



The Automation of Research in Digital Humanities: Artificial Researchers, Digital Archives, and the Democratization of Scholarly Access

ROBERTO DI QUIRICO

UNIVERSITÀ DEGLI STUDI DI CAGLIARI, DIPARTIMENTO DI INGEGNERIA MECCANICA, CHIMICA E DEI MATERIALI (DIMCM) – DIQUIRICO@UNICA.IT

ABSTRACT (ENGLISH)

Research automation via artificial researchers, defined as AI-driven agents capable of autonomous or semi-autonomous research activities, has shifted from experimental to operational status in high-return scientific domains during 2024-2025. The present study examines the implications for Digital Humanities and social sciences, with a particular focus on digital archives as pivotal sites where automated systems are poised to fundamentally reconfigure access and institutional relationships.

Based on exploratory discussions with senior staff at major European archival institutions and the author's participation in international research automation projects, this analysis identifies two primary integration models: external access via APIs that enable researcher-controlled systems, and internal artificial archivist agents that mediate user interactions. Both models offer significant public engagement benefits by reducing economic, skill-based, and temporal barriers to archival research, thereby expanding potential users beyond specialist academics to include citizen researchers, educators, and community groups.

Nevertheless, significant challenges constrain deployment, including technical limitations in humanities applications, infrastructural heterogeneity that impedes interoperability, and institutional tensions between access democratization and the preservation of symbolic capital invested in physical facilities. This paper contends that, despite these obstacles, archival institutions must engage in anticipatory planning by coordinating digital transformation, developing governance frameworks, and advocating for interoperability standards. Such proactive measures are necessary to shape, rather than merely react to, the transformations redefining scholarly research infrastructures.

Keywords: artificial researchers; research automation; digital archives; public engagement; Digital Humanities

ABSTRACT (ITALIANO)

L'automazione della ricerca nelle Digital Humanities: Ricercatori Artificiali, Archivi Digitali e la democratizzazione dell'accesso alla ricerca

L'automazione della ricerca attraverso ricercatori artificiali—agenti guidati dall'intelligenza artificiale capaci di svolgere attività di ricerca autonome o semi-autonome—è passata da uno stato sperimentale a operativo in ambiti scientifici ad alto ritorno economico nel 2024-2025. Questo contributo esamina le implicazioni per le Digital Humanities e le scienze sociali, concentrandosi sugli archivi digitali come luoghi critici in cui i sistemi automatizzati riconfigureranno radicalmente l'accesso e le relazioni istituzionali.

Sulla base di colloqui esplorativi con dirigenti di importanti istituzioni archivistiche europee e del coinvolgimento dell'autore in progetti internazionali di automazione della ricerca, l'analisi identifica due modelli di integrazione: l'accesso esterno tramite API, che abilita sistemi controllati dai ricercatori, e gli agenti archivisti artificiali interni, che mediano le interazioni degli utenti. Entrambi promettono benefici sostanziali per il public engagement, riducendo le barriere economiche, competenziali e temporali alla ricerca archivistica e espandendo i potenziali utenti dagli accademici specialisti ai citizen researchers, agli educatori e ai gruppi comunitari. Tuttavia, sfide significative ne vincolano l'implementazione: limiti tecnici nelle applicazioni umanistiche, eterogeneità infrastrutturale che impedisce l'interoperabilità e tensioni istituzionali tra la democratizzazione dell'accesso e la preservazione del capitale simbolico investito nelle strutture fisiche. Il contributo sostiene che, nonostante queste sfide, le istituzioni archivistiche devono impegnarsi ora in una pianificazione anticipatoria—coordinando la trasformazione digitale, sviluppando quadri di governance e sostenendo standard di interoperabilità—per plasmare anziché reagire semplicemente alle trasformazioni che ridefiniscono le infrastrutture di ricerca.

