ARCHEION
2024, T. CXXV, s. 55–78
ISSN 0066-6041
e-ISSN 2658-1264

DOI 10.4467/26581264ARC.24.006.20201

Artificial Intelligence and Machine Learning at the Intersection of Privacy and Archives[1]

Iori Khuhro

The University of British Columbia (Canada)
ORCID 0009-0002-6403-4149

Erin Gilmore

San José State University (USA)
ORCID 0009-0008-0249-1954

Jim Suderman

InterPARES Trust AI Project (Canada)
suderman.mawg@gmail.com

Darra L. Hofman

San José State University (USA)
darra.hofman@sjsu.edu, ORCID 0000-0002-1772-6268

ABSTRACT
As records are increasingly born digital – and thus, at least ostensibly, potentially much more accessible – archivists find themselves struggling to enable general access while providing appropriate privacy protections for the torrent of records being transferred to their care. In this article, the authors report the results of an integrative literature review study, examining the intersection of AI, archives, and privacy in terms of how archives are currently coping with these challenges and what role(s) AI might play in addressing privacy in archival records. The study revealed three major themes: 1) the challenges of – and possibilities beyond –

KEYWORDS
archival science, privacy, artificial intelligence, machine learning, named entity recognition

[1] The authors are grateful that this work was supported by International Research on Permanent Authentic Records in Electronic Systems (InterPARES) Trust AI, an international research partnership led by Drs. Luciana Duranti and Muhammad Abdul-Mageed, University of British Columbia. InterPARES Trust AI is supported in part by funding from the Social Sciences and Humanities Research Council of Canada (SSHRC). The authors would like to thank Kisun Kim (Okanagan College) and Carlos Quevedo, previous InterPARES Trust AI Graduate Research Assistants, for their contribution to this work.

defining "privacy" and "AI"; 2) the need for context-sensitive ways to manage privacy and access decisions; and 3) the lack of adequate "success measures" for ensuring the actual fitness for purpose of privacy AI solutions in the archival context.

Sztuczna inteligencja i uczenie maszynowe na styku prywatności i archiwów

STRESZCZENIE

W miarę tego, jak dokumentacja w coraz większym stopniu tworzona jest w formacie cyfrowym – a tym samym, przynajmniej pozornie, jest potencjalnie bardziej dostępna – archiwiści zmagają się z zapewnieniem powszechnego dostępu do niej, przy jednoczesnym zagwarantowaniu odpowiedniej ochrony prywatności w odniesieniu do ogromnej liczby dokumentów przekazywanych pod ich opiekę. W niniejszym artykule autorzy przedstawiają wyniki przeglądu literatury, badając obszar styku pomiędzy sztuczną inteligencją (AI) a archiwami i prywatnością pod kątem tego, jak placówki archiwalne radzą sobie współcześnie z tymi wyzwaniami i jaką rolę może odegrać sztuczna inteligencja w ochronie prywatności w archiwach. Badanie ujawniło trzy główne obszary tematyczne: 1) wyzwania związane z definicjami „prywatności" i „sztucznej inteligencji" oraz otwierające się nowe możliwości; 2) potrzebę powstania wrażliwych na kontekst sposobów zarządzania prywatnością i decyzjami dotyczącymi dostępu; oraz 3) brak odpowiednich „mierników sukcesu", które gwarantowałyby faktyczną przydatność rozwiązań AI chroniących prywatność w kontekście archiwalnym.

SŁOWA KLUCZOWE
archiwistyka, prywatność, sztuczna inteligencja, uczenie maszynowe, rozpoznawanie nazwanych jednostek

Introduction

Archivists and records professionals have long endeavoured to balance privacy and access. According to the U.N., privacy laws, in one form or another, exist in 137 of 194 countries[2]. The combination of privacy protection laws and digital technologies have made striking that balance increasingly difficult. It is no longer possible to presume "privacy by obscurity", which is the assumption that analogue records remain private, in part, because it is just too much trouble for anyone without a strong (and presumably legitimate) interest in those records to go to where they are stored and leaf through them[3]. The advent of digital technologies has allowed for mass records aggregation, transmission, and filtering to be performed quickly, easily, and inexpensively. It is within the capacity of private organizations and state actors alike to assemble and maintain vast dossiers on private individuals. This new landscape has led to growing concerns around individual privacy and data protection.

[2] UNCTAD. Data Protection and Privacy Legislation Worldwide, https://unctad.org/page/data-protection-and-privacy-legislation-worldwide [access: 5.11.2024].

[3] D.J. Solove, *Access and Aggregation: Privacy, Public Records, and the Constitution,* "Minnesota Law Review" 2002, vol. 86, pp. 1137–1217.

Recent advances in artificial intelligence (AI), especially machine learning and generative AI, while holding out the promise that machines will be able to accomplish what humans have not has, in fact, further complicated the privacy-access relationship. This is, in part, because the privacy/technology relationship is multifaceted. Researchers from Australia surveyed machine learning and privacy research initiatives from three perspectives:

- privacy of the machine learning model and related data;
- the use of machine learning models to enhance privacy protection;
- the use of machine learning models to breach privacy protections[4].

Because new information technologies, including AI, will be applied in all three ways, technology alone seems unlikely to solve the privacy protection challenge – at least not in the foreseeable future. Although all organizations find themselves faced with complying with privacy and data protection regulations, archival institutions, whose *raison d'être* includes providing access to records, must navigate digital privacy concerns with an eye to both current and future access needs, preservation and records trustworthiness, and not just legal, but ethical obligations to a number of stakeholders, including data subjects. In 1998, Paul Sillitoe noted that, "[i]n this fast-moving environment, archives and records services are about to be caught in new legislation for which they were not the primary target. Data protection today, freedom of information tomorrow. Whether we like it or not, we are involved"[5]. This study seeks to understand the relationship between privacy and archives since the emergence of ubiquitous AI.

Methods

This study utilized an integrative literature review as it is an approach that allows researchers to "create initial or preliminary conceptualizations and theoretical models […][and] to combine perspectives and insights from

[4] B. Liu et al., *When Machine Learning Meets Privacy: A Survey and Outlook*, "ACM Computing Survey" 2021, vol. 54, no. 2, article 31, pp. 1–36, https://doi.org/10.1145/3436755 [access: 5.11.2024].

[5] P. Sillitoe, *Privacy in a public place: Managing public access to personal information controlled by archives services,* "Journal of the Society of Archivists" 1998, vol. 19, no. 1, pp. 5–15, p. 5.

different fields or research traditions"[6]. The initial research questions posed for the literature review, listed below, were set to try and capture the multifaceted relationships between archives, privacy, and AI:

1. How are archival institutions dealing with protecting privacy in digital records containing Personal Information (PI)[7] when providing access to them?
2. How could AI tools and techniques contribute to the challenges faced by archival institutions in providing access to these kinds of records?
3. What are the implications of using AI tools and techniques to deal with privacy issues in records?
4. How effectively can machine learning (ML), natural language processing (NLP), and named entity recognition (NER) enable the identification and location of personal information in large digital textual collections?

