

Data Ethics

Katerina Allmendinger
Nickoal Eichmann-Kalwara

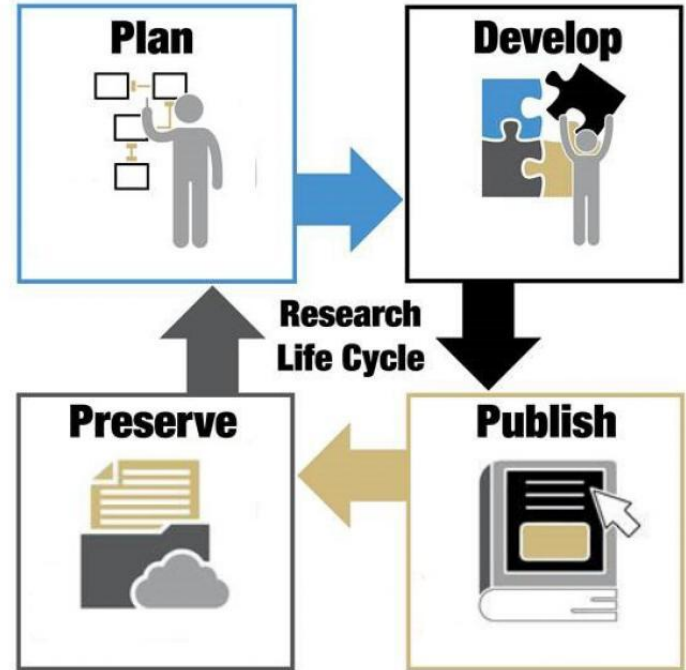
Data Camp 2024



Data are constructed - they are the products of biased human choices and perspectives.

Data Ethics and the Research Lifecycle

- What does ethics [data / research] mean to you?
- At which stages should you critically engage with your data?



Data Ethics and the Research Lifecycle

- Uphold applicable statutes, regulations, professional practices, and ethical standards
- Respect the public, individuals, and communities
- Respect privacy and confidentiality
- Act with honesty, integrity, and humility
- Hold oneself and others accountable
- Promote transparency

Data as Texts/Artifacts

- Data are not neutral or objective.
- Data do not represent a self-evident, discoverable Truth.

Rather,

- Data are constructed, the products of biased human choices and perspectives.
- The interpretation, use, and analysis of data, like any text, is shaped by our biases and perspectives.

Code4Lib Journal, December 2023

Alma - Circulation Statistics

Follow

Document Type

Other

Publication Date

2023

Abstract

Power BI Alma Circulation Template Dashboard

Recommended Citation

"Alma - Circulation Statistics" (2023). *Library Scholarship*. 91.
[redacted].edu/librarian_fac/91

Additional Files

[C4LBibs.pbix](#) (13 kB)

Power BI Alma Bibliographic Template Dashboard

Download

39 DOWNLOADS

Since November 08, 2023

 PLUMX METRICS

Bringing it All Together: Data from Everywhere to Build Dashboards

This article will talk about how [redacted] University approached building a data dashboard bringing together various datasets, from MySQL, vendor emails, Alma Analytics and other sources. Using Power BI, Power Automate and a Microsoft gateway, we can see the power of easy access to data without knowing all of the disparate systems. We will discuss why we did it, some of the how we did it, and privacy concerns.

Open Letter to the *Code4Lib* Journal Editorial Board

We, the undersigned, **no longer have confidence in the *Code4Lib* Journal editorial board to protect library patron privacy and confidentiality**. We do not believe there is a way for the editorial board to regain our trust without **transparent actions** bringing considerably more **editorial attention to privacy, confidentiality, and research ethics** than has been apparent to date.

We, the undersigned, believe that **if patron privacy and confidentiality are not protected**, including from librarians and researchers, in library research and assessment practices that use patron data, such research and assessment practices **are unethical, must not be published, and must be retracted** where they have been published, including in *Code4Lib Journal*.

 to CODE4LIB ▾

Dear all,

as one of the editors of Code4Lib Journal I would like beg your pardon for the security incident.

Since the journal is edited by a group of volunteers and we do not have any formal organizational structure, we as a journal do not have yet an common answer, but I can tell you my private opinion. Right now we are considering the suggestions of the open letter. Some of them could be implemented and there is a high chance that will be implemented. In this particular case we made a couple of editorial, communication related and technical mistakes, but we are aware of the importance of the problem, and I personally disagree that the data breach happened because we did not take care of the ethical concern. In this case - and again speaking only from my part - I did not have the necessary knowledge to check the content of files in a particular (Power BI) format, and thus I was not aware of the real content of that files (the article itself doesn't tell details about the content of the attached file).

It is sure we are taking care of this issue and the open letter, and we act accordingly. I hope that in the following days we will also have a better statement than mine, that will reflect the opinion of all editors.

