I Trust AI, el Nuevo Proyecto de investigación de InterPARES

Luciana Duranti[*]
Muhammad Abdul-Mageed[**]
Darra Hofman[***]
Peter Sullivan[****]

Resumen

El nuevo proyecto InterPARES, I Trust AI, tiene por obje-
tivo diseñar, desarrollar y aprovechar la IA para respaldar la dis-
ponibilidad y accesibilidad continuas de documentos públicos
confiables mediante la formación de una colaboración sostenible
y continua que produzca investigaciones originales, capacite a
estudiantes y otro personal altamente calificado, generando un
círculo virtuoso entre la academia, las instituciones archivísticas,
los profesionales de documentos gubernamentales y la industria.
Con alrededor de 200 participantes y 87 socios, el enfoque es

[*] Principal Investigator and Co-Director of I Trust AI project. School of
Information of the University of British Columbia (Vancouver, Canada).
Luciana.Duranti@ubc.ca

[**] Co-Director of I Trust AI project. School of Information of the University of
British Columbia (Vancouver, Canada). muhammad.mageed@ubc.ca

[***] School of Information of the San Jose State University (San Jose, USA). darra.
hofman@sjsu.edu

[****] School of Information of the University of British Columbia (Vancouver,
Canada).prsull@student.ubc.ca

completamente interdisciplinario para respaldar el examen integral de las dimensiones administrativas, de archivo, tecnológicas, éticas, legales y sociales de la implementación de IA en el control y acceso a documentos públicos confiables.

El desafío al que se enfrenta este proyecto nunca antes había sido abordado de manera sistemática y global. No obstante, y si bien se desconocen los riesgos de usar la IA para resolver los problemas de administrar el incesante crecimiento de los documentos públicos a lo largo de su ciclo de vida, los riesgos de no actuar de manera concertada para hacerlo son inaceptables: pérdida de la capacidad de garantizar los derechos de las personas; de la evidencia como base para la toma de decisiones; y de la memoria histórica. Esa es la razón por la cual el proyecto es extremadamente importante para las agencias gubernamentales y los archivos, para los especialistas en IA y para las universidades que educan a los profesionales de archivos y a los expertos en IA del futuro.

Palabras clave: Inteligencia Artificial y Archivos, Machine Learning y Archivos, Tecnologías y Archivos

I Trust AI, the latest InterPARES research project

Abstract
The new InterPARES project, *I Trust AI*, is addressed to design, develop, and leverage AI to support the ongoing availability and accessibility of trustworthy public records by forming a sustainable, ongoing Partnership producing original research, training students and other highly qualified personnel (HQP), generating a virtuous circle between academia, archival institutions, government records professionals, and industry. With

about 200 participants and 87 partners, the approach is fully interdisciplinary in order to support the comprehensive examination of the administrative, archival, technological, ethical, legal, and social dimensions of implementing AI to control and provide access to trustworthy public records.

The challenge facing with this project has never before been systematically and globally dealt with. However, while the risks of using AI to solve the problems of managing the ever-growing of public records throughout their lifecycle are unknown, the risks of not acting in concert to do so are unacceptable: loss of the ability to secure people's rights; of evidence as a foundation for decision making; and of historical memory. That is the reason why the project is extremely significant to government agencies and archives, to AI specialists, and to universities educating the records and archival professionals and the AI experts of the future.

I Trust AI, the latest InterPARES research project

Supporting archival work by using Artificial Intelligence (AI) is an old idea. Rhoads wrote in 1969 about the "cybernetic archivist:" "we would not use the computer simply as a filing clerk and typist but rather as a cybernetic extension of the resear-

cher himself."[1] Hirtle discussed "Artificial Intelligence, Expert Systems, and Archival Automation" in 1987.[2] Stielow compared the development of archival theory with the cutting edge of AI in 1991, characterising both as "the search for the obvious."[3] Benson examined the use of AI to represent archival photographs in 2009.[4] Yet, in 2019, only 44% of records professionals agreed that their organizations use "automated tools […] to locate and preserve relevant information," 49% manually delete emails, and 58% manually delete records from mobile devices.[5] Organizations are drowning in records and information. Smallwood found that "71 percent of organizations surveyed have no idea of the content in their stored data".[6] In short, organizations are facing unprecedented uncertainty, struggling with the unsustainable task of keeping control of the explosive number of digital records now present.[7]

[1] Rhoads, J. B. (1969). The Historian and the New Technology. *The American Archivist*, *32*(3), 209–213.