Parole chiave: ricercatori artificiali; automazione della ricerca; archivi digitali; coinvolgimento pubblico; Digital Humanities

1. INTRODUCTION

In recent years, research automation has transitioned from speculative construct to operational reality. AI-driven agents capable of conducting autonomous or semi-autonomous research activities (hereinafter, Artificial Researchers, or AR), or AI platforms such as AlphaFold, are now ready for deployment across multiple scientific domains. These systems perform extensive literature reviews, identify research gaps, formulate hypotheses, design methodological approaches, and in some cases autonomously generate research outputs (Tang et al., 2025; Beel et al., 2025). The emergence of fully functional automated research platforms, particularly during 2024-2025, marks a critical threshold: research automation is no longer experimental but increasingly operational.

Recent contributions illustrate that Artificial Researchers do not constitute a single architectural model but rather a heterogeneous family of systems characterized by different degrees of autonomy, coordination mechanisms, and roles assigned to human researchers. The *AI co-scientist* framework conceptualizes the Artificial Researcher as a **human-in-the-loop epistemic collaborator**, based on a multi-agent architecture that emphasizes hypothesis generation, structured scientific debate, tournament-based selection, and iterative refinement under explicit human guidance (Gottweis et al., 2025). By contrast, the *ROBIN* system demonstrates a more **integrated and semi-autonomous research pipeline**, in which AI agents not only generate hypotheses but also interpret experimental data and iteratively update research directions within a lab-in-the-loop setting, significantly reducing human cognitive intervention while still relying on physical experimentation conducted by researchers (Ghareeb et al., 2025). Finally, the *AI-Researcher* framework advances the most radical configuration, proposing a **fully autonomous end-to-end Artificial Researcher** capable of literature review, idea generation, algorithmic implementation, experimental validation *in silico*, and publication-ready manuscript production, coupled with a dedicated benchmark (Scientist-Bench) to evaluate autonomous scientific performance (Tang et al., 2025). Taken together, these cases show that research automation is not converging toward a single model but toward a spectrum of Artificial Researchers whose architectures vary along key dimensions such as autonomy, modularity, validation strategies, and human oversight.

To date, these developments have concentrated in sectors characterized by high economic returns and advanced automation infrastructures. Materials engineering, pharmacology, and computational sciences—particularly machine learning—have witnessed the deployment of sophisticated platforms capable of coordinating robotic laboratories, managing complex experimental workflows, and autonomously generating scientific publications (Tang et al., 2025). In machine learning specifically, where research can be conducted entirely computationally, recent platforms demonstrate capabilities approaching full research autonomy (Beel et al., 2025; Lu et al., 2024).

This study situates these broader transformations within the context of Digital Humanities and social sciences, contending that automated research platforms and artificial researchers will soon become integral components of scholarly research infrastructures in these fields. As primary sources and archival materials increasingly originate in digital formats, the distinction between traditional humanities research and computationally mediated research is progressively diminishing (Chapinal-Heras & Díaz-Sánchez, 2023). Digital archives, in particular, constitute a critical domain where the integration of artificial researchers can fundamentally reconfigure access modalities, expand user demographics, and redefine institutional relationships among archives, researchers, and the broader public.

This contribution focuses specifically on archival contexts to examine how artificial researchers may integrate with existing infrastructures, the potential benefits—particularly in terms of public engagement—such integration offers, and the institutional, technical, and methodological challenges that persist. Importantly, this analysis underscores the necessity of anticipatory planning: archival institutions and research communities must begin preparing for these transformations now, even as the technologies remain partially nascent, to ensure that future digital reorganization aligns with both scholarly requirements and public access imperatives.

2. CONTEXT AND METHODOLOGY

Research automation operates along a continuum. At one extreme lie fully autonomous systems capable of independently defining research objectives, selecting methodologies, executing research pipelines, and

disseminating results (Tang et al., 2025). At the other extreme are directed automation models, in which human researchers define research questions or identify specific gaps in the existing literature, delegating execution to artificial research agents (Beel et al., 2025). Both models are operational, and both hold significant implications for humanities and social science research.