Based upon these initial questions, we began an iterative review of the literature. In screening for inclusion, our initial criteria included: date, peer review, type of publication, research setting, and research design. Searches were conducted in six databases and, of the almost 35,000 results returned, 52 were included in the chart[8].

Figure 1: Relevance criteria applied to articles for inclusion in the scope of the study

Criterion	Initial Requirements	Expanded
Date	2017 and subsequent; initially chosen due to the breakthroughs in AI	Yes
Type of publication and peer review	Peer-reviewed journal articles and conference proceedings	Yes
Research setting	Inclusive	No
Research design	Inclusive	No

Author's own elaboration.

[6] H. Snyder, *Literature review as a research methodology: An overview and guidelines*, "Journal of Business Research" 2019, vol. 104, pp. 333–339, https://doi.org/10.1016/j.jbusres.2019.07.039 [access: 5.11.2024].

[7] The glossary of the International Association of Privacy Professionals (IAPP) notes that the terms "Personal Information" and "Personal Data" are synonymous. "Personally Identifiable Information" (PII), while not indicated as synonymous to the other two terms, likewise refers to "any information […] that can be used to distinguish or trace an individual's identity". IAPP. Glossary, https://iapp.org/resources/glossary/#paperwork-reduction-act-2 [access: 5.11.2024].

[8] The databases consulted were: 1) ACM Digital Library; 2) IEEE; 3) Jstor; 4) LISTA; 5) ProQuest; 6) Taylor & Francis.

When this research began in 2021, the AI revolution was at an earlier stage and the academic literature was therefore not as deep. In particular, academic literature discussing the ways in which archives, specifically, were balancing access and privacy was not as rich as it has since become. Furthermore, particularly within the privacy literature, it became apparent that fundamental work that preceded our initial cutoff date strongly influences the ways in which privacy is understood. As can be seen, our initial review of the data led us to expand our inclusion criteria to include earlier publications (particularly in the realms of privacy and digital archives) and relevant grey literature (particularly within the AI/privacy domain).

We initially charted the objectives, research questions, core concepts, research setting, research design, key findings, and implications for each article into a shared spreadsheet, with at least two researchers independently charting each article. The team reviewed the literature in an iterative fashion, and points of disjunction in the analysis were addressed through team consultation and reconciliation.

In the refined version of our data analysis, we also charted "type of study" (archival/legal/computer science); jurisdiction; privacy scope (from the very broad, such as "private user data" to very specific types of personal data, such as "email addresses, email messages, and headers"); how the study deals with privacy; success measures; whether human intervention was needed; and novel model ideas.

Results

This study, which is still underway to further triangulate the findings with more recent literature, has revealed three major themes: 1) the challenges of – and possibilities beyond – defining "privacy" and "AI"; 2) the need for context-sensitive ways to manage privacy and access decisions; and 3) the lack of adequate "success measures" for ensuring the actual fitness for purpose of privacy AI solutions in the archival context.

What is "Privacy"? What is "Artificial Intelligence"?

It is almost *pro forma* for privacy articles to begin by noting that privacy is difficult to define. What emerged from this study is not just about the challenge of defining the term but about "definition" versus "understanding" because "privacy" is a contested, polysemous term. The legal scholar Dan Solove writes of privacy: "Privacy is a concept in disarray. Nobody can articulate what it means"[9]. Haejung Yun *et al.* confirm this by stating that "[t]he nature and degree of [personal information privacy] PIP concerns may vary in different contexts because privacy means different things to different individuals in different contexts"[10].

Most people have a sense of what they mean by "privacy", but it is a term that takes a multitude of concepts, technologies, and approaches in its sweep. A number of taxonomies have been presented for privacy[11], with entire literatures devoted to different privacies, such as informational privacy (the primary concern of archivists), bodily privacy, and relational privacy – often with the caveat that these categories typically overlap. Legal scholar Joshua Fairfield questions not just the need, but the effectiveness of trying to grapple the definition of privacy to the ground: "A word's meaning is not its definition but its use"[12]. Despite the definitional challenges, archivists continue to receive, preserve, and provide (or deny) access to records, doing their best to comply with relevant laws and meet ethical obligations to protect privacy.

With regard to use, the relationship between privacy and records is long and deep. The Fair Information Practice Principles (FIPPs, also referred to as the Fair Information Practices, FIPs), which form the foundation of much modern privacy regulation and practice, originated from a 1973 report from the U.S. Department of Health, Education, and Welfare (HEW) entitled "Records, Computers and

[9] D.J. Solove, *A Taxonomy of Privacy*, "University of Pennsylvania Law Review" 2006, vol. 145, no. 3, pp. 477–564, https://doi.org/10.2307/40041279 [access: 5.11.2024].

[10] H. Yun, G. Lee, D.J. Kim, *A Chronological Review of Empirical Research on Personal Information Privacy Concerns: An Analysis of Contexts and Research Constructs*, "Information & Management" 2019, vol. 56, no. 4, pp. 570–601, p. 571, https://doi.org/10.1016/j.im.2018.10.001 [access: 5.11.2024].

[11] D.J. Solove, *A Taxonomy*…; B.J. Koops, B.C. Newell, T. Timan, T. Chokrevski, *A Typology of Privacy*, "University of Pennsylvania Journal of International Law" 2017, vol. 38, no. 2, pp. 483–578; J. Heurix, P. Zimmermann, T. Neubauer, S. Fenz, *A taxonomy for privacy enhancing technologies,* "Computers & Security" 2015, vol. 53, pp. 1–17.

[12] J.A. Fairfield, *"You Keep Using That Word": Why Privacy Doesn't Mean What Lawyers Think*, "Osgoode Hall Law Journal" 2002, vol. 59, pp. 249–290.

the Rights of Citizens: Report of the HEW Advisory Committee on Automated Personal Data Systems"[13] The FIPPs, which range from accountability and authority to security and transparency, "operationalize important values like dignity and autonomy"[14]. These values are incorporated into the principles of the General Data Protection Regulation (GDPR) and other data protection regulations globally.

These old principles remain foundational to an archival approach to privacy in part because the potential privacy problems attendant on computerized technologies, including technologies under the artificial intelligence umbrella, have long been known. Privacy scholar Alan Westin noted in his seminal *Privacy and Freedom*, "The fact that »machine-to-machine reporting« is spreading data from agency to agency through the [U.S.] federal system was […] noted by a congressional study in 1963"[15]. The HEW report identified three ways in which "computerization" changes recordkeeping to the detriment of privacy. These changes are further entrenched in the use of artificial intelligence technologies. Consider the following passage from the HEW report; one could easily replace the word "computerization" with "artificial intelligence" and reproduce much of the work currently done around privacy and AI:

"The dangers latent in the spread of computer-based personal-data record keeping stem […][from the fact that]:

- Computerization enables an organization to enlarge its data-processing capacity substantially.
- Computerization greatly facilitates access to personal data within a single organization, and across boundaries that separate organizational entities.
- Computerization creates a new class of *record keepers* whose functions are technical and whose contact with original suppliers and ultimate users of personal data are often remote"[16] (emphasis added).