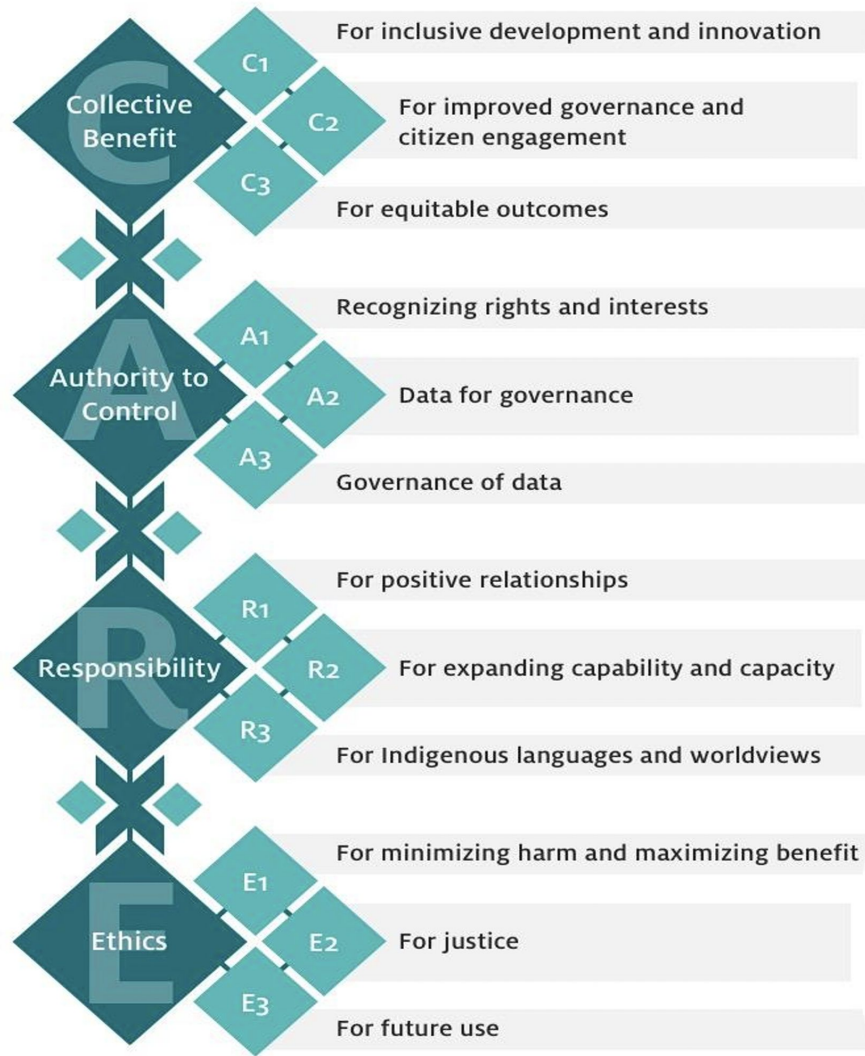
I beg your pardon again,



*machine
actionability*

*social
actionability*

<https://www.gida-global.org/care>



CARE Data Principles

- From the perspective and ethical values of the involved community, what are the potential benefits and harms of collecting, using, or sharing these data now and in the future?
- Does the data play into harmful stereotypes and stigmas, or does it empower the community to assert their rights and interests?
- Are the provenances, purposes, and use limitations stated in the metadata or other documentation?
- Do metadata standards incorporate the (Indigenous) communities values and concepts?



Consent

Privacy

Care work

Harm Reduction

Sovereignty

Power

Privacy

Trust

Justice

Data Camp 2024



Power

- Who is collecting and analyzing the data?
- What's counted, or not?
- How and when something is counted and why?
- Current social and political concerns as influencers?
- Right to be Forgotten (GDPR)

Privacy & Security

- Consent: Permission to use these data (ownership?)
- Anonymization: removing sensitive and identifiable information
- Confidentiality: Whose/what information is protected and for how long?

Trust

- Forms of Knowledges
- Accountability
- Documentation
- Reproducibility

Justice

- Social impact
- Missing and counter data
- Protection
- Co-liberation



Ethics for Artificial Intelligence (AI)

So while AI is novel, it often seems to continue long-standing paradigms of technology in the service of capital ... scholarship suggests that the human harms documented in recent AI-driven initiatives are not merely “teething problems,” but part of a broader paradigm of capitalist and colonialist values at the core of our current economic and technological systems. (Munn, 2023)

Bias



a biological scientist

a chief executive officer

Image from Sun, L. et al. (2023)

Me: "What does an economics professor look like?"