The empirical foundation for this analysis is based on the author's role as principal investigator of an international research consortium comprising four European universities. This consortium investigates research automation applications across multiple disciplinary domains, including those pertinent to Digital Humanities and archival sciences. The project provides an institutional context for examining how automated research systems may integrate with humanities research infrastructures.

This analysis also draws on exploratory discussions conducted in 2025 with senior archival staff at the Historical Archives of the European Union (HAEU) in Fiesole and the Historical Archives of the Bank of Italy in Rome. These conversations were informal in nature and not designed as a systematic empirical inquiry; they served rather as consultations aimed at gauging institutional perspectives on digital transformation and the prospect of automated research systems. Both institutions are currently transitioning toward predominantly digital archival collections, a shift that necessitates new computational approaches to access, management, and research support (Colavizza et al., 2022; Shinde et al., 2024).

The perspective developed in this paper should be situated transparently. The author does not write from within archival science or as a specialist in archival theory, but as a researcher whose competence lies in the study of research automation and artificial researchers across heterogeneous scientific domains. The contribution offered here is therefore deliberately external to the archival profession: it consists in importing into the archival debate a conceptual framework—centred on artificial researchers and their integration models—that has taken shape in adjacent scientific fields where research automation has progressed more rapidly. This positioning is consistent with the interdisciplinary character of Digital Humanities, which has historically advanced through the cross-pollination of methods and frameworks between domains. The aim is not to displace archival expertise, which remains indispensable for any meaningful integration of automated systems into archival practice, but to bring into the archival conversation a set of developments and analytical categories that, if left unaddressed, risk shaping archival infrastructures from the outside without institutional participation.

3. CURRENT STATE AND TECHNICAL CHALLENGES

Despite recent advances, significant technical, methodological, and institutional challenges constrain the deployment of artificial researchers in humanities and social science contexts. The platforms emerging in 2024-2025, while demonstrative of feasibility, remain concentrated in domains where automation is either technically simpler or economically prioritized (Tang et al., 2025; Lu et al., 2024).

Disciplinary differences profoundly shape the efficacy and efficiency of research automation. In experimental sciences with robotic laboratory integration, artificial researchers can execute standardized protocols with high reliability (Merchant et al., 2023). By contrast, humanities and social science research engages heterogeneous source materials, interpretive ambiguities, multilingual corpora, and complex historical contexts—conditions that substantially complicate automation (Colavizza et al., 2022).

Furthermore, the structure of services and infrastructures with which artificial researchers must interface critically determines system feasibility and reliability. Institutional repositories, publisher-controlled bibliographic databases, library systems, archives, and proprietary AI-based services each present different access protocols, metadata standards, query capabilities, and usage restrictions (Groppe et al., 2025; Nguyen et al., 2025).

These factors mean that artificial researchers capable of sophisticated operations in controlled computational domains may perform inadequately when confronted with archival complexity. Current AI systems, including large language models, demonstrate persistent limitations in sustained reasoning, factual accuracy, and handling of specialized domain knowledge—limitations that directly impact their utility in humanities research contexts (Beel et al., 2025; Marcus, 2024).

Consequently, although the trajectory toward automated research systems in the humanities is evident, substantial refinement of techniques, standardization of interfaces, and development of specialized capabilities remain necessary. The systems discussed herein should be regarded as emergent rather than mature, with significant development required before achieving the reliability and sophistication observed in more advanced automation domains.

4. INTERNAL AND EXTERNAL INTEGRATION OF ARTIFICIAL RESEARCHERS

Two primary integration models can be identified for artificial researchers within archival ecosystems, each with distinct technical architectures and institutional implications.

