Guidelines for AI and E-Resources Licensing

Document Status: Active
Prepared by: Chris Palazzolo, Jessie Copeland, Susan Kellett, John Morgenstern
Reviewed by: Collections Steering Committee, Cabinet
Approved by: Library Cabinet
Date Approved: 5 September 2024
Revision Cycle: Annual
Current as of 4 September 2024

Purpose and Scope

This document outlines the Emory Libraries' policies and guidelines for responding to AI language in vendor agreements for e-resources. As with text and data mining," AI learning models hold promise for advancing research through the analysis of large corpora of text (as data). This document will be revised as AI models and tools increase in prominence and new legal norms arise that may affect use of licensed library resources. This document does not cover the purchase of AI tools as supplements to existing resources or as stand-alone products, save for the recommendation to carefully review new AI products offered as part of existing licensed vendor content. It further does not cover policies surrounding the use of general university content for AI training content.

Principles

We subscribe to the following principles:

- In line with the [UC Berkeley Comment to the U.S. Copyright Office](#): "scholars' ability to access the underlying content [of licensed resources] to conduct AI training [must] be preserved."
- The use of AI training and machine learning tools using licensed material should be for the purposes of non-commercial research and teaching, provided that the output does not fully recreate copyrighted works or create a competing product and is done in a secure environment.
- AI clauses should not create unenforceable mechanisms for libraries and their authorized users, nor should they establish new liability clauses beyond those already established.
- No ontract clauses pertaining to use of AI or other computational analysis should be more restrictive than fair use allows.
- The use of AI tools or vendor-generated tools within existing resources to access licensed content should not come at a significant additional cost to the Libraries.

Licensing Language Recommendations

- Emory Libraries endorses the [language proposed by UC Berkeley's Scholarly Communications Office](#) for preserving AI in usage restrictions clauses, which builds upon established text and data mining clauses in vendor agreements for electronic resources.
    - These clauses cover the need for secure environments for loading and analyzing data and restricting the use of third-party generative AI tools (not institutionally acquired or subscribed).
- If Emory Libraries is unable to strike or change restrictive AI clauses, the vendor should at least be willing to acknowledge the need to review the clauses/language in light of potential changes to federal law governing the use of copyrighted materials for non-commercial, research and training uses of AI.
- If possible, licenses should include language describing how users can access underlying content for AI training purposes (APIs, etc.). Ideally, users should not be limited to the vendor's own environment. With that said, we recognize that licenses do NOT allow for general automated or systematic bulk downloading of content.
- Any changes to licenses related to TDM or AI should be negotiated at the time of renewal, rather than as an addendum or appendix in the middle of a contract year, save for exceptional circumstances (e.g., researcher requirement).
- We realize that in some cases—especially for vendors who are aggregators themselves—AI language may not be able to be changed in light of contractual obligations with providers that demand those restrictions. This might particularly be the case for statistical data providers. With that said, the Libraries should pose the question of AI use for clarity.

Criteria for Acquisition of Corpora for Data Mining and AI Training

Where legally and financially possible, Emory Libraries will liaise with vendors to accommodate researchers wishing to use corpora derived from licensed e-resources for computational analysis and AI learning. Criteria and guidelines for the supplemental purchase or delivery of licensed content for data mining or various machine-learning techniques are covered under a **separate set of guidelines** . Before using artificial intelligence tools (or access to large amounts of content) with any licensed library resource, researchers should contact the library for more information about governing terms.

Referenced Documents and Endorsements
- ["Licensing Research Content via Agreements that Authorize Uses of Artificial Intelligence."](#)
- [Authors Alliance. January 10, 2024](#)

- "ICOLC Statement on AI in Licensing." International Consortium of Library Consortia. March 3, 2024
- "Research Libraries Guiding Principles for Artificial Intelligence." Association for Research Libraries. May 3, 2024.
- Samberg, Rachel. UC Berkely Library to Copyright Office: Protect Fair Uses in AI Training for Research and Education. October 24, 2023.
- "All TDM & AI Rights Reserved? Fair Use & Evolving Publisher Copyright Statements." SPARC Open Access. March 28, 2024.