Data Privacy For AI Fraud Detection Models

A Framework For GDPR Compliant AI

Authors: Kadir Ider and Andreas Schmietendorf

Presenter: Kadir Ider, Delivery Hero SE, kadir.ider@deliveryhero.com

Date: October 14, 2020

Delivery Hero

Hochschule für Wirtschaft und Recht Berlin

Berlin School of Economics and Law

IARIA

Kadir Ider
Delivery Hero SE
Berlin, Germany
email: kadir.ider@deliveryhero.com

Global data protection manager with a background in business intelligence and process management and doctoral candidate with focus on technology compliance

Prof. Dr. Andreas Schmietendorf
Berlin School of Economics and Law
Berlin, Germany
email: andreas.schmietendorf@hwr-berlin.de

Professor for business informatics with focus on system development, implementation of integration architectures, quality assurance of software development based on metrics

Wordcloud. Source: Kadir Ider, 2020.

Common buzzwords of our research projects

Agenda

1. Topic Relevance
2. Research Development & Gaps
3. Data Protection and Privacy
4. AI Compliance Framework
5. Use Case: Fake Reviews
6. Project Milestones

1. Topic Relevance

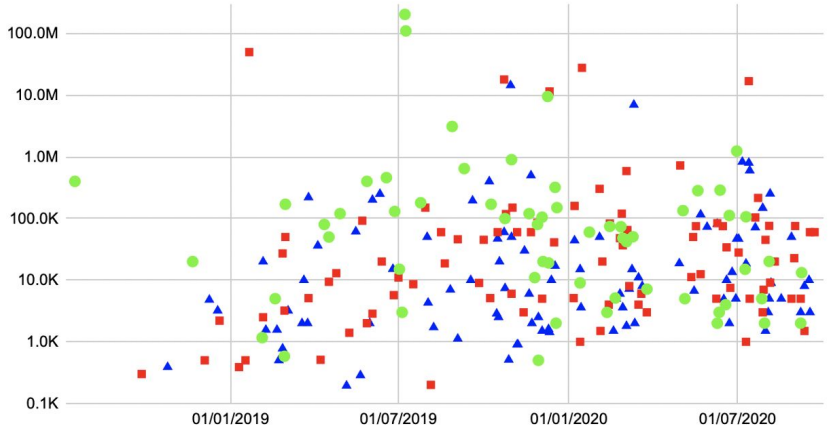
Economic and strategic importance of compliant AI models

- Financial damage in the US is estimated to amount to USD 42 billion, with a large fraction contributed by cybercrime
- False positive fraud detection three times higher than the detection of true positive (USD 120 billion)
- 50% of businesses experienced fraud within a 24 months period, of which 50% employ AI for fraud detection
- Effective algorithms require adequate input data, incl. personal identifiers

- Goals and challenges:
  - Increasing detection accuracy and decreasing false positive
  - Minimize monetary losses and accelerate business
  - Manage the trade-off between privacy and accuracy
  - Improve processing transparency and reduce blackbox problem
  - Manage data protection by design and by default, in accordance with Art. 25 GDPR
  - Promote principles relating to processing of personal data

400 GDPR fines due to non-compliance

- Non-compliance with data processing principles and data breach obligations (△)
- Insufficient legal basis for data processing (▫)
- Insufficient technical and organisational measures to ensure information security (◯)



Source: K. Ider, Effective Privacy Management Concepts: Increasing Privacy Control by Reducing Complexity, 2020.

2. Research Development & Gaps

Fraud detection segmentation

- **Blacklists:** reactive, static characteristic
- **Rule engines:** somewhat proactive, partially reactive, high maintenance
- **AI solutions:** proactive, prediction accuracy and multitude of input features , transparency
- **Common fraud activities:** identity theft, account takeovers, abuse of promotions, fake reviews or -listings

Research of AI based fraud detection models

- Major focus on development, assessment of features, comparison of fraud detection algorithms performance
- Comparison fraud detection algorithms performance
- Assessment of elements of trustworthiness in the usage of AI
- Classification techniques and improving AI models prediction accuracy