[2] Hirtle, P. B. (1987). Artifical Intelligence, Expert Systems, and Archival Automation. *Provenance, Journal of the Society of Georgia Archivists*, *5*(1), 8.

[3] Stielow, F. J. (1991). Archival Theory Redux and Redeemed: Definition and Context Toward a General Theory. *American Archivist*, *54*, 14–26.

[4] Benson, A. C. (2009). The Archival Photograph and Its Meaning: Formalisms for Modeling Images. *Journal of Archival Organization*, *7*(4), 148–187.

[5] Cohasset Associates, & ARMA International. (n.d.). *2019 Information Governance Benchmarking Report*. Retrieved from https://armai.informz.net/ARMAI/pages/Cohasset_Benchmarking_Survey_2019.

[6] Smallwood, R. (2015). *Information Governance: Concepts Strategies and Best Practices.* Wiley

[7] Hirtle, P. B. (1987). Artificial Intelligence, Expert Systems, and Archival Automation. *Provenance, Journal of the Society of Georgia Archivists*, *5*(1), 8; Stielow, F. J. (1991). Archival Theory Redux and Redeemed: Definition and Context Toward a General Theory. *American Archivist*, *54*, 14–26; Benson, A. C. (2009). The Archival Photograph and Its Meaning: Formalisms for Modeling Images. *Journal of Archival Organization*, *7*(4), 148–187; Smallwood, Robert. (2015). *Information Governance: Concepts Strategies and Best Practices.* Wiley.

The situation is no better when digital records are ready to cross the archival threshold. Email records, for example, can be critical to accountability, enforcing rights, and preserving historical memory. "Yet collecting email is difficult and poses a host of logistical and ethical challenges. Archives have struggled to develop reproducible strategies and methods to both screen email for sensitive, confidential or legally restricted information, and provide effective access to email, due to the sheer volume and complexity of the material".[8] In the same vein, metadata – the data about data which attest to the identity and integrity of a record – is beyond anyone's ability to manage manually; "[t]o achieve the volumes of metadata that may be needed, archivists and records managers will almost certainly have to [use] artificial intelligence systems that automatically analyse the form, content and context of objects and populate descriptive systems accordingly".[9]

Since Anne Gilliland first explored using an expert system to classify digital records in the early 90s,[10] there have been several projects looking at AI in archives; however, such projects are typically discrete, often looking at a particular tool in a specific context or even a single set of records. While these studies are limited in scope, only evaluating one or two AI techniques, they generally agree upon the potential for AI tools, such as recu-

---

[8] Schneider, J., Adams, C., DeBauche, S., Echols, R., McKean, C., Moran, J., & Waugh, D. (2019). Appraising, processing, and providing access to email in contemporary literary archives. *Archives and Manuscripts*, *47*(3), 305–326.

[9] Yeo, G. (2013). Trust and context in cyberspace. *Archives and Records*, *34*(2), 214–234.

[10] Gilliland, A.J., (2016). Designing Expert Systems for Archival Evaluation and Processing of Computer Mediated Communications: Frameworks and Methods. Scholarship.org.

rrent neural networks, for classification of the content of large collections of records, to aid records managers and archivists in their daily work.[11] In addition to easing the burden on these professionals' current responsibilities, some studies have considered the potential use of AI to expand access to archival information, for instance through recommendation systems that connect relevant images to digitized letters[12] by using handwritten text recognition (HTR) to make old documents searchable,[13] and even chatbots to help knowledge seekers find connected information.[14]

