(Data) Ethics and LLMs in the Cloud

de.KCD Summer School 2024

Lisa Schaak
Group Cloud & IT
Institute of Digital Health

17.09.2024

de.KCD

BIH Berlin Institute
of Health
@Charité

Agenda

1. Introduction Survey

2. Theoretical Concepts

3. Discussion Example

4. Tools, Laws & Regulations

1.

Introduction Survey

What is your most pressing ethical concern regarding LLM usage?

2.

Theoretical Concepts

(Artifical) Intelligence?

1. What is human intelligence?

2. Operational definition of AI

3. Deductive vs. inductive reason

Bugs or Features?

1. Bias

2. Explainability

3. Hallucination

Different Ways of Using LLMs in Research

1. LLM as research object

2. (Proprietary) LLM as helping tool

3. LLMs in primarily non-IT specific research fields

4. Developing new LLMs for everyday use

5. LLMs and potential dual use constellation

Theories of Morality

1. Ethics of Virtue

2. Deontology

3. Utilitarianism

Additional Influences

1. Cultural Differences

2. Stakeholders' Interest

3. Questions of Power

(Conflicting) Values in Using LLMs for Research

1. Quality of services

2. Privacy

3. Transparency

4. Personalisation

5. Solidarity

6. Convenience

7. Dignity

8. Explainability

9. Accuracy

10. Fairness

11. Autonomy

12. Equality

13. Efficiency

14. Sustainability...

Risk Assessments

1. Type 1 vs. type 2 error

2. Rate consequences

3. Control group

Possible Remedies

1. Human oversight

2. Education

3. Engagement

3.

Discussion Example

10.11.2024   (Data) Ethics and LLMs in the Cloud

Discussion Example: Pole Data Set

1. Choose a research question and a way of LLM usage.

2. Identify possible stakeholders and interests.

3. Name 3 pro and 3 con arguments using the 3 theories of morality.

4. What types of error can occur? How risky are those errors?

5. Bonus question: How ethically is it to use the data set as example?

4.

Tools, Laws & Regulations

10.11.2024 (Data) Ethics and LLMs in the Cloud

AI Governance & Best-Practice

- EU HLEG -> AI Act

- IEEE

- WHO

- Tripod

- Z-inspection

The unmet promise of trustworthy AI in healthcare:
why we fail at clinical translation (V. Madai et al, 2024)

- The first translational failure: many exploratory studies, little validation

- The second translational failure: the gap between ethical principles and ethical practice

de.KCD   BIH Berlin Institute of Health @Charité

Questions?

Thank your for your participation at
de.KCD Summer School 2024

Lisa Schaak
Group Cloud & IT
Institute of Digital Health

17.09.2024

de.KCD

BIH Berlin Institute
of Health
@Charité