

Preserving the Past, Generating the Future: How archivists can contribute to generative AI's development

Sebastian Rodriguez & Chloe Thierstein



Provenance and Transparency

- A key concern of generative AI is the question of data ownership: how are models trained, where does the data come from, and do companies possess the right to use such data? **Complex models like diffusion and large language models make tracing data ownership difficult**, impacting data integrity and public support.
- Archival concept of "provenance" can be adapted for generative AI, focusing on the history and context of data. **Provenance helps in attributing content to original authors** and in internal organization by linking data points contextually.
- Generative AI operates through prediction, associating input text with likely output. Establishing **provenance for data points can lead to more meaningful AI responses** by leveraging data context.
- Provenance is already used in computer science and information management to track data history and context (Yeo, 2013, p. 220). **Enhances judgment of data trustworthiness by maintaining and organizing data based on its source or creator.**
- Immediate value of source tracing in tasks like search queries (e.g., Bing using ChatGPT and citing URLs). Source tracing benefits all generative AI users, reducing "hallucinations" and improving model performance (Ji et al., 2022). **Provenance enhances reliability by ensuring data creation and use context is preserved** (Duranti, 1995, p. 6).
- Legal challenges, such as OpenAI facing lawsuits for copyright infringement, highlight the need for proper data usage rights (Reisner, 2023; Montgomery, 2023). **Provenance allows users to see data origins and helps copyright holders identify and manage their works.**

Data Governance and Accountability

- Organizations are increasingly implementing data governance plans with the rise of AI. Data governance is a knowledge management concept with various implementations depending on institutional focus. Focus areas include data availability, usability, quality, integrity, and security.
- Archivists have extensive experience in knowledge management and standards. Key contributions include **supporting data quality, integrity, and usability through appraisal practices**. Experience in developing institutional mandates aligns with supporting an organization's overarching mission.
- Archivists excel in **curation through appraisal, determining information with "enduring value"** (Craig, 1992, p. 176). They recognize the wide-reaching social implications of including or excluding records in archives and have expertise in identifying human bias in decision-making processes.
- **AI is biased**, as seen in historical and recent implementations (e.g., biased sentencing algorithms, biased model training). Companies like OpenAI address biases through trust and safety processes but face ethical issues (e.g., exploitation of underpaid labor) (Perrigo, 2023). Inclusion of archivists in data governance discussions can reduce biases and create more trusted AI systems.

Long-Term Preservation

- Preservation of AI methods, datasets, and software is crucial due to the transformative impact of AI on society. **Future generations must understand AI technology, its functionality, and its societal impact.**
- AI, as a digital tool, will be difficult for future generations to use and comprehend due to evolving technology and unsupported systems.
- **Collaboration between technologists and archivists has shown success** (e.g., GitHub's Arctic Code Vault). Launched in 2020, storing 21 terabytes of data on 186 reels of film in Svalbard, Norway ("Arctic Code Vault," 2020). Demonstrates the value of recording digital history and the contribution of archival studies to computer science.
- Most popular **AI models are proprietary, not open-source, posing a challenge for public archiving**. Companies like OpenAI, Google, and Meta need to work with archivists to preserve their technology's interfaces and underlying frameworks.
- Archiving front-end (user-facing layer) differs significantly from archiving back-end (logic and data management layer).
- Internet Archive's Wayback Machine captures static content but cannot preserve dynamic functionality requiring interaction with a server (Wayback Machine, n.d.). **Web crawling practices are insufficient for preserving AI**, which relies on dynamic server interactions. Collaboration with AI developers is essential to effectively preserve AI technology.

Conclusion & Future Directions

- Archival studies and information management offer significant value to AI development. **Provenance** helps identify information sources and maintain system reliability. Archival appraisal practices can improve AI's societal impact.
- **Archivists have extensive experience** in data governance. Incorporating archivists in AI data governance discussions can be a starting point.
- Archivists and technologists have collaborated to preserve software projects in the past. Ensures future generations understand the use of digital technology. **Preservation is difficult due to AI's closed-source, proprietary nature**. Recording AI technology is crucial for cultural memory.
- Including archivists may be challenging due to the tech industry's focus on computer scientists, mathematicians, and designers. **Archivists' critical perspectives and methods can enhance AI's societal impact.**
- Future research should examine specific instances of archival practices in AI development. Focus on open-source alternatives to proprietary AI models like OpenAI's ChatGPT and Google's Gemini.