Use of Artificial Intelligence for Special Collections & Archives Work

Version 1 - EAR/MT 2025-05-05

This document was created for internal use by and for the UC Irvine Libraries, Department of Special Collections & Archives. It is informed by our values and mission, and compliant with ethics and best practices for archives and librarianship.

Understanding how AI works and its implications

Privacy and security: Use of third-party tools

- Artificial Intelligence (AI) systems, in many cases, are designed without much consideration for security, making them susceptible to cybersecurity attacks, which can lead to your information being used outside of your knowledge [1]
- According to UCI's Information Security Division, "Not only does a third-party (those running the chatbot) have a copy of the information that is submitted, but also the AI large language model may be further trained on that information. Data submitted may be used in future responses to others using the AI chatbot" [2]
- As personal information about individuals is gathered in AI systems, the lack of transparency and control over how that information is used makes the possibility of identity theft and fraud increasingly possible

Ethics and values: Responsible stewards, responsible use

- The use of AI in archival work is an evolving and emerging area, considerations for ethics, biases, and environmental impact are all factors to consider when using AI
- As SCA staff we have a duty to not only protect the records we are entrusted to preserve and provide long term access, but also the people and communities we partner with
- From our SCA values: "We operate from a care-based approach in the ways we work with patrons and collections. We value accountability, commitment to professional

[1] Oseni, A., Moustafa, N., Janicke, H., Liu, P., Tari, Z., & Vasilakos, A. (2021). Security and privacy for artificial intelligence: Opportunities and challenges. Retrieved from https://arxiv.org/pdf/2102.04661
[2] UCI Information Security. Use Artificial Intelligence Chatbots Securely. Retrieved from https://www.security.uci.edu/how-to/ai-chatbot/

standards and ethics, a holistic approach to teaching and learning, stewardship of collections, and quality service to our users"

- AI requires significant electricity and water consumption to run and keep servers cool, which has negative impacts on the power grid and "strain[s] municipal water supplies and disrupt local ecosystems" [3]
    - Shaolei Ren, a professor of engineering at UC Riverside, "estimates that a person who engages in a session of questions and answers with GPT-3...drives the consumption of a half-liter of fresh water" (Berreby, 2024). Ren and colleagues suggest that "globally, the demand for water for AI could be half that of the United Kingdom by 2027" (Crawford, 2024) [4]
- There have been many studies that point out the inherent biases in AI technology, perpetuating harmful racial, gender-based discrimination and stereotypes [5]
- The uses of AI are far reaching and include the criminalization of certain communities by government authorities [6]

UC Protection Levels (P1-P4) and AI best practices

- The UC Protection Levels provide helpful guidelines on data protection and appropriate level of security
    - University of California Protection Levels
- UCI produced guidance on which AI chatbots are recommended for use according to the Protected levels
    - UC Irvine Guidelines for Using AI Chatbots Securely
- Below is a chart of typical records SCA staff encounter in their work and departmental guidance of permissible use of Artificial Intelligence and various options available

[3] Zewe, A. (2025, January 17). Explained: Generative AI's environmental impact. Retrieved from https://news.mit.edu/2025/explained-generative-ai-environmental-impact-0117
[4] UCI Libraries. Environmental impacts of generative AI. Retrieved from https://guides.lib.uci.edu/gen-ai/ethics#s-lg-box-32721905
[5] UN Women. (2025, February 5). How AI reinforces gender bias—and what we can do about it. Retrieved from https://www.unwomen.org/en/news-stories/interview/2025/02/how-ai-reinforces-gender-bias-and-what-we-can-do-about-it
[6] Caputo, M. (2025, March 6). Scoop: State Dept. to use AI to revoke visas of foreign students who appear "pro-Hamas". Retrieved from https://www.axios.com/2025/03/06/state-department-ai-revoke-foreign-student-visas-hamas

Protection Level	Typical records SCA encounters	Ok for AI?	Private/secure campus approved AI tools	Campus approved third-party AI tools
P1 – public	Brochures, website content, exhibit text, published finding aids, items/metadata published on Calisphere	Yes	UCI ZotGPT Chat UCI Microsoft Copilot UCI Google Gemini	OpenAI ChatGPT Microsoft Copilot (Non-UCI) Google Gemini (Non-UCI)
P2 – not public, but can cause minor impact	Meeting notes[7], digitized SCA materials that do not have any P3/P4 data, unpublished finding aids, transfer forms	Yes	UCI ZotGPT Chat UCI Microsoft Copilot UCI Google Gemini	
P3 – not public, but disclosure could cause moderate damage	Student records (FERPA), Attorney-client privilege records, Deeds of Gift and other donor documentation, contractual agreements, Technically not P3 records but use discretion: oral history recordings and transcripts, photographs of people, audiovisual recordings of people "This classification level can also include lower risk items that, when combined, represent increased risk."[8]	No, with some exceptions and consult with the department.	UCI ZotGPT Chat UCI Microsoft Copilot	
P4 – not public, can result in fines, civil or criminal violations	SSNs, PHI, credit card numbers, passwords or encrypted content, any donor imposed restricted content	No		

[7] It depends on the sensitivity of what is being discussed, if there are any FERPA violations (course recordings), etc.
[8] https://www.security.uci.edu/program/classification/protection-levels/

How Staff Are Currently Using AI

- Matching an original item with a scanned copy of it
- Learning and creating scripts for automations in software
- Recording and transcribing and summarizing meeting notes
- Drafting documents, reports
- Summarize articles - find bullet points, important ideas
- Creating a schedule for an intern, progress points, check ins
- Creating artwork/graphics (stickers!)