---

[11] Belovari, S. (2017). Expedited digital appraisal for regular archivists: an MPLP-type approach. *Journal of Archival Organization*, 14(1-2), 55-77; Lee, C. A. (2018). Computer-Assisted Appraisal and Selection of Archival Materials. *2018 IEEE International Conference on Big Data,* 2721–2724; Makhlouf Shabou, B., Tièche, J., Knafou, J., & Gaudinat, A. (2020). Algorithmic methods to explore the automation of the appraisal of structured and unstructured digital data. Records Management Journal, 30(2), 175–200; Rolan, G., Humphries, G., Jeffrey, L., Samaras, E., Antsoupova, T., & Stuart, K. (2019). More human than human? Artificial intelligence in the archive. *Archives and Manuscripts*, *47*(2), 179–203; van Hooland, S., & Coeckelbergs, M. (2018). Unsupervised Machine Learning for Archival Collections: Possibilities and Limits of Topic Modeling and Word Embedding. *Revista catalana d'arxivística*, *41*, 73; Vellino, A., & Alberts, I. (2016). Assisting the appraisal of e-mail records with automatic classification. *Records Management Journal 26*(3): 293 – 313.

[12] Vowell, Z., Kusters, E., Soares, L., Verkruyse, S. Wilson, J.; Khosmood, F. (2020). Morgan Papers: Exploring the Correspondence of California's First Female Architect. https://dblp.org/rec/conf/dihu/VowellKSVWK20.html.

[13] Goudarouli, E., Sexton, A., & Sheridan, J. (2019). The Challenge of the Digital and the Future Archive: Through the Lens of The National Archives UK. *Philosophy & Technology*, *32*(1), 173–183; Dunley, R. (2018). Machines reading the archive: handwritten text recognition software. Blogpost.https://blog.nationalarchives.gov.uk/machines-reading-the-archivehandwritten-text-recognition-software/.

[14] Gupta, A., & Kapoor, N. (2020). Comprehensiveness of Archives: A Modern AI-enabled Approach to Build Comprehensive Shared Cultural Heritage. *arXiv preprint arXiv:2008.04541*.

Despite all the above, few studies have produced tools enabling broad AI methods to be adopted by the records and archival community at large, showing a gap that could be filled by a more comprehensive approach aimed at putting these tools into practice. One of the few recent projects that did produce a tangible AI tool for records managers, BitCurator-NLP (Natural Language Processing, the broad term for AI dealing with human language), used a combination of Named Entity Recognition (NER), entity relations tools, and an unsupervised learning technique, topic modeling, from off the shelf machine learning packages to create a visualization tool for the types of data stored on disk images.[15] These NLP techniques are essential to any AI solution to archival problems, due to how ubiquitous language is in archival data; yet, the field of NLP has undergone a major revolution within the last few years, leaving traditional (non-deep learning) methods outclassed by the introduction of extremely powerful deep learning models, such as Transformer and its variants.[16] To our knowledge, no application of these newer deep learning models has occurred in an archival setting, leaving open the questions of how we can leverage them to meet archival needs.

---

[15] Lee, C. A. (2018). Computer-Assisted Appraisal and Selection of Archival Materials. *2018 IEEE International Conference on Big Data,* 2721–2724. https://ieeexplore.ieee.org/servlet/opac?punumber=8610059

[16] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. In *Advances in neural information processing systems* (pp. 5998-6008); Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805; Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P. & Agarwal, S. (2020). Language models are few-shot learners. arXiv preprint arXiv:2005.14165.

However, relying on existing off the shelf tools, as all the past studies on AI in archives have done, limits what challenges can be met, as it makes the needs of archives subservient to the larger field of machine learning. In some cases, this may be practical, but, with many tangible instances of bias found in modern machine learning models, often driven by laissez faire data collection practices,[17] it raises the questions of whether it is the best solution for the archival field and of what AI could look like if this power relationship between AI and archives were reversed, with archival theory informing the creation of AI tools themselves. What is lacking is comprehensive, systematic research into the use of AI to carry out the different archival functions in an integrated way and ensure the continuing availability of verifiably trustworthy records to prevent the erosion of accountability, evidence, history and cultural heritage. Thus, we should address the technological issues from the perspective of archival theory, by integrating the technology with complex human-oriented tools.[18]

In April 2021, the Social Sciences and Humanities Research Council of Canada (SSHRC) awarded a Partnership grant to the InterPARES team for a new project, *I Trust AI* (where I stands for InterPARES). The overall goal of this project is to