External Access Model: In this configuration, archives provide programmatic access—typically via Application Programming Interfaces (APIs)—enabling external automated research systems to query collections, retrieve metadata, and access digitized materials. Researchers interact with their own artificial research platforms, which in turn interface with archival infrastructures. This model requires archives to establish governance frameworks that regulate machine access, define query limits, specify permissible operations, and monitor usage patterns.

Internal Deployment Model: Here, archives develop and maintain internal artificial archivist systems—intelligent agents capable of receiving inputs from external users (whether human researchers or external automated systems) and executing complex operations including semantic search, contextual guidance, metadata enrichment, and research path recommendation (Rim et al., 2025; Groppe et al., 2025; Nguyen et al., 2025). These internal systems function as sophisticated intermediaries, translating user queries into archival-specific operations.

Both models shift archival access from expert-dependent to system-mediated processes. Traditional archival research demands considerable tacit knowledge, including understanding institutional organization, mastering finding aids, interpreting archival classification systems, and navigating physical repository structures. Artificial researchers and digital archivists can encode and operationalize this expertise, thereby making it accessible to a broader range of users. However, both models also raise important questions regarding trust, interpretive authority, and the transformation of archival expertise from human skill to algorithmic operation (Jaillant & Rees, 2023).

5. BENEFITS AND PROSPECTS FOR PUBLIC ENGAGEMENT

The integration of artificial researchers into archival systems carries profound implications for public engagement—a dimension increasingly central to both Digital Humanities practice and archival mission (Ridge, 2014; Owens, 2018).

Traditional archival research entails substantial financial barriers. A doctoral researcher conducting multi-site archival work across European institutions can easily incur substantial costs in travel, accommodation, and subsistence—often several thousand euros for a single multi-site research trip. Early-career scholars, independent researchers, and interested citizens often lack access to institutional funding mechanisms that enable such expenditures. These economic barriers directly constrain who can conduct archival research and what research questions can be feasibly pursued.

Artificial researcher-mediated access dramatically reduces these barriers. Remote interaction with digital archives via intelligent systems eliminates travel requirements, compresses research timelines, and reduces overall project costs to marginal expenses of connectivity and computational resources. This economic democratization expands potential researcher populations to include those previously excluded by resource constraints.

Beyond economic barriers, archival research requires specialized competencies: facility with finding aids, understanding of archival organization principles, knowledge of institutional histories, and often language skills for historical documents. These competencies constitute significant entry barriers for non-specialists.

Artificial archivist systems (in practice, an “internal artificial researcher”) can mediate these requirements, guiding users through complex archival structures and translating general research interests into specific document-retrieval strategies. Academic and non-academic researchers can interact with archival systems that interpret their queries, propose feasible research directions, and facilitate document access—without requiring years of archival training.

Crucially, these systems can perform feasibility assessment: rather than requiring users to independently determine whether adequate sources exist for a given research question, artificial archivists can evaluate available holdings and propose research topics known to be documentarily supported. This guidance function reduces wasted effort and frustration, particularly for inexperienced researchers who might otherwise invest substantial time pursuing questions for which sources prove inadequate. Understanding collection organization, navigating finding aids, and physically accessing materials consume substantial

time—time that constitutes another form of resource barrier. Professional researchers can dedicate this time; many potential users cannot.

Automated systems compress research timelines by performing rapid searches across large collections, identifying relevant materials through semantic analysis rather than manual review, and presenting synthesized results rather than requiring users to independently process extensive metadata. What might require many days of on-site work can potentially be accomplished in a few hours of remote interaction. These combined effects—economic, competential, and temporal accessibility—fundamentally expand potential archival user populations. This demographic expansion aligns with contemporary emphases on public engagement, citizen science, and participatory research (McClure et al., 2020). By lowering entry barriers, artificial researchers enable archives to function as genuinely public research infrastructures rather than specialist-dominated institutions. The democratization of AI tools more broadly reinforces these possibilities, as users become more familiar with conversational interfaces and AI-mediated information access (van der Sommen & de Groof, 2024).