The widespread adoption of AI exacerbates the challenges identified in the HEW report, as training and/or using AI requires extensive data processing

[13] W.H. Ware, *Records, computers and the rights of citizens* ["Report of the Secretary's Advisory Committee on Automated Personal Data Systems", Washington 1973], https://aspe.hhs.gov/reports/records-computers-rights-citizens [access: 5.11.2024].

[14] W. Hertzog, *Privacy's Blueprint: The Battle to Control the Design of New Technologies*, Cambridge, Massachusetts 2018, p. 61.

[15] A.F. Westin, *Privacy and Freedom*, New York 1967, p. 343.

[16] W.H. Ware, *Records, computers*… The quoted text is from "II. Latent Effects of Computer-based Record Keeping".

capacity, often facilitates the flow of information across organizational boundaries, and creates new record keepers whose contact with the suppliers and users of personal data are remote.

As with privacy, AI has become a sprawling concept, seemingly including in its boundaries whatever new thing we'd like computers to do. Will Heaven writes that "[...] AI has come to mean all things to all people, splitting the field into fandoms. It can feel as if different camps are talking past one another, not always in good faith"[17].

The ethical imperative to protect privacy and minimize harm within the AI field is evident in the extensive literature in the computer science domain regarding efforts to predict, identify, and automatically redact, anonymize or pseudonymize personal data. It is clear from this literature that AI has been deployed to protect privacy in many "low-hanging fruit" cases, such as the use of NER to identify personal information having a standard form such as (but not limited to) social insurance numbers or telephone numbers. Identifying substantial portions of "personal information", however, and especially "sensitive information" such as ethnicity or sexual orientation has often been beyond the scope of the available artificial intelligence tools because the personal/sensitive nature of the information is complicated as it is frequently contextually dependent. This does not mean that artificial intelligence could not be developed to better identify personal information based on context. However, recent failures, such as the suggestion provided by a generative AI tool that one glue cheese to a pizza to keep it from sliding off[18], and vulnerabilities, such as penetrating the model to gain information about the underlying training data[19], make clear that we're a long way from relying on AI for privacy.

[17] W.D. Heaven, *What is AI?,* "MIT Technology Review", 10 July 2024, https://www.technologyreview.com/2024/07/10/1094475/what-is-artificial-intelligence-ai-definitive-guide/ [access: 5.11.2024].

[18] S. Ovide, *Why Google's AI might recommend you mix glue into your pizza,* "The Washington Post", 24 May 2024, https://www.washingtonpost.com/technology/2024/05/24/google-ai-overviews-wrong/ [access: 5.11.2024].

[19] B. Liu, M. Ding, S. Shaham, *When Machine Learning Meets Privacy...*, pp. 7–11.

The Role of Context in Managing Access and Privacy

Privacy is both a legal and ethical dilemma for archivists when making decisions regarding access to records. Digital technologies have made managing privacy, i.e. effectively balancing the value of accessible records against the harm of releasing records containing personal information, dramatically more difficult in the wake of "access and aggregation"[20]. Indeed, archivist Malcolm Todd, reflecting on the challenges of privacy legislation, argued in 2006 that archivists "shall have to address concerns in juridical systems that are explicitly »about« privacy and personal data and neither understood nor articulated in archival terms"[21].

Privacy is not the only challenge in the way of providing access to archival records in the digital era, of course, but it is a significant one. Jason Baron and Nathaniel Payne noted in 2017 that the U.S. National Archives was preserving almost 300 TB of White House emails but that "none have been systematically opened by archivists for public access, nor is there any strategic plan for doing so in the immediate future"[22]. This highlights two key considerations. The first is that the acquisition and preservation of digital records are proceeding apace; substantial digital backlogs are developing. The second is that many (perhaps even most) archival institutions do not have any strategies in place to enable general access to those records. There are simply far too many records to ever be reviewed individually, a problem noted also in literature addressing opening court records[23]. Given this sobering reality – and putting aside the fact that AI

[20] D.J. Solove, *Access and Aggregation…*

[21] M. Todd, *Power, Identity, Integrity, Authenticity, and the Archives: A Comparative Study of the Application of Archival Methodologies to Contemporary Privacy*, "Archivaria" 2006, vol. 61, p. 185.

[22] J. Baron, N. Payne, *Dark Archives and E-democracy: Strategies for Overcoming Access Barriers to the Public Record Archives of the Future* [in:] *Conference for E-Democracy and Open Government (CeDEM)*, eds. P. Parycek, N. Edelmann, Krems 2017, pp. 3–11, https://doi.org/10.1109/CeDEM.2017.27 [access: 5.11.2024].

[23] See D. Ardia, A. Klinefelter, *Privacy and Court Records: An Empirical Study*, "Berkeley Technology Law Journal" 2015, vol. 30, no. 3, pp. 1807–1898; M. Tamper et al., *Anonymization Service for Finnish Case Law: Opening Data without Sacrificing Data Protection and Privacy of Citizens*, 2018, https://research.aalto.fi/en/publications/anonymization-service-for-finnish-case-law-opening-data-without-s [access: 5.11.2024]; A. Oksanen et al., *ANOPPI: A Pseudonymization Service for Finnish Court Documents* [in:] *Legal Knowledge and Information Systems*, eds. M. Araszkiewicz, V. Rodríguez-Doncel, Amsterdam 2019, pp. 251–254, https://helda.helsinki.fi/server/api/core/bitstreams/622773b4-8c6e-4558-8571-da432fe7ea8f/content [access: 5.11.2024]; I. Glaser, T. Schamberger, F. Matthes, *Anonymization of German legal court rulings* [in:] *Proceedings of the Eighteenth International Conference on Artificial Intelligence*

systems are themselves generating records that will need to be dealt with – can AI tools and techniques help reduce the backlog and increase records accessibility?