ChatGPT: "An economics professor can vary in appearance like any other individual, as physical characteristics are unique to each person. However, there are certain generalizations one might make. . .may have a well-groomed, salt-and-pepper beard, symbolizing experience and maturity. . .they might wear tailored suits, dress shirts and formal trousers, particularly when giving lectures or attending conferences. . . on more casual days, they might opt for a smart-casual look, combining dress pants with a button-down shirt or a sweater. . .they may be observed carrying a leather briefcase or a satchel" (ChatGPT 2023a).

Image: Quote excerpt from Gross, N. (2023)

Exploitation

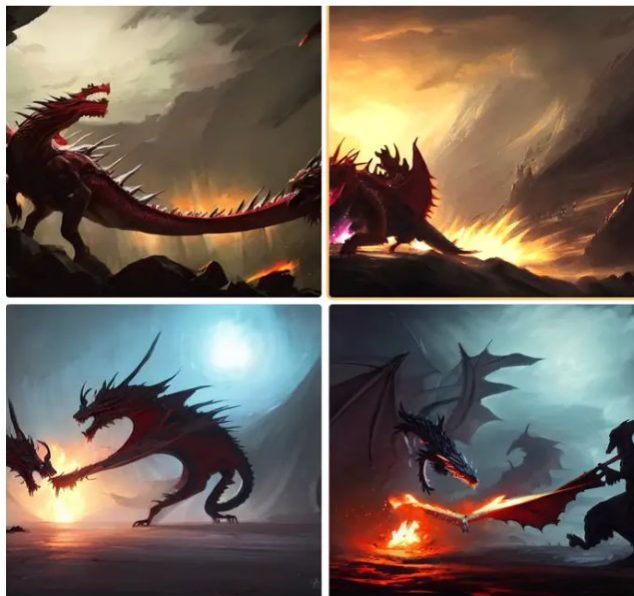


AI-generated art by Sandy60

Data Camp 2024



"Dragon Cage" by Greg Rutkowski.



Images created when Insider typed "Dragon battle with a man at night in the style of Greg Rutkowski" into Stable Diffusion. Stable Diffusion

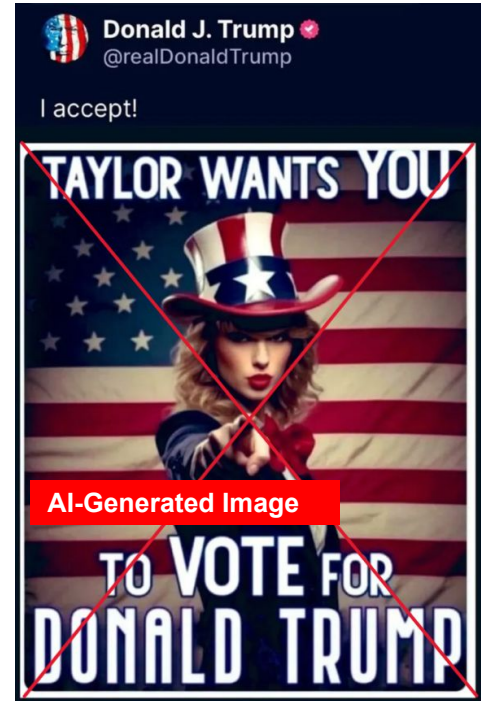
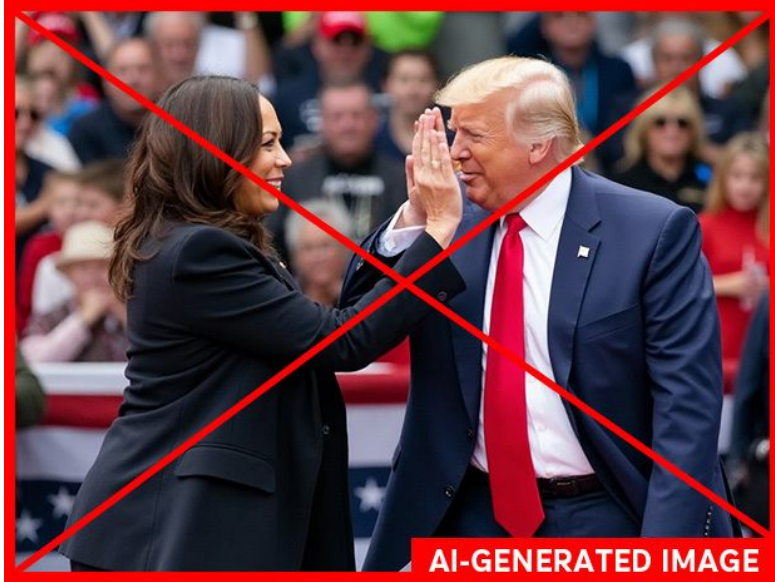
Privacy and Security

- **How is user input data treated?** Is it stored?
Anonymized? Shared with other companies?
- What are the **safeguards** against malicious attacks?
- Are these practices **transparent** and shared with users for **informed consent**?



