی، مدرس مدیریت هنر در دانشگاه گلداسمیتز، توضیح داد که در کار با هوش مصنوعی در بخش GLAM، امکان چیزی وجود دارد که او آن را «سوگیری به توان دو» نامید. مورفی توضیح داد که بسیاری از گالری‌ها، کتابخانه‌ها، موزه‌ها و آرشیوها از پیش دارای سوگیری‌هایی در مجموعه‌هایشان هستند؛ سوگیری‌هایی که در طول صدها سال استعمار شکل گرفته‌اند و ساختارهایی را پدید آورده‌اند که بر نحوه گردآوری و توصیف مجموعه‌ها تأثیر گذاشته‌اند. استفاده از هوش مصنوعی در چنین مؤسسه‌های خطر افزودن لایه‌های جدیدی از سوگیری را به همراه دارد که مسئله را پیچیده‌تر

۱. مصاحبه با هلن مک‌کارتی، ۱ ژوئیه ۲۰۲۱.

۲. human-in-the-loop

۳. مصاحبه با گراهام مک‌دونالد، ۲۳ ژوئن ۲۰۲۱.

۴. همان.

۵. همان.

۶. «بی‌طرفی» در کد خدمات دولتی، در دسترس از:

<https://www.gov.uk/government/publications/civil-service-code/the-civil-service-code> (دسترسی در ۲۵ آوریل ۲۰۲۲).

۷. «ایمنی» در کد رفتار UKRIO، در دسترس از:

<https://ukrio.org/publications/code-of-practice-for-research-principles-2-0/> (دسترسی در ۲۵ آوریل ۲۰۲۲).

می‌کند.^۱ این خطر نه تنها توسط فعالان و پژوهشگران بخش GLAM، بلکه توسط سیاست‌آ. کریر (مشاور سیاست‌گذاری و حکمرانی هوش مصنوعی) نیز تأیید شد. او اذعان داشت که آموزش ابزارهای هوش مصنوعی با استفاده از مجموعه‌های فرهنگی ممکن است با مسائلی مانند جنسیت‌گرایی مواجه شود.^۲ به طور مشابه، مورخانی مانند لینزی آکی بیان کردند که نسبت به آنچه ممکن است از نتایج جست‌وجو حذف شود، «کمی بدبین» هستند، زیرا «آرشیوها از پیش گزینش شده‌اند و صداها، روایت‌ها و تجربه‌های افراد از آرشیو حذف می‌شوند. بنابراین، این خطر وجود دارد که برخی ابزارهای هوش مصنوعی، افراد حاشیه‌نشین را بیش از پیش به حاشیه برانند».^۳

توان بالقوه هوش مصنوعی برای تقویت سوگیری، نگرانی واقعی بسیاری از ذی‌نفعان است. اگرچه سوگیری نباید به عنوان یک رویه نرمال پذیرفته شود، چندین مصاحبه‌شونده اذعان داشتند که سوگیری هرگز به طور کامل قابل حذف نیست. جین وینترز اشاره کرد که پژوهشگران علوم انسانی معمولاً درک می‌کنند که سوگیری همیشه وجود دارد و اینکه پذیرفتن سوگیری و فهم تأثیر آن بر پژوهش اهمیت دارد: «نمی‌توان سوگیری‌ها را حذف کرد، فقط باید پذیرفت که وجود دارند و فهمید این موضوع چه تأثیری بر نتایجی دارد که به دست می‌آورید [...]». اگر تلاش کنید سوگیری را حذف کنید، فقط سوگیری جدیدی ایجاد می‌کنید. پس بهتر است آن را بپذیرید و با آن کار کنید، نه اینکه بخواهید حذفش کنید».^۴ گرت میلورد، پژوهشگر در دانشگاه بیرمنگام، نیز تأکید کرد که پژوهشگران باید «نسبت به سوگیری‌هایی که ممکن است در برخی فناوری‌هایی که استفاده می‌کنیم نهادینه شده باشند، کاملاً آگاه باشند [...] و دست‌کم بتوانند آن‌ها را در زمینه مناسب تحلیل کنند، حتی اگر نتوانند به طور کامل حذفشان کنند».^۵