Victoria L. Lemieux and John Werner explain in their scoping review of privacy-enhancing technologies for archives that, despite experimentation with AI-enabled (predominantly NLP-based) approaches, effective ways to responsibly balance the provision of access with the protection of privacy remain elusive[24]. Legal scholar Paul Ohm asserts that the assumption "that anonymization protects privacy", on which most privacy regulation in the U.S. and Europe is based, is discredited. He also observes re-identification techniques make privacy protections in some laws, eg, the U.S. Health Insurance Portability and Accountability Act (HIPAA), "illusory and underinclusive because it deregulates the handling of types of data [eg. data pertaining to patient visits to hospitals], that can still be used to reidentify and harm"[25]. The same assumption makes Europe's GDPR "essentially boundless" as every new type of data used for re-identification must, from that point forward, come under its scope. Therefore, despite ever more sophisticated technological means of removing information that directly or indirectly identifies an individual, the possibility of re-identification is never eradicated. AI itself can and will be used to re-identify anonymized data. The research in this area shows that training models to identify personal information remains highly time-consuming, incomplete, and challenging due to the unavailability of appropriate training data sets containing personal information[26]. Deploying privacy tools that are insufficiently accurate could erode trust in both the tools and the archival institutions that might use them.

and Law, New York 2021, pp. 205–209, https://doi.org/10.1145/3462757.3466087 [access: 5.11.2024]; D. Garat, D. Wonsever, *Automatic Curation of Court Documents: Anonymizing Personal Data,* "Information" 2022, vol. 13, no. 27, https://www.mdpi.com/2078-2489/13/1/27 [access: 5.11.2024]

[24] V.L. Lemieux, J. Werner, *Protecting Privacy in Digital Records: The Potential of Privacy-Enhancing Technologies*, "Journal on Computing and Cultural Heritage" 2024, vol. 16, no. 4, article 83, pp. 1–18.

[25] P. Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, "UCLA Law Review" 2010, vol. 57, pp. 1741–1742.

[26] See for example P. Silva et al., *Using NLP and Machine Learning to Detect Data Privacy Violations* [in:] *IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, Toronto 2020, pp. 972–977, https://doi.org/10.1109/INFOCOMWKSHPS50562.2020.9162683 [access: 5.11.2024].

These challenges have not deterred archivists from continuing to make decisions on access to holdings in their care in ways that respect both the ethical and legal requirements of privacy. By approaching privacy, openness, and access as dimensions of how records are approached, rather than a hard binary – open or closed – archivists are searching for solutions to the problems of "access and aggregation" that allow the records, especially the born digital records, in their care to find transformative uses. A thorough examination of the privacy/access balance and the appropriateness of available tools including AI requires a critical consideration of such well-worn concerns as records' provenance, institutional mandate, and both ethical and legal obligations to the records' stakeholders. As Angeliki Tzouganaotu reminds us, "Questions about why to open up, whom to open up to, the level of openness and the quality of the process's nature for opening itself up are critical"[27].

With the outlined circumstances in mind, the theory of "contextual integrity" provides an alternative approach to privacy protection. Contextual integrity considers privacy as a relative rather than a static concept[28]. The use and dissemination of personal information can be appropriate or not based on their social settings, characterized by social norms, power structure, and internal values. Contextual integrity posits that one's privacy is not always violated when a certain piece of information is shared, but rather when it is shared in an unexpected context or way[29]. Steven Bingo suggests that, by using contextual integrity, archives can identify privacy risks by examining the contexts of the record creation, such as the creators' role and activities, during the appraisal process[30]. In other words, the provenance of a body of records should reveal the context and privacy norms that should govern access. Therefore, analyzing the provenance arguably could allow archives to evaluate the privacy risk without requiring a document-level review. Indeed, Joshua Fairfield points to NLP as a potential route for understanding "how people actually use privacy-related

[27] A. Tzouganatou, *Openness and privacy in born-digital archives: reflecting the role of AI development*, "AI & Society" 2022, vol. 37, p. 993.

[28] H.F. Nissenbaum, *Privacy in context: Technology, policy, and the integrity of social life*, Stanford 2009, p. 132.

[29] Ibidem, p. 140.

[30] S. Bingo, *Of Provenance and Privacy: Using Contextual Integrity to Define Third-Party Privacy*, "The American Archivist" 2011, vol. 74(2), pp. 506–521, https://doi.org/10.17723/aarc.74.2.55132839256116n4 [access: 5.11.2024].

language"[31], reminding us that NLP's major breakthrough was due to the availability of enormous amounts of training data which allowed the algorithms to identify and reproduce the patterns of human language use. However, the empirical work on contextual integrity thus far does not seem to have taken advantage of archival science's deeply developed models of records' contexts.

As noted above, the relationship between privacy, archives, and AI is multifaceted. Simply relying on AI solutions to solve the problem of balancing privacy and access risks further entrenching known issues in both AI and archives. However, combining AI-enabled approaches with archival knowledge and practice, such as rich description of provenance, may ameliorate existing problems but still fall short of regulatory compliance.

Meera Desai *et al.* note that the presence of personal information is an issue of common concern to Large Language Model (LLM) researchers with regard to pretraining datasets and describe some of the limitations of measures taken to assess privacy vulnerabilities as well as toxicity and data contamination. Two approaches to reduce privacy vulnerabilities are through redacting personal information and deduplication, but what counts as PII or duplication – and whether these are even sufficient to address privacy concerns – is often unaddressed in this work. Measuring privacy vulnerabilities with PII and duplicates assumes that privacy is discrete and that privacy leakages are the only form of privacy risk[32].

The authors note that despite meaningful differences, there are common features and similarities between pretraining datasets and archives: "both are collections of diverse sociocultural materials that mediate knowledge production and thereby confer power to those who select, document, and control access to them"[33]. They conclude by suggesting that developers of pretraining datasets might benefit from improved documentation detailing not only what data were used "but also why and how data were chosen, appraised, and excluded" and better tools for finding and assessing pretraining datasets. Such documents might contribute more effectively in minimizing harm resulting from LLMs trained on current datasets by helping LLM users to "understand the limitations of these models"[34].

[31] J.A. Fairfield, *"You Keep Using That Word"*..., p. 284.

[32] M.A. Desai, I.V. Pasquetto, A.Z. Jacobs, D. Card, *An Archival Perspective on Pretraining Data*, "Patterns" 2024, vol. 5, no. 4, pp. 1–11.

[33] Ibidem, p. 1.

[34] Ibidem, p. 6.

Defining “Success”: Setting meaningful privacy success measures

We sought to capture reported “success measures”, with a particular goal of understanding how the question of AI-privacy solution “effectiveness” was being addressed. Papers from the AI/ML field tended to report standard success measures for their field, including precision, recall, accuracy, and/or F1 scores, which serve as adequate measures for determining how well an algorithm identifies true and false positives or negatives. In many experiments, success was sought for a limited number of types of personal information or a fully automated process, i.e. one without human involvement.