---

[17] Prabhu, V. U., & Birhane, A. (2020). Large image datasets: A pyrrhic win for computer vision?. arXiv preprint arXiv:2006.16923; Sheng, E., Chang, K. W., Natarajan, P., & Peng, N. (2019). The woman worked as a babysitter: On biases in language generation. arXiv preprint arXiv:1909.01326; Tan, Y. C., & Celis, L. E. (2019). Assessing social and intersectional biases in contextualized word representations. In Advances in Neural Information Processing Systems (pp. 13230-13241).

[18] Rolan, G., Humphries, G., Jeffrey, L., Samaras, E., Antsoupova, T., & Stuart, K. (2019). More human than human? Artificial intelligence in the archive. *Archives and Manuscripts*, *47*(2), 179–203.

design, develop, and leverage AI to support the ongoing availability and accessibility of trustworthy public records by forming a sustainable, ongoing Partnership producing original research, training students and other highly qualified personnel (HQP), and generating a virtuous circle between academia, archival institutions, government records professionals, and industry, a feedback loop reinforcing the knowledge and capabilities of each party.

The objectives of *I Trust AI* are to:
1. Identify specific AI technologies that can address critical records and archives challenges;
2. Determine the benefits and risks of using AI technologies on records and archives;
3. Ensure that archival concepts and principles inform the development of responsible AI; and
4. Validate outcomes from Objective 3 through case studies and demonstrations.

Our approach is two-pronged, comprising: (1) the practical and immediate need to address large-scale existing problems, and (2) the longer-term need to have AI-based tools that are reliably applicable to future problems.
1. Our short-term approach focuses on identifying high impact problems and limitations in records and archives functions, and applying AI to improve the situation. This will be achieved via collaboration between records and archival scientists and professionals and AI researchers and industry experts.
2. Our long-term approach focuses on identifying the tools that records and archives specialists will need in the future to flexibly address their ever-changing needs. This includes decision support and, once decisions are made, rapid implementation of AI-based solutions to those needs.

Overarching both prongs will be a focus on the overall impact of the use of AI in the long-term context of public records, on a global basis. The question of how to support decision-making at the global, national, and local level involves a larger challenge presented by AI, which, despite its pervasive use, remains fraught. Osoba and Welser call "artificial agents whose results lead to incorrect, inequitable, or dangerous consequences" "misbehaving algorithms".[19] When algorithms misbehave, the results can be catastrophic, depriving people of due process[20] or even liberty,[21] "mediat[ing] access to opportunity"[22] in a way that retrenches and legitimizes bias[23] and circumvents laws.[24] Given the essential role of public records in governance, accountability, securing rights, and understanding the past, it is critical to identify the broader ethical, legal, and social implications of any solution implemented to create, manage, use, intellectually control, preserve, and provide access to such records.

Because of the critical importance of public records, we must find ways to capitalize on the efficiencies AI promises while avoiding unintended consequences, including risks to the very

[19] Osoba, O., Welser, W., IV, & RAND Reports. (2017). *An intelligence in our image: The risks of bias and errors in artificial intelligence*, 8.

[20] Citron, D. K. (2007). Technological due process. Wash. UL Rev., 85, 1249.

[21] Flores, A.W. (2016). False positives, false negatives, and false analyses: A rejoinder to machine bias: There's software used across the country to predict future criminals. and it's biased against blacks. *Federal probation*. , 2016, Vol.80, p.38; Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias. *ProPublica*, May, 23, 2016.

[22] Osoba, cit. 11.

[23] Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *California Law Review*, *104*(3), 671.

[24] Barocas, S., & Nissenbaum, H. (2014). Big data's end run around procedural privacy protections. *Communications of the ACM*, *57*(11), 31–33.

integrity of the public records infrastructure.[25] If this is done well, AI-enabled archives could support governance, evidence, and culture. Indeed, leveraging AI for archives is necessary for nothing less than the pursuit of comprehensiveness, equity, and justice.[26]

This project's interdisciplinary approach is based on cooperation between scholars and professionals in records and archives management, diplomatics, AI, law, communication and media, cybersecurity, information governance and assurance, digital forensics, computer engineering, robotics, and ethics, in order to support the comprehensive examination of the administrative, archival, technological, ethical, legal, and social dimensions of implementing AI to control and provide access to trustworthy public records.