6. INSTITUTIONAL PERSPECTIVES AND CRITICAL CONSIDERATIONS

From archival institutional perspectives, artificial researcher integration presents both opportunities and tensions. Understanding these dynamics is essential for a realistic assessment of likely trajectories. The exploratory discussions with senior staff at HAEU and the Historical Archives of the Bank of Italy conducted in 2025 suggest that research automation, while recognized as part of a broader digital transformation agenda, is not yet a prominent concern in the strategic planning of these institutions. This limited prominence appears to reflect two interrelated factors: ongoing regulatory uncertainty surrounding the use of AI in scholarly and cultural heritage infrastructures, and the absence of archival standards adapted to the opportunities and demands of automated research workflows. The institutional dialogue remains largely framed by the legacy agenda of digitization and metadata enrichment, rather than by the prospective integration of artificial researchers. The resulting gap between the trajectory of research automation in adjacent scientific domains and its current institutional visibility in the archival sector reinforces the case for anticipatory planning developed below.

Archives face mounting pressures regarding relevance, demonstrating impact, and justifying resources. Expanding user populations through artificial researcher-mediated access provides measurable impact metrics: increased access requests, broader demographic engagement, and enhanced public visibility. Furthermore, AI-mediated access can reduce operational costs associated with physical user accommodation—staff time devoted to user orientation, physical retrieval of materials, and facility maintenance for reading rooms. For institutions managing growing digital collections with constrained staff resources, automation offers operational efficiency.

However, the integration of artificial researchers also challenges traditional archival functions and professional identities. When intelligent systems mediate access, guide research, and facilitate retrieval, the archivist's expert mediating role is called into question. Professional archival work has traditionally centered on appraisal, arrangement, description, and reference services—functions that artificial systems may partially automate or reconfigure. This situation raises concerns not only about professional displacement but also about knowledge preservation, as archival expertise encompasses accumulated tacit knowledge regarding institutional histories, collection contexts, and interpretive frameworks. Encoding this knowledge algorithmically is a crucial step in the creation and training of artificial researchers and archivists.

Additionally, archival institutions must preserve their symbolic capital. For many archives, the physical facility serves as a material representation of institutional authority and prestige. Researchers visiting the archive experience not only document access but also an institutional encounter, reinforcing the significance of both the archival materials and the institutions that preserve them. Many institutions have invested substantial resources in archival facilities, and if access increasingly occurs remotely via artificial researchers, these investments may appear wasteful, with planned expansions or renovations potentially losing justification. At the same time, institutions recognize that expanded digital access enhances visibility and impact. However, algorithmic mediation introduces reputational risks, and traditional archival control over consultation becomes less effective in automated systems. Institutions must therefore develop governance mechanisms that preserve necessary oversight while enabling the expanded access promised by artificial researchers. They must also address significant technological investment requirements, as developing robust artificial archivist systems demands expertise in natural

language processing, semantic search, user interface design, and ongoing system maintenance. Smaller institutions may lack the resources for such development, potentially resulting in a two-tier system where major archives offer sophisticated automated access while others remain reliant on traditional methods. Finally, questions of authenticity, certification, and evidentiary value arise when access is algorithmically mediated. Mechanisms must be established to verify that retrieved materials accurately represent archival holdings and to ensure that artificial systems do not introduce interpretive biases or omissions. These concerns necessitate the development of new authentication and transparency protocols.

7. ANTICIPATORY PLANNING AND DIGITAL TRANSFORMATION

A central argument of this paper concerns timing: archival institutions and research communities must begin planning for the integration of artificial researchers *now*, even as technologies remain nascent. Several considerations justify this anticipatory approach.

First, the trajectory is clear even if the timeline remains uncertain. Artificial researchers have demonstrated feasibility in adjacent domains; extension to the humanities and social sciences is likely, not speculative.