The research conducted by Diego Garat and Dina Wonsever neatly illustrates the issues with success measures noted above. Their research focused “on the de-identification of proper names” with the expectation that the “automation or semi-automation of pre-publication tasks could not only reduce the workload and publication time but also have a great impact on the user experience while interacting with the jurisprudence database”[35]. They first applied “state-of-the-art Natural Language Processing (NLP) tools” for the NER task[36]. Finding that these tools performed poorly, they retrained the SpaCy NER module and achieved a F1-micro score of 90.21%[37]. The second task was one of co-referencing, a sense-making task involving the replacement of each proper name with a unique label so that it would be clear to the document reader when the same (or a different) person was referenced. Here they achieved a 95.95% ARI score[38]. Despite these high scores, the authors concluded that “the total number of documents where all names are completely and correctly recognized and linked with each other is still below 50% of the validation set”[39]. Their research effectively illustrates the challenge of identifying appropriate success measures. It also illustrates

[35] D. Garat, D. Wonsever, *Automatic Curation of Court Documents…*, p. 3.

[36] The tools listed in the article are CoreNLP, Freeling and SpaCy.

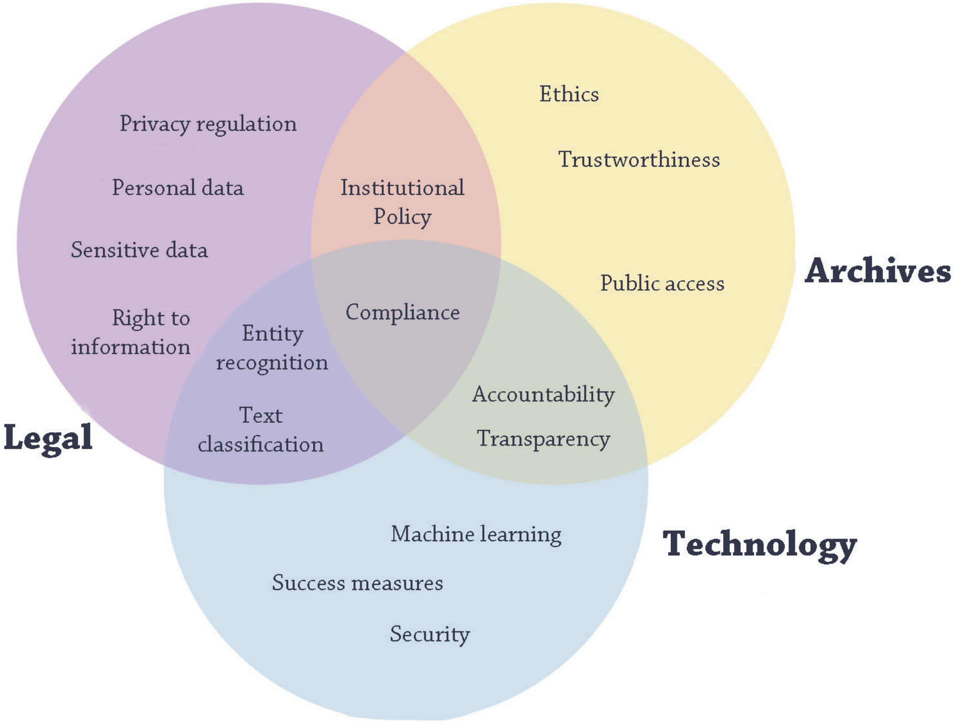
[37] The F1 metric “makes sense for multi-class data distributions” by combining “precision” and “recall’ scores. “»Precision« measures how many of the »positive« predictions made by the model were correct. »Recall« measures how many of the positive class samples present in the dataset were correctly identified by the model […][F1-micro] is calculated using »net« TP [True Positive], FP [False Positive], and FN [False Negative] values”. R. Kundu, *F1 Score in Machine Learning: Intro & Calculation*, 16 December 2022, https://www.v7labs.com/blog/f1-score-guide [access: 5.11.2024].

[38] An ARI (Adjusted Rand Index) score is a widely used metric for validating clustering performance.

[39] D. Garat, D. Wonsever, *Automatic Curation of Court Documents…*, p. 13.

the challenge of considering success measures for a much broader range of personal information than just the identification and removal of proper names. Determining whether or not information is personal, and to whom access to it should be given, seems to require something more. Given that privacy is used to represent a great diversity of human values, it is perhaps unsurprising that available success measures fall short of fully representing our sense of effective privacy management.

Figure 2: The Current Intersection of Archives, Technology, and Law for Privacy



Venn diagram created by the authors to reflect the privacy-focused intersection of archives, technology, and law.

The central zone of a notional representation of the dominant concepts unearthed in this study, shown in Figure 2 (above), contains only the relatively bounded “compliance”. In reality, ethics infuses both the legal and technological spheres, and archivists care deeply about security. But these issues – which all intersect in privacy – were generally treated in the reviewed literature as discrete, even siloed, concerns. Indeed, “success measures” as such were nearly exclusively discussed in the papers reporting the development of some form of AI/ML solution, regardless of whether those papers were published in computer science journals or archival journals.

"Success measures" in the AI/ML papers reviewed also assessed the performance of AI model's against that of a human, as is common in the development of supervised ML. Less reported in these studies was the measurement of the success of the *humans* annotating the data to train the models, typically measured by examining "inter-annotator agreement" (IAA). For example, the article by Gregory Rolan *et al*. describes a study where the New South Wales State Archives and Records (NSWSAR) piloted a program to use "off-the-shelf machine-learning software to the problem of classifying a corpus of unstructured data against a retention and disposal authority. The main aim was to test machine-learning algorithms on a corpus of records that had previously been manually sentenced against a disposal authority"[40]. NSWSAR used two machine-learning classification algorithms: Multinomial Naïve Bayes, which is a statistical model algorithm and the Multi-Layer Perceptron, which is a form of deep learning network. The Multi-Layer Perceptron had a success rate of 84% which, the authors noted, indicated that the technology would be adequate for "assisting with the classification and disposal of unclassified, unstructured data [but][…] is probably not yet human-level accuracy (though the actual human accuracy rate in this case is not known)"[41].

The advent of AI also heightens the urgency of longstanding questions around *who* defines success. As noted above, the most common measure of success for data annotation is inter-annotator agreement. In the case of archival annotation, we might assume that all annotators are archivists and therefore share certain professional notions that guide their annotation. However, given that privacy is not an archival concept *per se*, one must question whether that professional judgment is sufficiently developed amongst all archivists that IAA can actually serve as a measure of quality. Furthermore, the many meanings of privacy and the contextual nature of personal information means that there can be contradictory – but still correct – understandings of whether something is private, depending upon the context and the role, perspective, and lived experience of the annotator(s). This is particularly troublesome when one considers that "the

[40] G. Rolan et al., *More human than human? Artificial intelligence in the archive*, "Archives and Manuscripts" 2019, vol. 47, no. 2, p. 190.

[41] Ibidem, p. 193.

conceptual and practical dimensions of applying computational methods to digital archives may work conservatively to revivify notions of archival neutrality"[42], obscuring the contingent and ethically fraught nature that privacy decisions can carry beyond mere compliance.