همکاری‌ها و ایجاد اعتماد

بخش حاضر به بررسی این موضوع می‌پردازد که چگونه همکاری‌های عمیق‌تر میان کارمندان دولتی، متخصصان بخش GLAM و دانشگاهیان حوزه‌های علوم انسانی و علوم رایانه می‌تواند به ایجاد سطح مناسبی از اعتماد به ابزارهای هوش مصنوعی کمک کند و از کاربرد آن‌ها در دسترس‌پذیرتر و قابل استفاده‌تر کردن آرشیوهای دیجیتال حمایت کند. در حرفه‌های مختلف، از جمله خدمات دولتی، بخش GLAM و دانشگاه، کار گروهی و همکاری ارزشمند تلقی می‌شود. در کد اخلاقی خدمات دولتی، به همکاری و احترام با همکاران و شرکا و به در نظر گرفتن استانداردهای اخلاقی حاکم بر سایر حرفه‌ها اشاره شده است.^۶ کد اخلاقی شورای بین‌المللی آرشیوها (ICA) تصریح می‌کند که آرشیویست‌ها باید «با اعضای حرفه‌های مرتبط بر اساس احترام و درک متقابل همکاری کنند» (International Council on Archives, 1996: 3). همچنین، کد رفتار UKRIO توصیه می‌کند که پژوهشگرانی که به صورت مشترک فعالیت می‌کنند، باید با آگاهی از راهنمایی‌ها و استانداردهای اخلاقی همکارانشان عمل کنند.^۷ با این حال، در عمل، همکاری میان کارمندان دولتی،

۱. مصاحبه با اونا مورفی، ۳۰ ژوئن ۲۰۲۱.

۲. مصاحبه با سباستین آ. کریر، ۲ ژوئیه ۲۰۲۱.

۳. مصاحبه با لینزی آکی، ۱ ژوئیه ۲۰۲۱.

۴. مصاحبه با جین وینترز، ۲۴ ژوئن ۲۰۲۱.

۵. مصاحبه با گرت میلورد، ۷ ژوئیه ۲۰۲۱.

۶. «صداقت» در کد خدمات دولتی، در دسترس از:

<https://www.gov.uk/government/publications/civil-service-code/the-civil-service-code> (دسترسی در ۲۵ آوریل ۲۰۲۲).

۷. کد رفتار UKRIO، بند ۳.۵ درباره همکاری. در دسترس از: <https://ukrio.org/publications/code-of-practice-for-research-collaborative-working-5-3/standards-for-organisations-and-researchers-> (دسترسی در ۲۵ آوریل ۲۰۲۲).

متخصصان GLAM و دانشگاهیان برای گشودن آرشیوهای دیجیتال به ندرت اتفاق می افتد. برای پیشبرد این هدف، نیاز به گفت وگوهای گسترده و میان رشته ای وجود دارد؛ گفت وگوهایی درباره دغدغه های مشترک، اصول حرفه ای و اخلاقی مشترک و همچنین یافتن راه حل هایی با مشارکت دانشمندان رایانه ای که ابزارها را توسعه می دهند، نه صرفاً تمرکز بر موانع اجرای هوش مصنوعی.

سه تنش اصلی که در بخش پیشین بررسی شد (پاسخگویی، کنترل و سوگیری) به بی اعتمادی عمومی نسبت به هوش مصنوعی در میان متخصصان خدمات دولتی، بخش GLAM و دانشگاه منجر شده اند، اما در هر یک از این تنش ها، مصاحبه شوندگانی بودند که راه حل هایی مطابق با کدهای رفتاری و اصول اخلاقی حرفه ای پیشنهاد دادند. نکته کلیدی این است که همکاری های میان رشته ای نقش حیاتی در توسعه ابزارهای هوش مصنوعی قابل اعتماد دارند. جان شریدن (آرشیو ملی بریتانیا) خاطرنشان کرد: «همه چیز به انتقال دانش از بیرون به بخش ما مربوط می شود. هیچ کس چیزی را به طور خاص برای ما نمی سازد. همه چیز مربوط به چیزهایی است که در جهان ساخته می شود و بعد کسی می گوید: "اوه، می توانم این را در آرشیوها به کار ببرم." و فکر می کنم تنها کاری که واقعاً می توانیم انجام دهیم این است که خودمان را مجهز کنیم تا بفهمیم این فرآیند انتقال فناوری چگونه است و درک کنیم که پیامدهای آن چیست.»