In consideration of the fact that the project aims to discover solutions that support the records and archives professions in meeting society's needs for trustworthy records, the theoretical framework will be adapted from archival and diplomatic theory. "The first object of archival theory is the nature of archival documents or records. The archival discipline consists in building knowledge about archival documents and acting upon them in methodical ways to protect the properties that they have".[27] Foundational constructs from those bodies of theory, including "record" (document, information, data) and "trustworthiness" (reliability, authenticity—identity and integrity, and accuracy),

---

[25] Cheatham, B., Javanmardian, K., & Samandari, H. (2019). Confronting the risks of artificial intelligence. *McKinsey Quarterly*, 1-9.

[26] Gupta, A., & Kapoor, cit.

[27] Eastwood, T. (1994). What is archival theory and why is it important? *Archivaria*, 37, 122.

and concepts such as the characteristics of archival documents/records (impartiality, authenticity, naturalness, interrelatedness, and uniqueness; and the network of originary, necessary, and determined relationships between and among records (archival bond), will inform this work.

One of the objectives of this project is to understand the risks and benefits of AI adoption in recordkeeping and archival preservation. Thus, we will also use risk management, an area of study that informs decision making in an uncertain environment. Management of records is essentially about managing risk in the context of legislation, regulation, and reputation while weighing up costs and benefits.[28] Operationally, risk managers reference the *ISO 31000 Risk Management – Principles and Guidelines on Implementation* framework, and, like record managers, seek to build best practices into business processes in order to support "good governance and accountability processes in organizations".[29] Introducing AI in records and archives environments, especially if imbued with decision making authority, introduces risks, costs, and benefits that must be understood operationally.

This research project, which prioritizes a holistic understanding of the potential impacts of AI on critical records and archival challenges, is grounded in soft systems thinking methodology. **"**Systems thinking can be characterised as an attempt to find common principles that apply at different levels of scale and across different types of phenomena. This methodology "makes

[28] Moss, M., & Gollins, T. (2017). Our digital legacy: an archival perspective. *The Journal of Contemporary Archival Studies*, 4.
[29] Isa, A. M., & Nordin, N. M. (2012). Strategic records and risk management for the sustainability of organisations. In *ICA Congress* in Brisbane, Australia, 20-23.

possible the collection and organization of accumulated knowledge in order to increase the efficiency of our actions'".[30] Soft systems thinking deals with "abstract or conceptual systems constituted by intangibles such as plans, hypotheses and ideas, wherein symbols represent attributes and objects".[31] In the process of "sweeping in," this project will proceed through five phases. Each phase builds upon the knowledge gained in the previous phase. The table below shows the broad activities of each phase as well as methods that will be employed to support that activity.

Date Range	Research Objective	Activities
2021 - 2022	1: Identify specific AI technologies that can address critical records and archives challenges	• Identify critical challenges to be addressed by AI, adding to an initial survey of techniques o Surveys and interviews with practitioners within the global records and archives community Identify within each critical challenge the specific factors to be addressed and how AI might address them o Expert interviews and mapping • Identify and prototype candidate AI technologies o Candidate use cases • Create initial evaluation criteria for AI solutions for records and archival challenges, including a diverse set of challenge datasets focusing on specific issues

[30] Tippett, J., Handley, J. F., & Ravetz, J. (2007). Meeting the challenges of sustainable development—A conceptual appraisal of a new methodology for participatory ecological planning. *Progress in Planning*, *67*(1), 57.
[31] Georgiou, I. (2007). *Thinking through systems thinking*. Routledge, 29.