"Community privacy" provides a useful perspective on the challenges facing humans in making privacy decisions. It is a concept with limited legal recognition, perhaps clearest in relation to the traditional knowledge of Indigenous communities that may form part of their cultural or spiritual identities[43]. Community privacy can be threatened when community information is either under- or over-represented in training. The problem of under-inclusive data is well-known, with data bias occurring at every stage of the "AI development and deployment lifecycle [...][as] a sequela of human, machine, and systems factors"[44]. Various measures have been developed to determine the success of efforts to minimize bias in training data and models. However, while much of the work on ethical AI has focused on the inclusion of marginalized communities, many communities have been historically overdocumented, a problem which propagates into AI systems and can be of particular concern to archivists.

Legal scholar Frank Pasquale notes that: "»Inclusion« can be as problematic as exclusion, becoming »predatory« or even »creepy«"[45]. This general "creepiness" comes with a heavier burden for marginalized communities, about whom records and data have often been created with little or no meaningful input or consent from its members. The fact that archives exist to serve the public good does not mean that they are good for all of the public. As Ellen LeClere argues in her article on privacy in large-scale archival digitisation projects:

> "The claim that archives – and by extension, digital archives – serve public interests within a liberal democracy is not uncontroversial. Archives in liberal democracies create a sense of accountability, transparency and access to information, but maintaining these values comes at the expense of asking margi-

[42] D. Mordell, *Critical Questions for Archives as (Big) Data*, "Archivaria" 2019, vol. 87, p. 140.

[43] WIPO. Genetic Resources, Traditional Knowledge and Traditional Cultural Expressions, https://www.wipo.int/tk/en/ [access: 5.11.2024].

[44] J.W. Gichoya et al., *AI pitfalls and what not to do: mitigating bias in AI*, "The British Journal of Radiology" 2023, vol. 96, no. 1150, p. 2, https://doi.org/10.1259/bjr.20230023 [access: 5.11.2024].

[45] F. Pasquale, *New laws of robotics: defending human expertise in the age of AI*, Cambridge 2020, pp. 133–135.

nalized groups for higher contributions for fewer benefits. This argument is also uncontroversial – access to archives has been historically controlled by privilege and power”[46].

This is not just an archives problem, nor is it merely an AI problem. Rather, challenges of inclusion, exclusion, privacy, and access exist in multifaceted relationships between archives, AI, and privacy. The biases in training datasets reflect the underlying biases in the data themselves. Data – including datafied records – emerge from a particular context. In both cases – underrepresentation and overrepresentation – members of the misrepresented groups may be harmed, or at least fail to receive the expected benefits of AI tools developed on biased datasets. For archives, whose records are increasingly being explored as a source of data, the ethical dimensions of access and privacy, as well as description, are heightened in these scenarios.

Discussion

It is clear that archivists, knowing that their collections have become inaccessible in order to comply with privacy protection regulations[47], are beginning to investigate AI for its potential to automate privacy protection and restore access[48]. The question, upon synthesizing the literature, is no longer whether AI can identify and then redact, anonymize or pseudonymize

[46] E. LeClere, *Breaking Rules for Good? How Archivists Manage Privacy in Large-Scale Digitisation Projects*, “Archives and Manuscripts” 2018, vol. 46, no. 3, pp. 289–308, p. 300, https://doi.org/10.1080/01576895.2018.1547653 [access: 5.11.2024].

[47] C.A. Lee, K. Woods, *Automated redaction of private and personal data in collections* [in:] *Proceedings of Memory of the World in the Digital Age: Digitization and Preservation International Conference*, eds. L. Duranti, E. Shaffer, Vancouver 2012, pp. 298–313, https://ils.unc.edu/callee/p298-lee.pdf [access: 5.11.2024]; B. Goldman, T.D. Pyatt, *Security without obscurity: Managing personally identifiable information in born-digital archives*, “Library & Archival Security” 2013, vol. 26, no. 1–2, pp. 37–55, https://doi.org/10.1080/01960075.2014.913966 [access: 5.11.2024].

[48] See J.R. Baron, N. Payne, *Dark archives and E-democracy…* See also T. Hutchinson, P*rotecting Privacy in the Archives: Preliminary Explorations of Topic Modeling for Born-Digital Collections* [in:] *Proceedings of the 2017 IEEE International Conference on Big Data, 25–30 June 2017, Honolulu, Hawaii*, eds. G. Karypis, J. Zhang, Los Alamitos 2017, pp. 2251–2255, https://harvest.usask.ca/items/e237ebe9-5627-44ac-8b2f-a61fc2e4acc3 [access: 5.11.2024]; T. Hutchinson, *Protecting Privacy in the Archives: Supervised Machine Learning and Born-Digital Records* [in:] *Proceedings 2018 IEEE International Conference on Big Data, 10–13 December 2018, Seattle*, ed. N. Abe et al.,

personal information. It is already established that it can do so, albeit imperfectly, for recognizable named entities[49], but rather can archivists, legal professionals, and computer scientists look beyond the existing attempts to define and apply privacy and begin to develop sufficiently rich, applied understandings of privacy that draw on the context-driven, human-centered nature of records to support the development of robust privacy AI solutions (and privacy *for* AI solutions).

Furthermore, the ongoing accumulation of auxiliary, i.e. non-personal, non-sensitive data, and the increasing sophistication of new tools for data analysis suggests that tools and techniques for de-identification, even if satisfactory today, may be inadequate tomorrow as AI is harnessed for re-identification. Complicating the matter further is that the grounds for deciding to fully restrict or fully open a body of records may change over time, rendering the original decision problematic. The professional judgment of the archivist in making access decisions requires an understanding of both regulatory obligations and human considerations that are loaded with a tacit understanding of the contextual factors that influence the sense of privacy attached to any given piece of information. Much like selection and appraisal, privacy and access decisions in reality are *interpretive*. Interpretation can, and should, be an exercise of both professional judgment and deep human respect. Even when it is both, it can also be wrong. Todd asks if there is "[...] a public interest in the archived collective memory that is higher than some of the mantras of the privacy lobby"[50]. Resolving this question will likely be deferred as long as the conviction exists that technological success is possible. This study has also made clear that alongside the technological innovations underway, there is a long-standing interdisciplinary dialogue regarding an effective balance between privacy protections and the social good of maintaining archives.

Protection of privacy is central to the ethical principles of both the archival and AI communities. For archivists, privacy is often understood as being balanced with the archival imperative to enable the widest possible access to information. Similarly, AI ethics emphasize principles of fairness, accountability, and transparency. In AI development, privacy is often pitted against innovation.

Piscataway 2018, pp. 2696–2701, https://doi.org/10.1109/BigData.2018.8621929 [access: 5.11.2024].

