# Letter from the Special Issue Editor

*Data-centric Responsible AI* is becoming increasingly critical as AI is widely used in our everyday lives. In addition to simply improving model performance, it is important to make sure the trained model is trustworthy and responsible in the sense that it is fair, robust, private, secure, explainable, aligns with values, and more. Moreover, AI is only as good as its data, so we must take a data-centric approach and improve the data itself to fundamentally solve these problems. Recently applications like Large Language Models (LLMs) have remarkable performance largely because of the large amounts of data they are trained on, so data-centric research is only going to become more important in the future. This issue is thus timely and contains recent solutions by leading experts in this field.

The first three papers propose Data-centric Responsible AI methods that can be applied at different stages in the machine learning pipeline. The paper *Coverage-based Data-centric Approaches for Responsible and Trustworthy AI* by Shahbazi et al. proposes data coverage methods to identify and resolve misrepresentation of minorities in data. The goal is to identify and resolve insufficient data coverage and generate data-centric reliability warnings to help data scientists determine if a prediction is reliable. Recent generative AI and foundation models can benefit from these techniques to effectively augment datasets with synthetic data. Next, the paper *Overcoming Data Biases: Towards Enhanced Accuracy and Reliability in Machine Learning* by Zhu and Salimi explores how causal modeling can improve the data cleaning, preparation, and quality management for machine learning. Causal reasoning can effectively identify and correct data biases resulting from missing data, confounding variables, and measurement errors and thus improve the fairness and accuracy of machine learning models. Finally, the paper *Fairness and Robustness in Answering Preference Queries* by Roy outlines algorithmic challenges and directions for systematically changing the original aggregated output to satisfy different criteria related to fairness and robustness. The author considers different scenarios on how users provide their input preferences and how the individual preferences get aggregated.

The next two papers explore interesting domains that provide inspiration to further advance Data-centric Responsible AI. The paper *On the Robustness of ChatGPT: An Adversarial and Out-of-distribution Perspective* by Wang et al. performs a thorough evaluation of the robustness of ChatGPT and other LLMs from an adversarial and out-of-distribution perspective. While LLMs are receiving significant attention nowadays, their robustness to unexpected inputs is still understudied, which is a concern especially for safety-critical applications. The authors leverage multiple recent datasets for adversarial robustness and show that ChatGPT performs better and others, but also has much room for improvement. The paper *Red Onions, Soft Cheese and Data: From Food Safety to Data Traceability for Responsible AI* by Grafberger et al. makes the interesting analogy that data traceability for Responsible AI is akin to ensuring food safety. In particular, the U.S. Food and Drug Administration (FDA) detects outbreaks of foodborne illnesses, discovers contaminated food, and conducts traceback investigations through the food supply chain to determine the root cause and issue a comprehensive product recall. Taking inspiration from this process, the authors propose a data-centric vision for Responsible AI that involves prediction monitoring, data tracing, and identifying contaminated data and pipeline steps through audits.

Overall, these works represent the state-of-the-art data management approaches for Data-centric Responsible AI from various angles. We are just scratching the surface, and the data management community is well positioned to eventually realize this vision.

Steven Euijong Whang
Korea Advanced Institute of Science and Technology