2022 - 2023	2: Determine the risks and benefits of using AI technologies on records and archives	• Determine the requirements of public records compared to the capabilities of AI technologies o Doctrinal legal research o Development of a value structure for risks and benefits • Identify the limitations of each potential AI solution o Policy analysis o Expert interviews o Environmental Scans o Comparison studies of AI solutions on representative datasets • Develop list of threats and vulnerabilities o SWOT/PESTLE Analysis o Theoretical Analysis o Stakeholder Interviews o Expert Assessment o Error analysis of AI solutions based on performance on challenge datasets. • Iterate on validation criteria, for instance creating new versions of challenge datasets, to address any important factors discovered through threat and vulnerability analysis

2023 - 2024	3: Establish how archival concepts and principles can inform the development of responsible AI	• Establish archival principles to be used to inform AI development
		• Develop and improve AI tools based on these principles
		• Identify and mitigate biases present in training datasets and models
		   o Consistency Analysis
		• Determine whether AI informed by archival principles is more aligned with archival needs
		   o Experimental comparison of models on challenge datasets

2024 - 2025	4: Validate outcomes from Objective 3 through case studies and demonstrations.	• Deploy archival oriented AI tools
		• Measure AI solutions against the validation criteria developed in Phases 1 and 2
		• Examine feasibility, sustainability, bias, transparency, generalizability, and preservation of context in AI solutions
		   o Case studies
		   o Use cases
		   o Detailed error analysis of AI solutions in the context of case studies
		• Develop and validate tools including framework for evaluation and checklists for institutions considering AI implementation

2025 - 2026	5: Completion of Outputs	• Finalize overarching publication of outcomes • Packaged software (e.g. to automatically caption historical photos, sensitize descriptions of documents, or translate historical documents in indigenous languages.)

Throughout the research, we will also employ AI methods, including NLP and a diversity of machine learning methods. NLP models can be applied to a number of challenges attendant to textual records, including text classification, translation, dialect detection, and even identifying offensive or harmful language.[32] Machine learning methods, including both traditional and deep learning techniques, will also be utilized to resolve archival problems that have just begun to be explored. For example, improving OCR for better accessibility and usability of historical records raises issues ranging from statistical analysis of OCR output to linguistic diversity in OCR training datasets.[33]

The project will also employ CNN/ResNet architectures for image classification and a range of similar techniques in conjunction. While exploratory research on archival data has compared some traditional machine learning techniques (bag of words) to deep learning techniques (like convolutional neural nets),[34] it has not considered recent deep learning techniques showing promise in an archival context. For example, an extremely useful tool would be the automatic caption or description of the countless images that exist in archives. For images that

[32] Elmadany, A., Zhang, Chiyu, Abdul-Mageed, M. & Hashemi, A. (2020). Leveraging Affective Bidirectional Transformers for Offensive Language Detection. In *Proceedings of the Fourth Workshop on Open-Source Arabic Corpora and Processing Tools* (OSACT4); Zhang, C., & Abdul-Mageed, M. (2019, August). No army, no navy: Bert semi-supervised learning of Arabic dialects. In *Proceedings of the Fourth Arabic Natural Language Processing Workshop* (pp. 279-284).

[33] Smith, D. A., & Cordell, R. (2018). A research agenda for historical and multilingual optical character recognition. NUlab, Northeastern University. https://ocr.northeastern.edu/report. Accessed 10/20/2020.

[34] Patel, K., Caragea, C., Phillips, M. E., & Fox, N. T. (2020, August). Identifying Documents In-Scope of a Collection from Web Archives. In *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries* in 2020, 167-176.

have no descriptions, advances in automatic image captioning through attention-based Faster R-CNN models[35] could be utilized. Similarly, where a description exists in a different language, multitask learning can be used to improve the quality of the translation.[36] Improving tools that work with this type of archival object alone could rely on a number of different image processing techniques, such as object detection and image classification,[37] as well as NLP techniques, such as text generation,[38] text classification,[39] and machine translation.[40]

The challenge we face with this project has never before been systematically and globally dealt with; it is enormous and fraught, but critical. While the risks of using AI to solve the problems of managing the ever-growing, ever-more-diverse bo-

[35] Anderson, P., He, X., Buehler, C., Teney, D., Johnson, M., Gould, S., & Zhang, L. (2018). Bottom-up and top-down attention for image captioning and visual question answering. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 6077-6086.

