

Special Forum: AI through the AΩ: Theological Librarians Interact with Artificial Intelligence

Integrating Morality into Generative AI

Libraries Maintaining the Human in the AI Regime

by Michael Robert Smith

The future of postsecondary academic librarianship will be increasingly intertwined with artificial intelligence (AI). Unless higher education religious librarians respond wisely, there is a grave threat that AI's content could undermine the mission of higher education religious institutions. Specifically, wise theological librarians should value and emphasize AI literacy, proactively guard against the dangers of AI bias, and actively seek to shape a generative AI that is committed to advancing their religion's orthodoxy and morality. Acknowledging the severe risk that generative AI will produce biased and dangerous content, postsecondary religious librarians must engage this issue now, or they will soon find themselves on the wrong side of the AI wave with little recourse to rolling back the tide (Ridley and Pawlick-Potts 2021, 2).

BENEFITS OF GENERATIVE AI

Generative AI can help academic librarians perform many traditional librarian tasks, and as various generative AI programs improve, their applications will likely increase in breadth and quality of work product produced. Such tasks include descriptive cataloging, technical services, collection development, subject indexing, database searching, document delivery, bibliographic recommendation services, and drafting of website content, library advertisement content, and library instruction content (Faggioli, Galluzzi, and Weinberger 2020, 213; Adebayo, Bello, Kayode, and Yusuf 2022, 4). Furthermore, AI can provide “better analysis of datasets” than humans can, “eliminate repetitive and tedious tasks,” “discover unexplored concepts,” and “reduce human errors.” It can “help users search for information with ease, help retrieve information across various collections, and help with users’ queries” (Adebayo, Bello, Kayode, and Yusuf 2022, 4–7). AI will increasingly enhance productivity, reduce costs, and create content library patrons rely on in their research and content selection processes. So, then, considering all of these benefits, why worry?

FUNDAMENTALS OF GENERATIVE AI

Generative AIs like ChatGPT are exceedingly complicated and often profoundly opaque, multidimensional systems. Such systems are an amalgamation of an algorithm, a data set, and feedback. Of these three, the algorithm is of primary and central significance. The algorithm determines how the data set is to be interpreted, analyzed, valued, and utilized, and it is the lens through which feedback is likewise made meaningful and interpreted. An algorithm, depending on its function, can be understood to be a set “of instructions or sequences of logical steps for a computer to use on a body of data to accomplish a task” or “as a recipe, or a step-by-step guide that prescribes how to obtain a certain goal, given specific parameters” (Archambault 2023, 531). Algorithms are thus neither “neutral nor

Michael Robert Smith is library services coordinator at Colorado Christian University.

value free” as they are both animated by and limited by the decisions of the humans who design them (Archambault 2023, 531).

The biases “embedded in technology are more than mere glitches; they’re baked in from the beginning. They are structural biases, and they cannot be addressed with a quick code-update” (Broussard, 2023, 86). “Technology discriminates by default” because it must rank-order information to make sense of it and also to perform its function (Broussard, 2023, 88). So long as foundational elements of algorithmic code define and limit the AI’s generative output, eliminating biases through expanded data sets and feedback is an uphill battle at best. Just as the same ingredients can produce bagels, croissants, and cookies, depending on the recipe (instructions) the baker employs, so too the type of (biased) work-product created by generative AI depends not on the data set fed to the AI but largely depends instead on the algorithm’s instructions and definitions as applied to the “ingredients” (data set and feedback).

HAZARDS OF GENERATIVE AI

The big issue is that generative AI content will always directly result from the values, priorities, definitions, and formulas built into the AI’s animating algorithm. This would not be a concern, except that the algorithmic formulas currently being crafted are based on values and hierarchies that conflict with and often contradict numerous fundamental religious values. For example, “in an AI-dominated world, ‘intelligence’ implies only the analytic function of computation” (Al-Kassimi 2023, 3). Human intelligence involves more than crunching numbers; it incorporates “values such as empathy, care, compassion, and love” (Al-Kassimi 2023, 17). The reality is that “prudence is not possible without reference to an epistemology that distinguishes between good and bad for the human—or, more precisely, a science of human nature” (Al-Kassimi 2023, 18). Information without a coherent and reliable worldview is devoid of meaning (Al-Kassimi 2023, 5), so AI backend engineers must incorporate some type of values, priorities, and definitions into their formulas. Secular librarians and algorithmic engineers may want to “infuse into machine learning the democratic and humanitarian values that libraries uphold” (Faggioli, Galluzzi, and Weinberger 2020, 213). However, this position begs these questions: Is there agreement on what these democratic and humanitarian values are? What philosophical or worldview grounding supports such values? Do said values conflict with the religious values of the various major world religions?