[49] A. Oksanen et al., *ANOPPI: A pseudonymization service...*

[50] M. Todd, *Power, Identity, Integrity, Authenticity, and the Archives...*, p. 185.

However, this study shows that by understanding privacy in broader ways, we can understand it, not as a binary standing in opposition to access and innovation, but as part of access and innovation strategies.

Conclusion

Most of the literature on AI, archives, and privacy reviewed by this study focused on developing AI solutions to protect privacy even while recognizing the challenges and complexity of doing so. As a consequence, evolving strategies for the adoption of AI by archivists for privacy protection will require thoughtful consideration of how archival collections should be shared and made available, both as records and as data. It will also require informed and competent choices of which AI tools to adopt and how to apply them responsibly within the archival domain.

The speed with which this technology is evolving has the potential to disrupt archival practice in a way that may tie trust in archives to trust in technology. At this point, the hype around the capabilities of AI with regard to protecting privacy still exceeds the (remarkable) progress towards building models and developing techniques to do so. If archivists are to grapple with the profound ethical implications of archival privacy in the AI world, we may well be forced to ask if and when it is necessary to deeply confront what we know and do. The risk of regulatory capture of digital privacy by technologists, often advanced by arguments that deep change will stifle innovation, shifts the burden to people like archivists to question whether and when a solution is simply too risky, too biased, or too problematic, asking ourselves: "what do we lose by buying into a logic of reformism when far more profound change is necessary?"[51].

The work of privacy in archives is difficult. Not because it is intellectually or even technologically insurmountable. But because it requires moral courage. If archivists are to have a meaningful voice in the development of AI and its impact on access to the records in our care, we must be willing to engage in substantive conversations regarding privacy based on our unique professional expertise. In particular, archivists bear the responsibility to speak from the perspective – as

[51] F. Pasquale, *New laws of robotics…*, p. 124.

incomplete and flawed as it necessarily is[52] – of a profession concerned with bearing trustworthy evidence of the past safely into the future. As Hilary Jenkinson wrote, "[t]he most common fault is haste in dealing with Archives, due to anxiety to make them available for use"[53]. Surely, then, as much as we embrace the possibilities of AI, we should also be a steadying hand, asking that privacy be understood and respected in view of the broader human systems of which any records (or data from records) are part.

Limitations and Future Work

This study is initial, exploratory work that is compiling and interpreting privacy expertise from the often siloed domains of law, computer science, and archives. It should be noted that the primary area of training and expertise for all members of the study team lies within archival science, and the review necessarily reflects that disciplinary perspective. Future work will focus on analyzing the values and limitations of computational/technical success measures for AI models against what is considered an acceptable, humanist attempt at protecting privacy within archival institutions by archivists.

Bibliography

Ardia D., Klinefelter A., *Privacy and Court Records: An Empirical Study,* "Berkeley Technology Law Journal" 2015, vol. 30, no. 3, pp. 1807–1898.

Baron J.R., Payne N., *Dark archives and E-democracy: strategies for overcoming access barriers to the public record archives of the future* [in:] *Conference for E-Democracy and Open Government (CeDEM)*, eds. P. Parycek, N. Edelmann, Krems 2017, pp. 3–11.

Bingo S., *Of Provenance and Privacy: Using Contextual Integrity to Define Third-Party Privacy,* "The American Archivist" 2011, 74(2), pp. 506–521, https://doi.org/10.17723/aarc.74.2.55132839256116n4 [access: 5.11.2024].

[52] H. Booms, *Überlieferungsbildung: keeping archives as a social and political activity*, "Archivaria" 1991, vol. 33, pp. 25–33; V. Harris, *The archival sliver: power, memory, and archives in South Africa*, "Archival Science" 2002, vol. 2, pp. 63–86.

[53] H. Jenkinson, *A Manual of Archive Administration: Including the Problems of War Archives and Archive Making,* London 1922, p. 66.

Booms H., *Überlieferungsbildung: keeping archives as a social and political activity*, "Archivaria" 1991, vol. 33, pp. 25–33.

Desai M.A., Pasquetto I.V., Jacobs A.Z., Card D., *An Archival Perspective on Pretraining Data*, "Patterns" 2024, vol. 5, no. 4, pp. 1–11.

Fairfield J.A., *"You Keep Using That Word": Why Privacy Doesn't Mean What Lawyers Think*, "Osgoode Hall Law Journal" 2002, vol. 59, pp. 249–290.

Garat D., Wonsever D., *Automatic Curation of Court Documents: Anonymizing Personal Data*, "Information" 2022, vol. 13, no. 27, pp. 1–16, https://doi.org/10.3390/info13010027 [access: 5.11.2024].

Gichoya J.W., Kaesha T., Celi L.A., Safad N., Banerjee I., Banja J.D., Seyyed-Kalantari L., Trivedi H., Purkayastha S., *AI pitfalls and what not to do: mitigating bias in AI*, "The British Journal of Radiology" 2023, vol. 96, no. 1150, pp. 1–8, https://doi.org/10.1259/bjr.20230023 [access: 5.11.2024].

Glaser I., Schamberger T., Matthes F., *Anonymization of German legal court rulings* [in:] *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Law*, New York 2021, pp. 205–209, https://doi.org/10.1145/3462757.3466087 [access: 5.11.2024].

Goldman B., Pyatt T.D., *Security without obscurity: Managing personally identifiable information in born-digital archives*, "Library & Archival Security" 2013, vol. 26, no. 1–2, pp. 37–55, https://doi.org/10.1080/01960075.2014.913966 [access: 5.11.2024].

Harris V., *The archival sliver: power, memory, and archives in South Africa*, "Archival Science" 2002, vol. 2, pp. 63–86.

Hertzog W., *Privacy's Blueprint: The Battle to Control the Design of New Technologies*, Cambridge, Massachusetts 2018.

Heurix J., Zimmermann P., Neubauer T., Fenz S., *A taxonomy for privacy enhancing technologies,* "Computers & Security" 2015, vol. 53, pp. 1–17.

Hutchinson T., *Protecting Privacy in the Archives: Preliminary Explorations of Topic Modeling for Born-Digital Collections* [in:] *Proceedings of the 2017 IEEE International Conference on Big Data*, *25–30 June 2017, Honolulu, Hawaii*, eds. G. Karypis, J. Zhang, Los Alamitos 2017, pp. 2251–2255, https://harvest.usask.ca/items/e237ebe9-5627-44ac-8b2f-a61fc2e4acc3 [access: 5.11.2024].

Hutchinson T., *Protecting Privacy in the Archives: Supervised Machine Learning and Born-Digital Records* [in:] *Proceedings 2018 IEEE International Conference on Big Data*, *10–13 December 2018, Seattle*, ed. N. Abe, H. Liu, C. Pu, X. Hu, N. Ahmed, M. Qiao, Y. Song, D. Kossmann, B. Liu, K. Lee, J. Tang, J. He, J. Saltz, Piscataway 2018, pp. 2696–2701, https://doi.org/10.1109/BigData.2018.8621929 [access: 5.11.2024].