Image credit: Photo by [Lianhao Qu](#) on [Unsplash](#)

Misinformation



From left: 1) AI-generated image that falsely depicts Kamala Harris and Donald Trump smiling and high-fiving 2) Social media post from Donald Trump sharing an AI-generated image that

Using Generative Artificial Intelligence

Questions to consider:

- How was this tool created and how is it maintained? By whom?
- Who can use this tool and for what purposes?
- How does my use of this tool impact me, my immediate communities, creators of input data, larger social structures, ecosystems, and intersections of those groups?
- ... and what do we do if we cannot answer these questions?

How we work towards Ethical AI

- Frameworks must center Indigenous knowledges and experiences:
 - [Indigenous Protocol and Artificial Intelligence Position Paper](#) (Lewis et al., 2020)
 - [The Five Tests of Māori scholar Sir Hirini Moko Mead](#) (Munn, 2023)
- Be a continual learner - have curiosity towards information you encounter (and generate)
- Is there ethical AI under capitalism?

Transformative and radical uses of AI

A note of hope ...

- [Beth Coleman's Octavia Butler AI](#)
- [Aaron Tucker's work with GANs looking at identity](#)
- [Potential for Indigenous language revitalization](#)
- The work of [Dr. Tiara Roxanne \(Purépecha Mestiza\)](#)
- Here at CU Boulder: [Generative AI and Transness at the Identity Lab](#)

Data are constructed - they are the products of biased human choices and perspectives.

Reflections

- What are examples of data ethical concerns that you encounter in your daily work, scholarship, and/or teaching?
- In what ways do you think *power* is reflected in your data?
- What data are you *missing*? Was it intentional? How might this missing data be beneficial?
- How might thinking about your data as a *text/artifact* change how you relate to, document, and share your work?

References

- Carroll, SC, Garba, I, Figueroa-Rodríguez, OL, Holbrook, J, Lovett, R, Materechera, S, Parsons, M, Raseroka, K, Rodriguez-Lonebear, D, Rowe, R, Sara, R, Walker, JD, Anderson, J and Hudson, M. The CARE Principles for Indigenous Data Governance. *Data Science Journal* (2020). <https://doi.org/10.5334/dsj-2020-043>
- Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>
- Munn, L. The five tests: designing and evaluating AI according to indigenous Māori principles. *AI & Society* (2023). <https://doi.org/10.1007/s00146-023-01636-x>
- Iliadis, A., & Russo, F. (2016). Critical data studies: An introduction. *Big Data & Society*, 3(2). <https://doi.org/10.1177/2053951716674238>
- D'Ignazio, C. & Klein, L. *Data Feminism*. 2020. <https://data-feminism.mitpress.mit.edu/>
- Fair Principles. *GO FAIR*. <https://www.go-fair.org/fair-principles/>
- Open Letter to the Code4Lib Journal Editorial Board, December 2023. <https://docs.google.com/document/d/1c-ADFaum9pB-W4AWQIxE6BPYVJs8goMlrVLoC00ucKc/edit?usp=sharing>

References

- Center for Countering Digital Hate. Fake Image Factories: How AI image generators threaten election integrity and democracy. <https://counterhate.com/research/fake-image-factories/>
- Sun, L., Wei, M., Sun, Y., Suh, Y. J., Shen, L., & Yang, S. (2023). *Smiling Women Pitching Down: Auditing Representational and Presentational Gender Biases in Image Generative AI* (arXiv:2305.10566). arXiv. <http://arxiv.org/abs/2305.10566>
- Gross, N. (2023). What ChatGPT Tells Us about Gender: A Cautionary Tale about Performativity and Gender Biases in AI. *Social Sciences*, 12(8), Article 8. <https://doi.org/10.3390/socsci12080435>
- Mylrea, M., & Robinson, N. (2023). Artificial Intelligence (AI) Trust Framework and Maturity Model: Applying an Entropy Lens to Improve Security, Privacy, and Ethical AI. *Entropy*, 25(10), 1429. <https://doi.org/10.3390/e25101429>
- Ferrara, E. (2024). The Butterfly Effect in artificial intelligence systems: Implications for AI bias and fairness. *Machine Learning with Applications*, 15, 100525. <https://doi.org/10.1016/j.mlwa.2024.100525>
- Meighan, P. J. (2021). Decolonizing the digital landscape: The role of technology in Indigenous language revitalization. *AlterNative: An International Journal of Indigenous Peoples*, 17(3), 397–405. <https://doi.org/10.1177/11771801211037672>

THANKS!

Data Ethics

Katerina Allmendinger
katerina.allmendinger@colorado.edu

Nickoal Eichmann-Kalwara
nickoal.eichmann@colorado.edu

Data Camp 2024