Consider, for example, an AI built on the definition of human nature as derived from an evolutionary, atheistic worldview; it may likely conceptualize human essence and telos concerning the goals of “eliminating death, aging, and suffering” (Al-Kassimi 2023, 3). When underlying worldview premises are taken to their logical conclusions, human beings are mere cosmic accidents, products of cause and effect playing out in the physical realm, devoid of any objective purpose, of any objective basis for morality, and of any ultimate meaning (Sire 2009, 242). Any “significance” adopted ad hoc is a glaring contradiction and comical absurdity. Furthermore, materialist atheistic postmodernity proclaims that all truth claims are relative and that objective truth is illusory. Building a moral generative AI on the premises and logical conclusions of a postmodern atheistic materialist worldview would prove illogical, self-contradictory, and destined to a slippery footing on unstable ground.

Acknowledging that attempting to ground an ethical AI on an evolutionary atheistic worldview is unstable, postmodern materialists might argue instead that AI should be taught in each major world religion to seek to draw out a “universal ethic.” This, too, is an illogical and impossible task. All major world religions make ontological and teleological truth claims that contradict one another (Sire 2009, 19). The major global religious traditions—Hinduism, Buddhism, Confucianism, Daoism, Judaism, Islam, and Christianity—cannot be reconciled, and no substantial universal ethic

can be derived from them that does not fragment upon definitional-comparative-analysis (Smith 1991, 385). Ultimately, any supposed universal ethic as a grounding for an AI would be a recipe for disaster for postsecondary academic librarians of particular religious affiliations seeking to provide their patrons with sound and reliable information and resources about their religions.

A RECOMMENDED RESPONSE

This predicament facing postsecondary religious academic librarians requires, at minimum, a two-tiered response. First, academic librarians must embrace the need to develop a baseline of algorithmic literacy. Second, academic librarians should seek to proactively participate in producing AI tools intentionally designed to create content in line with the orthodox beliefs of their religious traditions.

Concerning algorithmic literacy, academic librarians should affirm that “libraries have traditionally played a central role in making emerging technologies accessible to their communities” and that “an informed, algorithmically literate public is better able to negotiate and employ the complexities of AI” (Ridley and Pawlick-Potts 2021, 1, 5). AI literacy has been defined as “a set of competencies that enables individuals to critically evaluate AI technologies; communicate and collaborate effectively with AI; and use AI as a tool online, at home, and in the workplace” (Long and Magerko, “What Is AI Literacy?,” 2, quoted in Ridley and Pawlick-Potts 2021, 4). Essentially, those using AI would be well served to understand what a particular AI can do, what it can’t do, how it operates (“as relational, contingent, [and] contextual”), and the extent to which human agency plays a primary role in the “human-machine interaction” (Kitchin, 2017, 15, quoted in Ridley and Pawlick-Potts 2021, 2, 8). This, however, can be challenging because the algorithms that animate the AI are controlled “by those who build and deploy algorithms, not those who use them” (Ridley and Pawlick-Potts 2021, 4). Further, the algorithms are often “complex, opaque, invisible, [and] shielded by intellectual property protection,” and there is a “lack of assessment instruments” in the marketplace (Ridley and Pawlick-Potts 2021, 7). Even still, “algorithms are the new power brokers in society,” and it is incumbent on librarians to endeavor to inform themselves and their patrons about these new realities by preparing and teaching algorithmic literacy (Diakopoulos, 2014, 2, quoted in Ridley and Pawlick-Potts 2021, 9).

Second, higher-education religious academic librarians must proactively seek to participate in forming the generative AIs they will eventually incorporate into their work and work products. Specifically, librarians should initially seek out, use, and promote explainable AI, which has “the ability to explain their rationale, characterize their strengths and weaknesses, and convey an understanding of how they will behave in the future” (Turek, 2016, quoted in Ridley and Pawlick-Potts 2021, 8). Such explainable AI values transparency, understanding, and accountability (Ridley and Pawlick-Potts 2021, 8). Unfortunately, because of competition between AI creators, transparency is elusive, and identifying which AI to adopt and support is tricky. It is easier to determine which AI should not be used and supported in its growth, based on the biased content created, than to determine which AI are most “explainable.”