چالش های مهارت، اعتماد و مشارکت در استفاده از هوش مصنوعی

جان شریدن ادامه داد که «برای ما واقعاً سخت است که بدانیم چگونه باید در این جهان حرکت کنیم؛ جایی که اگرچه مجموعه هایی داریم، اما از نظر بودجه و توانمندی، ما مثل کک هایی در اطراف فیل ها هستیم. اما تنها نیستیم. این مسائل فراگیر هستند و فکر می کنم مهم است که نه فقط با یکدیگر، بلکه با دیگرانی که دغدغه های مشابه دارند، گفت وگو کنیم»^۱. دانیل ون استرین، متصدی دیجیتال در کتابخانه بریتانیا نیز تأکید کرد که بخش GLAM باید فعالانه در نحوه به کارگیری ابزارهای هوش مصنوعی در مجموعه ها مشارکت داشته باشد. اگر این ابزارها بدون تعامل با متخصصان GLAM توسعه یابند، ممکن است در عمل برای این بخش کاربردی نباشند.^۲ درحالی که برخی مصاحبه شوندگان نسبت به ایده استفاده از هوش مصنوعی بی اعتماد بودند، برخی دیگر به تصمیمات آرشیوها در زمینه کاربرد بالقوه هوش مصنوعی اعتماد داشتند. گرت میلورد (دانشگاه بیرمنگام) توضیح داد: «اگر آرشیو ملی ابزاری برای جست و جو در آرشیو خودش در اختیار من قرار دهد، به این اعتماد دارم که آن ها درباره اش فکر کرده اند و این بخشی از فرآیند آرشیوی آن هاست [...]، همان طور که باید به روش فهرست نویسی شان اعتماد کنم»^۳.

با این حال، در عمل، بسیاری از کاربران دانش و توانایی استفاده از فناوری های پیشرفته مانند هوش مصنوعی برای کاوش در آرشیوها را ندارند. این شکاف مهارتی، مانعی برای گفت وگوهای مشارکتی ایجاد می کند. یکی از متخصصان در یک مؤسسه آرشیوی بزرگ در لندن گفت: «چند بار تلاش کردیم رویدادهایی برای پژوهشگران برگزار کنیم تا بفهمیم چه انتظاراتی دارند یا چه می خواهند. اما [...] بیشتر مخاطبان اصلی پژوهشی ما هنوز با اسناد دیجیتال متولد شده کار نمی کنند». این موضوع برای آرشیویست ها و متخصصان حفظ دیجیتال که سال هاست روی این نوع اسناد کار کرده اند، ناامیدکننده است. «رویدادها تبدیل به جلسات آموزشی درباره آنچه حرفه آرشیو از ۱۰ سال تلاش برای حفظ دیجیتال آموخته می شوند، نه جلساتی برای استخراج انتظارات پژوهشگران». برای این متخصص، دیدن تقاضای اندک پژوهشگران برای استفاده از ابزارهای دیجیتال در کاوش آرشیو دیجیتال، شگفت انگیز بود: «همچون سپرده گذاران، پژوهشگران هم

۱. مصاحبه با جان شریدن، ۲۹ ژوئن ۲۰۲۱.

۲. مصاحبه با دانیل ون استرین، ۲ ژوئیه ۲۰۲۱.

۳. مصاحبه با گرت میلورد، ۷ ژوئیه ۲۰۲۱.

دنیای کاغذی و روش های پژوهش کاغذی شان را به دنیای دیجیتال منتقل می کنند»^۱.