Second, meaningful integration necessitates coordinated digital transformation across multiple dimensions. Artificial researchers require machine-readable materials, making it imperative for institutions to prioritize digitization efforts and establish standards that ensure materials are not only digitized but also structured for computational access. Additionally, rich and standardized metadata is essential for semantic search and intelligent retrieval, requiring investment in metadata schema development, controlled vocabularies, and entity linking. External access models also demand the creation of robust, well-documented APIs, while archival professionals must acquire new competencies through targeted training programs. Furthermore, institutions must establish policies governing automated access, including the definition of permissible operations, protection of sensitive materials, and assurance of ethical use. The development of these frameworks requires consultation among stakeholders and institutional deliberation.

Third, anticipatory planning enables collective action toward interoperability standards. One of the most significant current obstacles to the adoption of artificial researchers is the heterogeneity and incompatibility among systems. Institutional repositories employ divergent metadata schemas; publishers' platforms, even when universities hold legitimate subscription-based access rights, frequently deploy technical barriers that obstruct automated querying; archival systems lack standardised APIs; and proprietary AI services operate within closed ecosystems. These fragmented infrastructures compel artificial researchers to navigate multiple protocols, limiting both efficiency and reliability.

8. CONCLUSIONS

This paper has argued that artificial researchers and automated research platforms represent a structural transformation rather than marginal innovation within Digital Humanities and social science research. The archival domain offers a particularly instructive case. As collections become predominantly digital-native, traditional access modalities become inadequate. Artificial researchers—operating both as external systems interfacing with archival APIs and as internal digital archivists mediating user interactions—promise to fundamentally reconfigure access, expanding user demographics and reducing economic, competential, and temporal barriers to participation.

These developments carry profound implications for public engagement. By dramatically lowering entry barriers, artificial researcher-mediated archival access enables genuinely democratized primary source research. Archives can transition from specialist-dominated institutions to public research infrastructures serving diverse communities—from professional academics to citizen historians, educators, journalists, and community groups.

However, realizing these opportunities requires confronting significant technical challenges, navigating institutional tensions, and undertaking coordinated digital transformation. Artificial researcher systems remain partially nascent, particularly for humanities applications. Substantial refinement of techniques, standardization of interfaces, and development of domain-specific capabilities are necessary. Archival institutions face strategic decisions that balance democratizing access with preserving scholarly rigor and professional expertise.

Most critically, institutions must engage in anticipatory planning now. The trajectory toward automated research systems is sufficiently clear to justify proactive preparation, even as specific timelines remain uncertain.

ACKNOWLEDGEMENTS

The author wishes to thank the senior staff of the Historical Archives of the European Union (HAEU) in Fiesole and of the Historical Archives of the Bank of Italy in Rome for their availability and for the insights shared during the exploratory discussions that informed this analysis.

This contribution has developed in the context of the ATHENA SEED research project (University of Cagliari), which investigates the application of research automation and artificial researchers across epistemologically heterogeneous scientific domains.

AI assistance (Claude, Anthropic) was employed for initial literature synthesis, structural organization, and drafting support. All content has been reviewed, verified, and substantially revised by the author. Empirical material derives from direct research activities and institutional consultations conducted by the author.