Given these challenges, and perhaps to move beyond them, religious librarians should consider supporting the creation of a unique generative AI that explicitly incorporates the orthodox teachings of their faith. Suppose postsecondary religious librarians want a generative AI to produce content aligned with their faith’s orthodoxy and guide patrons to the best and most appropriate content. In that case, they should advocate for and support the production of religion-specific generative AIs. Of course, such religion-specific generative AIs would intrinsically be biased, but then again, the goal cannot be to create an unbiased generative AI, for that is impossible (as demonstrated above). What

such an AI would “look like” in detail and how it could be built is beyond the scope of this article. However, those serious about engaging the possibility should seriously consider Mark Graves’s superb 2022 article, “Theological Foundations for Moral Artificial Intelligence,” and reach out to the leaders of their respective faiths to raise awareness of this issue, hopefully prompting a productive response.

CONCLUSION

Inevitably, academic librarians must face the brave new world that includes generative AI tools in performing their duties. As technology changes and economies squeeze budgets, libraries will increasingly be pressured to “show value for money and demonstrate cost-effective practices” (Adebayo, Bello, Kayode, and Yusuf 2022, 8). So, the adoption and incorporation of generative AI into library practice is almost inevitable. In such a world, religious academic librarians should value and actively pursue algorithmic literacy, proactively teach patrons algorithmic literacy, prioritize explainable AI, and assume a hands-on role in seeking out and helping to develop a unique generative AI grounded on the tenets of their faith. Academic librarians will be on the right side of history by empowering patrons in this brave new world and endeavoring to protect patrons from the spiritual dangers of biased generative AI.

REFERENCES

- Adebayo, O. A., L. A. Bello, J. O. Kayode, and T. I. Yusuf. 2022. “Adoption of Artificial Intelligence for Effective Library Service Delivery in Academic Libraries in Nigeria.” *Library Philosophy and Practice*. <https://digitalcommons.unl.edu/libphilprac/6804/>.
- Al-Kassimi, Khaled. 2023. “A Postmodern (Singularity) Future with a Post-Human Godless Algorithm: Trans-Humanism, Artificial Intelligence, and Dataism.” *Religions* 14, no. 8: 1–26. <https://doi:10.3390/rel14081049>.
- Archambault, Susan G. 2023. “Expanding on the Frames: Making a Case for Algorithmic Literacy.” *Communications in Information Literacy* 17, no. 2: 530–53. <https://doi:10.15760/comminfolit.2023.17.2.11>.
- Broussard, Meredith. 2023. “How to Investigate an Algorithm: Algorithmic Auditing Has the Potential to Decrease Bias and Prevent or Fix Harms Caused by Artificial Intelligence.” *Issues in Science and Technology* 39, no. 4: 85–95. <https://doi:10.58875/oake4546>.
- Diakopoulos, Nicholas. “Algorithmic Accountability Reporting: On the Investigation of Black Boxes” (New York: Tow Center for Digital Journalism, Columbia University, 2014). <https://doi.org/10.7916/D8ZK5TW2>.
- Faggioli, Chiara, Anna Galluzzi, and David Weinberger. 2020. “When the Best Librarian Is an AI.” *AIB studi* 60, no. 2: 213–217.
- Graves, Mark. 2022. “Theological Foundations for Moral Artificial Intelligence.” *Journal of Moral Theology* 11: 182–211. <https://doi.org/10.55476/001c.34130>.
- Kitchin, Rob. 2017. “Thinking Critically about and Researching Algorithms.” *Information, Communication & Society* 20, no. 1: 14–29. <https://doi.org/10.1080/1369118X.2016.1154087>.

- Long, Duri, and Brian Magerko 2020. "What Is AI Literacy? Competencies and Design Considerations." *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Honolulu, HI: Association for Computing Machinery. <https://doi.org/10.1145/3313831.3376727>.
- Ridley, Michael, and Danica Pawlick-Potts. 2021. "Algorithmic Literacy and the Role for Libraries." *Information Technology and Libraries* 40, no. 2: 1–15. <https://doi:10.6017/ital.v40i2.12963>.
- Sire, James W. 2009. *The Universe Next Door: A Basic Worldview Catalog 5th ed*. Nottingham, England: InterVarsity Press.
- Smith, Huston. 1991. *The World's Religions: Our Great Wisdom Traditions*. San Francisco: HarperSanFrancisco.
- Turek, Matt. 2016. "Explainable Artificial Intelligence (XAI)." Arlington, VA: DARPA. <https://www.darpa.mil/program/explainable-artificial-intelligence>.