اندرو رایلی (مرکز آرشیو چرچیل) نیز به کمبود مهارت هایی اشاره کرد که مانع از کار مؤثر کاربران در مجموعه های دیجیتال می شود. او تجربه خود از حضور در کارگاهی در وزارت امور خارجه را که توسط مورخ هلن مک کارتی درباره اسناد دیجیتال متولد شده برگزار شده بود، شرح داد. رایلی دریافت که مانند بسیاری از آرشیویست های «سنتی» آموزش دیده برای پردازش اسناد کاغذی، مورخان حاضر در کارگاه نیز در استفاده از مواد دیجیتال احساس ناامنی و تردید داشتند: «فکر می کردم اعتماد به نفس بیشتری داشته باشند، اما همه شان نسبت به مهارت های اساسی خود به عنوان مورخ در استفاده از این نوع مواد تردید داشتند و به دلایل مشابه، امیدوار بودند قبل از اینکه مجبور شوند زیاد با این ها کار کنند، بازنشسته شوند که واقعاً برایم تعجب آور بود»^۲.

در چنین زمینه ای، جای تعجب نیست که بسیاری از پژوهشگران به مجموعه هایی روی می آورند که دسترسی و استفاده از آن ها آسان تر است. مجموعه های کاغذی یا مجموعه های دیجیتالی شده که با جست و جوی کلیدواژه ای قابل کاوش هستند. نگرانی ها درباره داده های کلان و هوش مصنوعی می تواند از طریق درک مفهومی تکنیک های مورد استفاده کاهش یابد. با این حال، پژوهشگران علوم انسانی فرصت های نهادی اندکی برای پر کردن شکاف دانشی خود در زمینه هوش مصنوعی دارند. برای مثال، مدرسه تابستانی «علوم انسانی دیجیتال در آکسفورد» در سال ۲۰۲۲م هیچ دوره ای درباره هوش مصنوعی نداشت.

بدون فشار قابل توجه از سوی پژوهشگران برای گشودن مجموعه های دیجیتال متولد شده و بدون راه حل های آسان برای مسائل حساسیت و حق نشر، منطقی به نظر می رسد که آرشیوها فعلاً بسته باقی بمانند. جیمز لاپین از تصمیم NARA برای حفظ آرشیو دیجیتال کاخ سفید - حتی اگر هنوز «تاریک» باشد - تمجید کرد. حفظ این آرشیو به عنوان مهم ترین اقدام معرفی شد، زیرا امکان دسترسی در آینده را فراهم می سازد.^۳ در مقابل، برندان نولتون (عضو تیم دولتی) خواستار شفافیت بیشتر شد و گفت: «وقتی درباره بخش عمومی صحبت می کنید، کاری که ما انجام می دهیم با بودجه عمومی تأمین می شود یا توسط پارلمان، روزنامه نگاران و دیگران مورد بررسی قرار می گیرد. بیشتر کارهای ما باید به طور پیش فرض عمومی باشد [...]». فکر می کنم اگر کارهای بیشتری از ما عمومی بود، مردم بیشتر آن را خسته کننده می دانستند. اما همچنین، در صورت افشای داده ها، دیگر خبری از تیتراهی مثل "اوه، گنجینه مخفی ایمیل های دفتر کابینه توسط دستیار منتشر شد" نخواهد بود»^۴.

همان طور که دیدیم، پاسخ گویی و شفافیت از اصول محوری در کدهای اخلاقی کارمندان دولتی هستند، اما این اصول باید در کنار لزوم حفظ امنیت ملی و جلوگیری از افشای اسناد حساس مورد توجه قرار گیرند. برگزاری گفت وگوهای مشترک با سایر ذی نفعان، از جمله آرشیویست ها و دانشگاهیان می تواند دیدگاه های تازه ای در اختیار متخصصان دولتی قرار دهد و به آنان کمک کند تا در مورد انتقال اسناد به مخازن آرشیوی، تصمیمات آگاهانه تری اتخاذ کنند.