REFERENCES

- Beel, J., Kan, M.-Y., & Baumgart, M. (2025). Evaluating Sakana's AI Scientist for autonomous research. arXiv preprint. <https://doi.org/10.48550/arXiv.2502.14297>
- Chapinal-Heras, D., & Díaz-Sánchez, C. (2023). A review of AI applications in human sciences research. *Digital Applications in Archaeology and Cultural Heritage*, 30, e00288. <https://doi.org/10.1016/j.daach.2023.e00288>
- 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), Article 13. <https://doi.org/10.1145/3479010>
- Ghareeb, A. E., Chang, B., Mitchener, L., Yiu, A., Szostkiewicz, C. J., Laurent, J. M., Razzak, M. T., White, A. D., Hinks, M. M., & Rodriques, S. G. (2025). *ROBIN: A multi-agent system for automating scientific discovery*. arXiv. <https://arxiv.org/abs/2505.13400>
- Gottweis, J., Weng, W.-H., Daryin, A., Tu, T., Palepu, A., Sirkovic, P., Myaskovsky, A., Weissenberger, F., Rong, K., Tanno, R., Saab, K., Popovici, D., Blum, J., Zhang, F., Chou, K., Hassidim, A., Gokturk, B., Vahdat, A., Kohli, P., Matias, Y., Carroll, A., Kulkarni, K., Tomasev, N., Dhillon, V., Vaishnav, E. D., Lee, B., Costa, T. R. D., Penadés, J. R., Peltz, G., Xu, Y., Pawlosky, A., Karthikesalingam, A., & Natarajan, V. (2025). *Towards an AI co-scientist*. arXiv. <https://arxiv.org/abs/2502.18864>
- Groppe, J., Marquet, A., Walz, A., & Groppe, S. (2025). Automated archival descriptions with federated intelligence of LLMs. arXiv preprint. <https://arxiv.org/abs/2501.08955>
- Jaillant, L., & Rees, A. (2023). Applying AI to digital archives: Trust, collaboration, and shared professional practices. *Digital Scholarship in the Humanities*, 38(2), 626–648. <https://doi.org/10.1093/llc/fqac087>
- Lu, C., Lu, C., Lange, R. T., Foerster, J., Clune, J., & Ha, D. (2024). The AI Scientist: Towards fully automated open-ended scientific discovery. arXiv preprint. <https://arxiv.org/abs/2408.06292>
- Marcus, G. (2024). What does it mean when an AI fails? *Communications of the ACM*, 67(1), 42–45. <https://doi.org/10.1145/3611097>
- McClure, E. C., Sievers, M., Brown, C. J., Buelow, C. A., Ditria, E. M., Hayes, M. A., Pearson, R. M., Tulloch, V. J. D., Unsworth, R. K. F., & Connolly, R. M. (2020). Artificial intelligence meets citizen science to supercharge ecological monitoring. *Patterns*, 1(7), 100109. <https://doi.org/10.1016/j.patter.2020.100109>
- Merchant, A., Batzner, S., Schoenholz, S. S., Aykol, M., Cheon, G., & Cubuk, E. D. (2023). Scaling deep learning for materials discovery. *Nature*, 624, 80–85. <https://doi.org/10.1038/s41586-023-06735-9>
- Nguyen, H. D., Nguyen, T.-H. A., & Nguyen, T. B. (2025). LLM-based smart search for archive systems: A new approach to information retrieval. arXiv preprint. <https://arxiv.org/abs/2501.09876>
- Owens, T. (2018). *The Theory and Craft of Digital Preservation*. Johns Hopkins University Press.
- Ridge, M. (2014). *Crowdsourcing Our Cultural Heritage*. Ashgate Publishing.

- Rim, K., King, O. C., Lynch, K., Verhagen, M., & Pustejovsky, J. (2025). AI-assisted archival metadata generation and enrichment. *Lecture Notes in Computer Science*, 13918, 234–248.
https://doi.org/10.1007/978-3-031-40839-2_17
- Shinde, G., Kirstein, T., Ghosh, S., & Franks, P. C. (2024). Artificial intelligence in archival science: A systematic literature review. arXiv preprint. <https://arxiv.org/abs/2411.01140>
- Tang, J., Xia, L., Li, Z., & Huang, C. (2025). AI-Researcher: Autonomous end-to-end scientific research through large language models. arXiv preprint. <https://arxiv.org/abs/2501.02471>
- van der Sommen, F., & de Groof, J. (2024). Risks and rewards of AI democratization in healthcare and beyond. *United European Gastroenterology Journal*, 12(4), 427–428.
<https://doi.org/10.1002/ueg2.12563>