Ground Truth: AI Threats to Critical Dam Infrastructure

Although AI has been used to assist with safer dam construction and to locate risky dams, leverage deep learning techniques to analyze operational patterns, and evaluate effectiveness, such reliance on these technologies—often linked to online computer systems—has also created a new level of threats. The most obvious of these threats comes from cyberhackers, as seen in the infamous 2013 incident when Iranian hackers breached a New York dam and tried to open the floodgates, but were thwarted by a sluice valve that had been disconnected for maintenance.

Currently, there is a fractured national patchwork of practices related to dam management, with only 5% of the 91,827 dams in the United States falling under federal regulation. Nearly two-thirds of U.S. dams are privately owned, meaning those owners are responsible for upkeep and repairs, including complying with inspectors' requests. As more and more utilities companies look to AI to reduce costs by eliminating personnel, potentially not properly training existing personnel on AI applications, the threats will only increase. And incorporation of AI into any water or utilities system collectively heightens the risk of network-wide failures.

As an emerging field, the use of AI is also largely unregulated, and its proliferation could trigger serious and unexpected problems, including system-wide compromise owing to design errors and malfunction. As an example, a Cornell University study noted that "an AI algorithm tasked with minimizing damages in the event of a dam failure could inadvertently prioritize reduction of economic losses at the expense of human life if it were accidentally programmed to optimize for wrong or overly narrow goal ranges."

AI is becoming increasingly reliable in engineering, but it still requires human supervision to ensure results are accurate and any potential biases or errors in the data are identified and corrected. AI models are only as good as the data they are trained on, so it is important to ensure any data used to train the model is accurate, unbiased, and representative of real-world conditions.

Additionally, AI models still need to be validated and verified by human engineers, who should be *trained* to interpret and explain the results and predictions of those AI models and should also be *retained* to oversee successful implementation and practice.

In the case of safety assessment of dams and hydropower projects, it is crucial for human engineers to validate the AI results and predictions, as the sanctity and well-being of people, property, and entire communities depend upon it as a matter of life and death.

References:

Chris Riotta, “Hacking the Floodgates: US Dams Face Growing Cyber Threats,” Bank Info Security, April 18, 2024

Susan Partain, “The Risks and Possibilities of AI for Utilities,” American Public Power Association, April 17, 2024

Dario Amodei, Chris Olah, Jacob Steinhardt, Paul Christiano, John Schulman, and Dan Mané, “Concrete Problems in Safety,” Cornell University, July 25, 2016.

M. Amin Hariri-Ardebili, Golsa Mahdavi, Larry K. Nuss, and Upmanu Lall, “The role of artificial intelligence and digital technologies in dam engineering: Narrative review and outlook.” Engineering Applications of Artificial Intelligence, November 2023.

Jill Castellano, Tracy Loew and Rosalie Murphy, “Oroville crisis highlights risky dams, spotty inspections around U.S.,” *USA Today*, February 16, 2017.

“What Are AI Hallucinations?” IBM.com.