نتیجه گیری

این پژوهش نشان داد که فقدان ارتباط مؤثر میان کارمندان دولتی، متخصصان بخش GLAM و دانشگاهیان اغلب منجر به بی اعتمادی به سایر ذی نفعان می شود. مصاحبه شوندگان پیش بینی کردند که در چنین شرایطی ممکن است

۱. مصاحبه با [نام محفوظ]، ۲۷ مه ۲۰۲۱.

۲. مصاحبه با اندرو رایلی و کریستوفر نولز، ۷ ژوئیه ۲۰۲۱.

۳. مصاحبه با جیمز لاپین، ۲۹ ژوئن ۲۰۲۱.

۴. مصاحبه با برندان نولتون، ۱ ژوئیه ۲۰۲۱.

داده‌ها افشا شوند، اعتبار افراد آسیب ببیند و در بدترین حالت، امنیت ملی به خطر بیفتد، اما خیانت بدون اعتماد ممکن نیست و دقیقاً همین فقدان اعتماد به سایر ذی‌نفعان و فناوری‌های نوین بود که در مصاحبه‌ها برجسته شد. بی‌اعتمادی به فناوری، اجرای ابزارهای هوش مصنوعی را دشوار می‌سازد. نگرانی‌هایی که مصاحبه‌شوندگان مطرح کردند - در زمینه‌های پاسخ‌گویی، کنترل و سوگیری - فرآیند توسعه هوش مصنوعی در حوزه آرشیوها را کند می‌کند. این مقاله توصیه می‌کند که گفت‌وگوی میان‌بخشی و میان‌رشته‌ای افزایش یابد تا اصول اخلاقی حرفه‌ای مشترک میان ذی‌نفعان آشکار شود. درک عمیق‌تر میان انسان‌ها گامی ضروری برای به‌کارگیری هوش مصنوعی در اسناد دیجیتال متولد شده است. همچنین، نیاز به پژوهش‌های بیشتر در مرزهای میان‌رشته‌ای وجود دارد. مسائل مربوط به فناوری‌های نوین در آرشیوهای دیجیتال اغلب تنها در چارچوب رشته‌هایی مانند مطالعات آرشیوی یا علوم رایانه بررسی شده‌اند. درحالی‌که کاربرد هوش مصنوعی در آرشیوها نیازمند مجموعه‌ای متنوع از مهارت‌هاست، نه تنها مهارت‌های محاسباتی، بلکه دانش حوزه‌ای، یعنی شناخت دقیق از موجودیت‌هایی که الگوریتم‌ها با آن‌ها سروکار دارند. تفکر در چارچوب‌های محدود رشته‌ای، توان جمعی ما را برای به‌کارگیری فناوری در اسناد دیجیتال تضعیف می‌کند. علوم انسانی دیجیتال که ذاتاً رشته‌ای میان‌رشته‌ای است، نقش مهمی در پر کردن شکاف میان تولیدکنندگان اسناد، آرشیویست‌ها، پژوهشگران و سایر کاربران ایفا می‌کند.

منابع مالی

این پژوهش با حمایت مالی گروه پروژه‌های سازمانی (EPG) در دانشگاه لافبرو انجام شده است. منابع مالی EPG از «صندوق نوآوری آموزش عالی» تأمین می‌شود.

