AI-driven Value Creation vs. Information Privacy: An unbridgeable Discrepancy?

Analyse des Konflikts zwischen Privatsphäre und KI-Systemen

Aufgrund der steigenden Menge an verfügbaren Daten, bezahlbaren hohen Rechenkapazitäten und entscheidenden technologischen Durchbrüchen, ist Künstliche Intelligenz (KI) nicht mehr nur Teil von großen wissenschaftlichen Projekten, sondern wird von vielen Unternehmen bereits gewinnbringend eingesetzt: sei es durch die Effizienz-Steigerung von Unternehmensprozessen oder die Aufwertung des eigenen Produktportfolios – der Einsatz von KI entwickelt sich immer mehr zum entscheidenden Wettbewerbsvorteil.

Aus technischer Perspektive, basieren KIs auf bestehenden Zusammenhängen, die aus Daten mittels Algorithmen extrahiert werden. Um KIs zu erstellen wird daher eine bestimmte Menge an Daten aus problem-spezifischen Domänen benötigt, die entsprechende Zusammenhänge abbilden. Wenn ein zu lösendes Problem allerdings voraussetzt, dass eine KI auf Grundlage von personenbezogenen Daten erstellt wird, können KI-basierte Lösungen eine Gefahr für die Privatsphäre des Einzelnen darstellen. Beispielsweise besteht die Gefahr, dass Dritte unberechtigten Zugriff auf den zugrundeliegenden Datensatz erhalten oder bereitgestellte Informationen durch Interaktion mit der KI nachgebildet werden können. Um dieser potentiellen Gefahr entgegenzuwirken werden immer mehr Initiativen zur Sicherung der Privatsphäre ins Leben gerufen. Zum Beispiel verlangt die neue Datenschutz-Grundverordnung (DSGVO/GDPR) der EU, die Anonymisierung und das gezielte Löschen bestimmter personenbezogener Datensätzen, die von Unternehmen gesammelt wurden.

Während das grundlegende Ziel hinter solchen Initiativen die Reduzierung der Speicherung verfügbarer personenbezogener Informationen darstellt und somit den Schutz von Privatsphäre verfolgt, können sie jedoch die Umsetzung von KIs gefährden: Die Anonymisierung und Löschung vereinzelter Datensätze kann zum Verlust von Informationen führen, die Rückschlüsse auf Zusammenhänge enthalten, die für die Erstellung einer bestimmten KI essentiell sein können. Entsprechend kann der Schutz der Privatsphäre des Einzelnen die Qualität von angestrebten KIs signifikant beeinträchtigen und in vielen Fällen sogar die Realisierung verwehren. Bedingt also der Einsatz von KI in bestimmten Domänen den Verzicht auf Privatsphäre des Einzelnen? Oder sind die Ziele beider Seiten letztendlich doch miteinander vereinbar?

Zielsetzung der Arbeit

Gegenstand dieser Arbeit ist die Identifikation von Dimensionen im Hinblick auf den Informationsgehalt von Daten bzgl. Anforderungen Privatsphäre-schützender Initiativen einerseits und Anforderungen der KI-Technologie andererseits zu erarbeiten und gegenüberzustellen. Die Bearbeitung ist auf Deutsch oder Englisch möglich.

Fragestellung

- Welche Anforderungen stellt die KI Technologie an den Informationsgehalt von Daten, um einsetzbare KI Lösungen erzielen zu können? Welche Privatsphäre-Risiken können durch den Einsatz von KI entstehen?
- Wie beeinflusst die Gewährleistung der Privatsphäre des Einzelnen die Verwendbarkeit von Informationen?
- Inwiefern lässt sich der Einfluss von Maßnahmen zur Gewährleistung der Privatsphäre des Einzelnen auf Daten einerseits mit den Anforderungen der KI Technologie an Daten andererseits vereinbaren? Bedingt der Einsatz von KI den Verzicht auf Privatsphäre des Einzelnen?

Methodik / Vorgehensweise

Beispielsweise kann, je nach Ausgestaltung, eine oder beide der folgenden Methoden verwendet werden:

- **Strukturierte Literaturrecherche** zum Einfluss von Maßnahmen zur Gewährleistung der Privatsphäre auf Daten einerseits und der Anforderungen der KI Technologie an Daten sowie potentielle verbundene Privatsphäre-Risiken andererseits
- **Experten Interviews** zur Erweiterung der literaturbasierten Erkenntnisse und konkreten Untersuchung der Kombinierbarkeit des datenbezogenen Privatsphäre-Einflüsse und KI-Anforderungen

Beginn / Betreuer

Beginn ab sofort möglich. Bei Interesse bitte melden bei: Timo Sturm (sturm@is.tu-darmstadt.de, S1|02 237a)

WIRTSCHAFTS INFORMATIK

Bachelor/Seminar/Master Thesis

TECHNISCHE UNIVERSITÄT DARMSTADT

AI-driven Value Creation vs. Information Privacy: An unbridgeable Discrepancy?

Analysis of the Conflict between an Individual's Privacy and AI-driven Systems

Due to the increasing amount of available data, affordable high computing capacities and decisive technological breakthroughs, Artificial Intelligence (AI) is no longer only part of large scientific projects, but is already being used profitably by many companies: whether by increasing the efficiency of company processes or upgrading their own product portfolio - the use of AI is increasingly developing into a decisive competitive advantage.

From a technical perspective, AIs are based on existing associations that are extracted from data using algorithms. Thus, in order to create AIs, a certain amount of data from problem-specific domains is required to capture the corresponding connections. However, if a targeted problem requires an AI to be based on personal data, AI-based solutions can represent a threat to an individual's privacy. For example, it may be possible for third parties to gain unauthorized access to the underlying data set or to reproduce information provided through interaction with the AI. In order to address this potential risk, more and more initiatives are being launched to ensure information privacy. For example, the EU's new data protection regulation (DSGVO/GDPR) requires the anonymization and systematic deletion of certain personal data records collected by companies.

While the basic objective behind such initiatives is to reduce the storage of available personal information and thus to protect privacy, they may jeopardize the implementation of AIs: The anonymization and deletion of individual data records can lead to the loss of information that contains inferences about relationships that can be essential for the creation of a particular AI. Accordingly, the protection of an individual's privacy can significantly impair the quality of targeted AI use cases and in many cases even prevent their realization. So does the use of AI in certain domains mean that the privacy of the individual has to be sacrificed? Or are the goals of both sides after all compatible with each other?

Objective of the Thesis

The object of this work is to identify and compare dimensions with regard to the amount of information contained in data with regard to the requirements of privacy protection initiatives on the one hand and the requirements of AI technology on the other. It is possible to write the thesis in German or English.

Research Questions

- Which requirements does AI technology require of the amount of information contained in data in order to achieve applicable AI solutions? What privacy risks can arise from the use of AI?
- How does ensuring information privacy affect the usability of information?
- To which extent can the impact of information privacy policies be reconciled with the data requirements of AI technology? Does the use of AI imply the renunciation of the individual's privacy?

Methodologies / Research Process

For example, the thesis may include one or both of the following methodologies:

- **Structured literature review** on the impact of privacy policies on data on the one hand and the requirements of AI technology on data and potential related privacy risks on the other hand
- **Expert interviews** to expand the literature-based findings and to investigate the combinability of data-related privacy impacts and AI requirements.

Start / Supervisor

Start immediately possible. If you are interested, please contact: Timo Sturm (sturm@is.tu-darmstadt.de, S1|02 237a)