Jenkinson H., *A Manual of Archive Administration: Including the Problems of War Archives and Archive Making,* London 1922.

Koops B.J., Newell B.C., Timan T., Chokrevski T., *A Typology of Privacy*, "University of Pennsylvania Journal of International Law" 2017, vol. 38, no. 2, pp. 483–578.

LeClere E., *Breaking Rules for Good? How Archivists Manage Privacy in Large-Scale Digitisation Projects*, "Archives and Manuscripts" 2018, vol. 46, no. 3, pp. 289–308, https://doi.org/10.1080/01576895.2018.1547653 [access: 5.11.2024].

Lee C.A., Woods K., *Automated redaction of private and personal data in collections* [in:] *Proceedings of Memory of the World in the Digital Age: Digitization and Preservation International Conference,* eds. L. Duranti, E. Shaffer, Vancouver 2012, pp. 298–313, https://ils.unc.edu/callee/p298-lee.pdf [access: 5.11.2024].

Lemieux V.L., Werner J., *Protecting Privacy in Digital Records: The Potential of Privacy-Enhancing Technologies*, "Journal on Computing and Cultural Heritage" 2024, vol. 16, no. 4, article 83, pp. 1–18, https://doi.org/10.1145/3633477 [access: 5.11.2024].

Liu B., Ding M., Shaham S., Rahayu W., Farokhi F., Lin Z., *When Machine Learning Meets Privacy: A Survey and Outlook,* "ACM Computing Survey" 2021, vol. 54, no. 2, article 31, pp. 1–36, https://doi.org/10.1145/3436755 [access: 5.11.2024].

Mordell D., *Critical Questions for Archives as (Big) Data*, "Archivaria" 2019, vol. 87, pp. 140–161.

Nissenbaum H.F., *Privacy in context: Technology, policy, and the integrity of social life*, Stanford 2009.

Ohm P., *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, "UCLA Law Review" 2010, vol. 57, pp. 1701–1777.

Oksanen A., Tamper M., Tuominen J., Hietanen A., Hyvöonen E., *ANOPPI: A pseudonymization service for Finnish court documents* [in:] *Legal Knowledge and Information Systems*, eds. M. Araszkiewicz, V. Rodríguez-Doncel, Amsterdam 2019, pp. 251–254, https://helda.helsinki.fi/server/api/core/bitstreams/622773b4-8c6e-4558-8571-da432fe7ea8f/content [access: 5.11.2024].

Pasquale F., *New laws of robotics: defending human expertise in the age of AI*, Cambridge, Massachusetts 2020.

Rolan G., Humphries G., Jeffrey L., Samaras E., Antsoupova T., Stuart K., *More human than human? Artificial intelligence in the archive*, "Archives and Manuscripts" 2019, vol. 47, no. 2, pp. 179–203, https://doi.org/10.1080/01576895.2018.1502088 [access: 5.11.2024].

Sillitoe P., *Privacy in a public place: Managing public access to personal information controlled by archives services,* "Journal of the Society of Archivists" 1998, vol. 19, no. 1, pp. 5–15.

Silva P., Goncalves C., Godinho C., Antunes N., Curado M., *Using NLP and Machine Learning to Detect Data Privacy Violations* [in:] *IEEE Conference on Computer Communications*

Workshops (INFOCOM WKSHPS), Toronto 2020, pp. 972–977, https://doi.org/10.1109/INFOCOMWKSHPS50562.2020.9162683 [access: 5.11.2024].

Snyder H., *Literature review as a research methodology: An overview and guidelines*, "Journal of Business Research" 2019, vol. 104, pp. 333– 339, https://doi.org/10.1016/j.jbusres.2019.07.039 [access: 5.11.2024].

Solove D.J., *A Taxonomy of Privacy*, "University of Pennsylvania Law Review" 2006, vol. 145, no. 3, pp. 477–564, https://doi.org/10.2307/40041279 [access: 5.11.2024].

Solove D.J., *Access and Aggregation: Public Records, Privacy, and the Constitution*, "Minnesota Law Review" 2002, vol. 86, no. 6, pp. 1137–1209.

Tamper M., Oksanen A., Tuominen J.A., Hyvönen E.A., Hietanen A., *Anonymization Service for Finnish Case Law: Opening Data without Sacrificing Data Protection and Privacy of Citizens*, 2018, https://research.aalto.fi/en/publications/anonymization-service-for-finnish-case-law-opening-data-without-s [access: 5.11.2024].

Todd M., *Power, Identity, Integrity, Authenticity, and the Archives: A Comparative Study of the Application of Archival Methodologies to Contemporary Privacy*, "Archivaria" 2006, vol. 61, pp. 181–214.

Tzouganatou A., *Openness and privacy in born-digital archives: reflecting the role of AI development*, "AI & Society" 2022, vol. 37, pp. 991–999, https://doi.org/10.1007/s00146-021-01361-3 [access: 5.11.2024].

Westin A.F., *Privacy and Freedom*, New York 1967.

Yun H., Lee G., Kim D.J., *A Chronological Review of Empirical Research on Personal Information Privacy Concerns: An Analysis of Contexts and Research Constructs*, "Information & Management" 2019, vol. 56, no. 4, pp. 570–601, https://doi.org/10.1016/j.im.2018.10.001 [access: 5.11.2024].

Netography

Heaven W.D., *What is AI?*, "MIT Technology Review", 10 July 2024, https://www.technologyreview.com/2024/07/10/1094475/what-is-artificial-intelligence-ai-definitive-guide/ [access: 5.11.2024].

IAPP. Glossary, https://iapp.org/resources/glossary/#paperwork-reduction-act-2 [access: 5.11.2024].

Kundu R., *F1 Score in Machine Learning: Intro & Calculation*, 16 December 2022, https://www.v7labs.com/blog/f1-score-guide [access: 5.11.2024].

Ovide S., *Why Google's AI might recommend you mix glue into your pizza*, "The Washington Post", 24 May 2024, https://www.washingtonpost.com/technology/2024/05/24/google-ai-overviews-wrong/ [access: 5.11.2024].

UNCTAD. Data Protection and Privacy Legislation Worldwide, https://unctad.org/page/data-protection-and-privacy-legislation-worldwide [access: 5.11.2024].

Ware W.H., *Records, computers and the rights of citizens*, ["Report of the Secretary's Advisory Committee on Automated Personal Data Systems", Washington 1973], https://aspe.hhs.gov/reports/records-computers-rights-citizens [access: 5.11.2024].

WIPO. Genetic Resources, Traditional Knowledge and Traditional Cultural Expressions, https://www.wipo.int/tk/en/ [access: 5.11.2024].