منابع

- Agar, J. (2020) What is science for? The Lighthill report on artificial intelligence reinterpreted. *The British Journal for the History of Science*, 53(3): 289–310. <https://doi.org/10.1017/S0007087420000230>
- Angelova, L., Collomosse, J., Smillie, R., Cubey, R., and Norman, J. (2020) Deep discoveries: Interim Report. Foundation Projects, Towards a National Collection, Arts and Humanities Research Council. <https://www.nationalcollection.org.uk/sites/default/files/2021-02/Deep%20Discoveries.pdf> (accessed 4 November 2022)
- Baron, J. R., Sayed, M. F., and Oard, D. W. (2020) Providing more efficient access to government records: a use case involving application of machine learning to improve FOIA Review for the deliberative process privilege. *Journal on Computing and Cultural Heritage*, 15(1):1–19. <https://doi.org/10.1145/3481045>
- Barons, M., Bhatia, S., Double, J., et al. (2021) Safeguarding the nation's digital memory: towards a Bayesian model of digital preservation risk. *Archives and Records*, 42(1): 58–78. <https://doi.org/10.1080/23257962.2021.1873121>
- Cave, S., Coughlan, K., and Dihal, K. (2019) 'Scary robots': Examining public responses to AI. *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, Association for Computing Machinery Honolulu, HI, USA. <https://doi.org/10.1145/3306618.3314232>
- Christie, L. (2020) Interpretable machine learning (POSTnote No. 633). UK Parliament, London. <https://post.parliament.uk/research-briefings/post-pn-0633/> (accessed 4 November 2022)
- Cordell, R. (2020) Machine Learning in Libraries. Library of Congress, Washington, D.C. <https://labs.loc.gov/static/labs/work/reports/Cordell-LOC-ML-report.pdf?locl=blogs> (accessed 4 November 2022)

- Corrado, E. M. and Sandy, H. M. (2017) Digital Preservation for Libraries, Archives, and Museums. London: Rowman & Littlefield Publishers
- Deegan, M. and Tanner, S. (eds). (2006) Digital Preservation. London: Facet
- Delve, J. and Anderson, D. (eds) (2014) Preserving Complex Digital Objects. London: Facet
- Fašt, E. and Horvitz, E. (2017) Long-term trends in the public perception of artificial intelligence. AAAI (Association for the Advancement of Artificial Intelligence) Conference, pp. 963–69
- Gabriel, I. (2020) Artificial intelligence, values, and alignment. *Minds and Machines*, 30: 411–37. <https://doi.org/10.1007/s11023-020-09539-2>
- Hagendorff, T. (2020) The ethics of AI ethics: an evaluation of guidelines. *Minds and Machines*, 30: 99–120. <https://doi.org/10.1007/s11023-020-09517-8>
- Harris, F. and Lyon, F. (2013) Transdisciplinary environmental research: building trust across professional cultures. *Environmental Science & Policy*, 31: 109–19. <https://doi.org/10.1016/j.envsci.2013.02.006>
- Harvey, R. and Weatherburn, J. (2018) Preserving Digital Materials. Washington DC: Rowman & Littlefield
- Hockx-Yu, H. (2014) Access and scholarly use of web archives. *Alexandria: The Journal of National and International Library and Information Issues*, 25(1–2): 113–27. <https://doi.org/10.7227/ALX.0023>
- International Council On Archives. (1996) Code of Ethics. International Council on Archives
- Jackson, P. C. (2019) Introduction to Artificial Intelligence. Mineola, NY: Dover Publications
- Jaillant, L. (2022a) How can we make born-digital and digitised archives more accessible? Identifying obstacles and solutions. *Archival Science*. <https://doi.org/10.1007/s10502-022-09390-7>
- Jaillant, L. (ed.) (2022b) Archives, Access and Artificial Intelligence: Working with Born-Digital and Digitized Archival Collections. Bielefeld, Germany: Transcript. <https://doi.org/10.14361/9783839455845>
- Jaillant, L. (2019) After the digital revolution: working with emails and born-digital records in literary and publishers' archives. *Archives and Manuscripts*, 47(3): 285–304. <https://doi.org/10.1080/01576895.2019.1640555>
- Jaillant, L. and Caputo, A. (2022) Unlocking digital archives: cross-disciplinary perspectives on AI and born-digital data. *AI & Society*. <https://doi.org/10.1007/s00146-021-01367-x>
- Jobin, A., Ienca, M., and Vayena, E. (2019) Artificial Intelligence: the global landscape of ethics guidelines. *Nature Machine Intelligence*, 1: 389–99. <https://doi.org/10.1038/s42256-019-0088-2>
- Kelley, P. G., Yang, Y., Heldreth, C., et al. (2021) Exciting, useful, worrying, futuristic: public perception of artificial intelligence in 8 countries. *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, Association for Computing Machinery. <https://doi.org/10.1145/3461702.3462605>
- Kosciuk, M. and Myska, M. (2019) Copyright law challenges of preservation of born-digital digital content as cultural heritage. *European Journal of Law and Technology*, 10(1): 1–21
- Lawton, D., Stacey, R., and Dodd, G. (2014) eDiscovery in digital forensic investigations (No. 32/14). CAST Publication, Home Office UK, London. <https://www.gov.uk/government/publications/ediscovery-in-digital-forensic-investigations> (accessed 4 November 2022)
- Lomas, E. J. (2017) Navigating confidentiality, legal privilege and GDPR to maintain legal records for future generations: the case for archiving. Presented at Legal Records at Risk, London. Available at: <https://discovery.ucl.ac.uk/id/eprint/10026237>
- McDonald, G., Macdonald, C., and Ounis, I. (2020a) How the accuracy and confidence of sensitivity classification affects digital sensitivity review. *ACM Transactions on Information Systems (TOIS)*, 39(1): 1–34. <https://doi.org/10.1145/3417334>
- McDonald, G., Macdonald, C., and Ounis, I. (2020b) Active Learning Stopping Strategies for Technology-As-