Best practices for AI use for SCA work

- AI technology is in its infancy—double check the search results as they can often be incorrect
- Regularly evaluate AI outputs and take into consideration any possible errors and biases to ensure both accuracy and reliability of the AI system you choose to use
- When deciding whether to use AI or not for work, think about…
    - Does this material have any restrictions (donor-imposed, copyright)?
    - Who is documented in this record? Do they represent a vulnerable community? What is the potential risk of harm?
        - Examples of vulnerable communities include: undocumented/immigrant communities, visa holders, activists, people who experience(d) incarceration, LGBTQIA+ communities, communities of color. These communities are at greater risk for legal and other harm.
    - Consider environmental impact—does the use of AI help you fix a significant problem?
    - Consider the long-term impact it could have on our roles within Special Collections and Archives (e.g. labor, resource implications)
- Be aware of where data is stored, how it might be used, and how data policies change over time
- Do not share P4 and/or most P3 level data. This may include:
    - Oral histories
    - Recordings of peoples' voices (audio/video)
    - Photographs depicting people
    - Transcripts
    - Donor documentation
    - Vendor documentation
    - Deeds of gift
- Use the campus approved tools first
    - Note: make sure you are signed in with your UCINetID when using Copilot
- Stay up to date with emerging AI technologies, news relating to the AI systems you choose to use, and any security updates or concerns regarding these tools

- Here are some effective ways to use AI while protecting the privacy of our records creators, and users:
    - Automating tasks that do not involve sensitive data
    - Creating bibliographies (verify sources for accuracy)
    - Summarization of content (verify sources for accuracy)
    - Troubleshooting technical issues
    - Generating code
    - Creating images (graphics)

Pros & Cons of Third Party AI Models

Aim to use UCI-endorsed tools. In the use of third party tools, discretion is strongly advised. Below is a graph of popular third-party AI models, their strengths and weaknesses.

AI Tool	Description	Pros	Cons
DeepSeek	A free AI-powered chatbot that has numerous capabilities including generating text, handling large-scale tasks, and analyzing data.	-Great for technical tasks, especially math -Runs more efficiently than other tools -Low cost to train and maintain [9]	-Terms of Use allows it to collect a broad range of information about device, network, and personal activity -Personal data can be stored in international servers
Grok	An AI-powered assistant integrated with X platform that can help answer questions, draft emails, debug code, and generate ideas.	-Integration with X allows for responses based on real-time discussions -Can be useful for analyzing trends and interactions on social networks	-Dependence on data from X can result in informational bias -X users' posts were used for model training without their consent [10]
ChatGPT	Tool that uses natural language processing to generate conversational responses, often used for	Studies showed it to perform better than other LLMs in terms of accuracy and	-Algorithm bias [12] -Violation of the use of personal data [13]

[9] Robison, K. & Lopatto, E. (2025, January 28). Why everyone is freaking out about DeepSeek. Retrieved from https://www.theverge.com/ai-artificial-intelligence/598846/deepseek-big-tech-ai-industry-nvidia-impac
[10] O'Flaherty, K. (2024, September 9). Retrieved from https://www.wired.com/story/grok-ai-privacy-opt-out/
[12]Jegham, N., Abdelatti, M., & Hendawi, A. (2025). Visual Reasoning Evaluation of Grok, Deepseek's Janus, Gemini, Qwen, Mistral, and ChatGPT. Retrieved from https://arxiv.org/pdf/2502.16428
[13] Wach, K. (2023). The dark side of generative artificial intelligence: A critical analysis of controversies and risks of ChatGPT. Retrieved from https://eber.uek.krakow.pl/eber/article/view/2113/852

	its writing, coding, and research capabilities.	reasoning [11] -Can build chatbots that provide human-like responses for various purposes such as customer service and tutoring	
Gemini	Developed by Google Deepmind, this tool is capable of processing, interpreting, and generating text, images, code, audio, and videos.	-Can be integrated with Google apps and services -Complex reasoning capabilities allow it to analyze data and solve problems	-May be challenging to use outside of Google tools and platforms -Data collected is used to develop Google products -Data is retained for up to three years after deleting Gemini activity
Claude	Chatbot trained to generate conversational responses that can be used for summarization, editing, and coding.	-Better for coding support (Artifacts) -Claims to have been developed around Constitutional AI to ensure outputs follow predefined ethical principles and values [14]	-No image generation -No web browsing
Copilot	AI assistant that can be used for writing, image generation, and summarization, among other tasks. It is deeply integrated into Microsoft 365 applications, enhancing AI-powered features into these programs.	-Excels in transcribing and summarizing calls and meetings	-Limitations in processing long documents and large files -Limited functionality outside of Microsoft 365 apps
Canva	AI-powered features integrated into Canva provide tools for image generation and graphic design.	-Can automate tasks traditionally done manually in Photoshop -Can choose image variations based on	-Devalues professional graphic design work

[11] Jegham, N., Abdelatti, M., & Hendawi, A. (2025). Visual Reasoning Evaluation of Grok, Deepseek's Janus, Gemini, Qwen, Mistral, and ChatGPT. Retrieved from https://arxiv.org/pdf/2502.16428
[14] Bai, Y. (2022). Constitutional AI: Harmlessness from AI Feedback. Retrieved from https://arxiv.org/pdf/2212.08073

		preferred style and composition	
Grammarly	AI-based grammar tool used to check for spelling and grammar errors.	-Can help with brainstorming ideas -Aside with detecting errors in spelling, grammar, and punctuation, can also make suggestions for improving clarity	-Must pay for Premium and Business plan
Notebook LM	An experimental note-taking tool developed by Google that can summarize numerous uploaded files and draw connections between them.	-Has a podcast feature that creates a podcast that summaries information, useful for people that understand information orally better	-Lack of customization options to link your notes