# Big data Analytics

L0 Mohamed Diana
Document 2

This article proposes a participatory data-centric approach to AI Ethics by Design, emphasizing the role of a bridge builder to facilitate collaboration between machine learning developers and domain experts. By focusing on data activities and related ethical and epistemic challenges, this approach aims to create strong communication channels and orchestrate different types of domain-specific. In critical domains like healthcare and law enforcement, where data is scarce, careful curation of datasets is crucial.
The early adoption of a "design turn" in IT ethics advocated for a proactive approach to system development, prioritizing upfront ethical considerations over post-hoc analyses. Value sensitive design, proposed by Friedman aimed to integrate ethics into Human-Computer Interaction and Computer Supported Cooperative Work.
Nissenbaum introduced the concept of "engineering activism," emphasizing the involvement of humanities and social sciences in system development, down to technical details. Value sensitive design initially distinguished itself by involving stakeholders and employing a tripartite iterative method focusing on ethical and human values.
However, recent iterations of value sensitive design have shifted focus away from technical investigations, primarily emphasizing philosophical evaluations of existing technologies and conceptual exploration of values, rather than proactive collaboration with developers to embed values into IT systems, the success of AI has come at the expense of awareness regarding issues such as validation, verification, and explainability. As AI is increasingly adopted in critical societal domains, it is crucial to anticipate ethical and epistemic issues at an early developmental stage.
The process of ethical AI design involves various stages, including data exploration, system design, and deployment, each requiring careful consideration of ethical principles and stakeholder involvement.
Tools like datasheets and model cards are used to document datasets and model performance characteristics, facilitating transparency and accountability throughout the system lifecycle. The ethical requisites of system design are formulated based on value identification, and prototypes are developed to test the system's resilience to value-related challenges. However, challenges remain in ensuring interdisciplinary collaboration and addressing AI-specific issues, such as fairness and transparency, throughout the development process. A participatory data-centric approach to AI Ethics by Design emphasizes the importance of engaging domain experts in data activities, but further attention is needed to address challenges related to model training and verification.
It emphasizes the role of a "bridge builder" with interdisciplinary competences in computer science and ethics to facilitate collaboration between stakeholders. Overall, integrating ethical considerations into AI design requires a holistic approach that spans from data exploration to system deployment, with a focus on interdisciplinary collaboration and stakeholder involvement.
The quality of data preparation, particularly in less formalistic domains like public administration, is crucial for machine learning model performance. Challenges include accurately labeling data and defining socially constructed phenomena. Drawing from human computer interaction traditions, participatory design practices are suggested to democratize system development and engage users as competent practitioners. Collaborative workshops and tools like conceptual sketching can help reveal domain knowledge and empower domain experts. Value elicitation activities are proposed to explore ambiguous insights concerning value issues in data. The role of the bridge builder extends to raising awareness of power mechanisms affecting ground truth in data.
projects intersect with complex socio-political issues, such as documentation demands in the elderly care sector, requiring attention to broader implications beyond technical aspects.
Machine learning systems accumulate technical debt, requiring ongoing maintenance and management. Decision-makers need to be aware of the organisational commitment required to manage technical debt effectively.
The participatory data-centric approach offers a practical means of addressing ethical challenges in AI design and deployment. Bridge building between stakeholders is essential for fostering communication and integrating domain-specific knowledge into machine learning projects.