sisted Sensitivity Review. In Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. New York, NY, USA: Association for Computing Machinery, 2053–56. Available at: <https://doi.org/10.1145/3397271.3401267> (Accessed: 9 August 2021)

- Millar, L. (2017) Archives: Principles and Practices. London: Facet
- Mittelstadt, B. (2019) Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, 1: 501–07. <https://doi.org/10.1038/s42256-019-0114-4>
- Moss, E., Watkins, E. A., Singh, R., Elish, M. C., and Metcalf, J. (2021) Assembling Accountability: Algorithmic Impact Assessment for the Public Interest. Available at SSRN: <https://ssrn.com/abstract=3877437> or <http://dx.doi.org/10.2139/ssrn.3877437> (accessed 4 November 2022)
- Murphy, O. and Villaespesa, E. (2020) AI: A Museum Planning Toolkit. The Museums þ AI Network, Goldsmiths, University of London. <https://themuseumsai.network/toolkit/> (accessed 4 November 2022)
- Nix, A. and Decker, S. (2021) Using digital sources: the future of business history? *Business History*, 1–24. <https://doi.org/10.1080/00076791.2021.1909572>
- Owens, T. (2018) The Theory and Craft of Digital Preservation. Baltimore: Johns Hopkins University Press
- Padilla, T. (2019) Responsible Operations: Data Science, Machine Learning, and AI in Libraries. OCLC Research, Dublin, Ohio. <https://www.oclc.org/content/dam/research/publications/2019/oclcresearch-responsible-operations-datascience-machine-learning-ai.pdf> (accessed 4 November 2022)
- The National Archives UK. (2016) The application of technology-assisted review to born-digital records transfer, inquiries and beyond. www.nationalarchives.gov.uk/documents/technology-assisted-review-to-born-digital-records-transfer.pdf (accessed 4 November 2022)
- The National Archives UK. (2018) A guide to archiving personal data. <https://cdn.nationalarchives.gov.uk/documents/information-management/guide-to-archiving-personal-data.pdf> (accessed 4 November 2022)
- Vlassenroot, E., Chambers, S., Di Pretoro, E., et al. (2019) Web archives as a data resource for digital scholars. *International Journal of Digital Humanities*, 1: 85–111. <https://doi.org/10.1007/s42803-019-00007-7>
- Wadsworth, C., Vera, F., and Piech, C. (2018) Achieving fairness through adversarial learning: an application to recidivism prediction. <https://doi.org/10.48550/arXiv.1807.00199>
- Wilkinson, M.D. et al. (2016) The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1): 160018. <https://doi.org/10.1038/sdata.2016.18>
- Winters, J. and Prescott, A. (2019) Negotiating the born-digital: a problem of search. *Archives and Manuscripts*, 47(3): 391–403. <https://doi.org/10.1080/01576895.2019.1640753>