[36] Elliott, D., & Kádár, A. (2017). Imagination improves multimodal translation. arXiv preprint arXiv:1705.04350.

[37] Cireşan, D. C., Meier, U., & Schmidhuber, J. (2012, June). Transfer learning for Latin and Chinese characters with deep neural networks. In *The Proceedings of the 2012 International Joint Conference on Neural Networks* (IJCNN), 1-6. IEEE.

[38] Fedus, W., Goodfellow, I., & Dai, A. M. (2018). MaskGAN: Better text generation via filling in the_. arXiv preprint arXiv:1801.07736.

[39] Abdul-Mageed, M., & Ungar, L. (2017, July). Emonet: Fine-grained emotion detection with gated recurrent neural networks. In *Proceedings of the 55th annual meeting of the association for computational linguistics* (volume 1: 718-728).

[40] Nagoudi, E. & Abdul-Mageed, M. (2020). Growing Together: Modeling Human Language Learning With n-Best Multi-Checkpoint Machine Translation. In *Proceedings of The 4th Workshop on Neural Generation and Translation (WNGT 2020);* Przystupa, M., & Abdul-Mageed, M. (2019, August). Neural machine translation of low-resource and similar languages with backtranslation. *In Proceedings of the Fourth Conference on Machine Translation* (Volume 3: 224-235).

dies of public records throughout their lifecycle, from creation to preservation and access, are unknown, the risks of not acting in concert to do so are unacceptable: loss of the ability to secure people's rights; of evidence of past acts and facts to serve as a foundation for decision making; and of historical memory.

This project will bring significant impact to society in several areas. (1) Records-keeping in local and national government agencies is a vital part of our society's ability to maintain oversight on and accountability of governance, but with the inability to handle the vast quantities of digital records, public bodies risk undermining their own legitimacy as oversight entities if they can not appropriately process and make accessible information in a timely fashion. By helping address this crisis through the development, evaluation, and contextualization of AI techniques we can contribute to the ability of agencies and institutions to maintain their place in our democracy. (2) Automation techniques can potentially aid in the economic viability of many cash-starved records offices and archival institutions by ensuring that professional records management and archival expertise is used wisely, with classification tools and TAR able to allow a quick review and assessment of vast quantities of records. Similarly, with businesses depending on records agencies for routine activities, improved speed in responding to queries will bring a positive effect to the economy. (3) AI techniques have the potential to aid in the accessibility of records and archives material by new audiences, for instance by translating and indexing historical materials written in indigenous languages, sensitising problematic archival descriptions, helping patrons find connected items, or captioning historical photographs. These techniques have both a cultural significance, by providing better

access to historical material, and a social and scientific significance, by making current records easier to organise, retrieve and use by both their creators and the public at large. (4) While there have been numerous calls to action to systematically explore the application of AI techniques to the records and archives field, AI also currently faces major ethical challenges that will benefit from an archival theory perspective, for instance in dealing with bias and personal information. By exploring further, the connections between AI and archives, this project will contribute to the intellectual progress of both fields.

The project has generated a great amount of enthusiasm among participants (about 200) and partners (87), as well as organizations that do not have the capacity to participate but look forward to outcomes they can use, because it deals with issues that are already dramatically changing the way we act, behave and think. The knowledge we have developed in the previous four InterPARES projects[41] is a good place to start as we move from a focus on designers and builders of records and archives systems to a focus on developing efficiency and effectiveness in carrying out archival functions while protecting professional, social and cultural values. The objectives of the project are extremely significant to government agencies and archives to simply fulfill their mandates, to AI specialists to develop tools that meet social and ethical requirements, and to universities educating the records and archival professionals and the AI experts of the future.

We have a unique and essential contribution to make, because we have the means of creating knowledge ensuring that digital data and records are controlled and made accessible in

---

[41] See www.interpares.org.

a trustworthy, authentic form wherever they are located; are promptly available when needed; duly destroyed when required; and accessed only by those who have a right to do so. To learn more about the I Trust AI project, please go to its website at www.interparestrustai.org.

---

[*] Director del Servei d'Arxiu Municipal de Lloret de Mar -SAMLM. E-mail: jdaban@lloret.cat