Research Gap from a compliance perspective

- Transparency and accountability for PII adherent to the GDPR marginal
- Technology introduces new risks to data but more importantly to individuals,
- Development of an AI privacy framework for AI models reduce present shortcomings and improve the accountability requirements pursuant to Art. 5 (2) GDPR

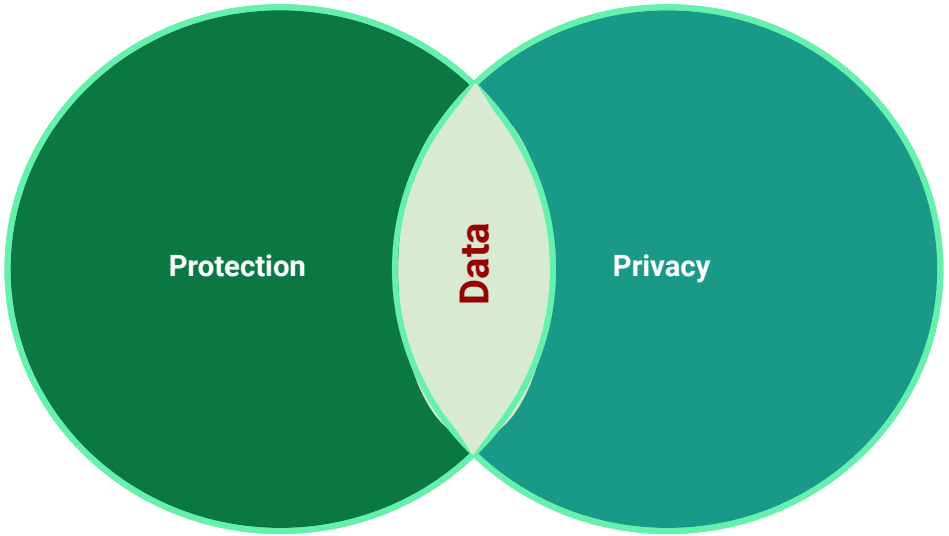
3. Data Protection and Privacy

Data Protection

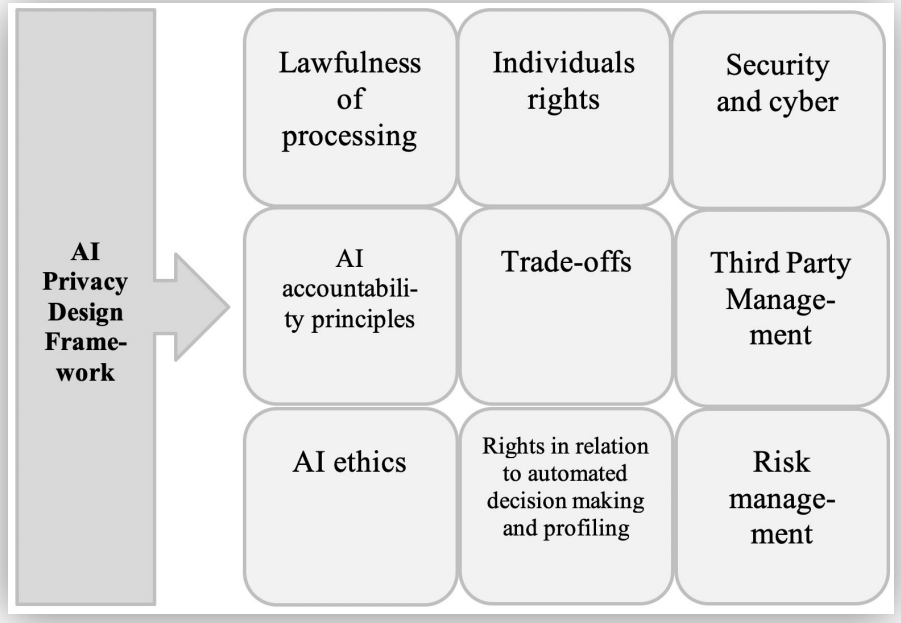
- Legal mechanism (e.g. GDPR)
- Ensures lawful processing
- Basis for data privacy
- Not individual centric, i.e. one “umbrella” for all individuals

Data Privacy

- Defines guidelines for purpose and means of processing
- Ensures user rights (to control own data)
- It is a right of every individual, i.e. “umbrella” for each person



4. AI Compliance Framework



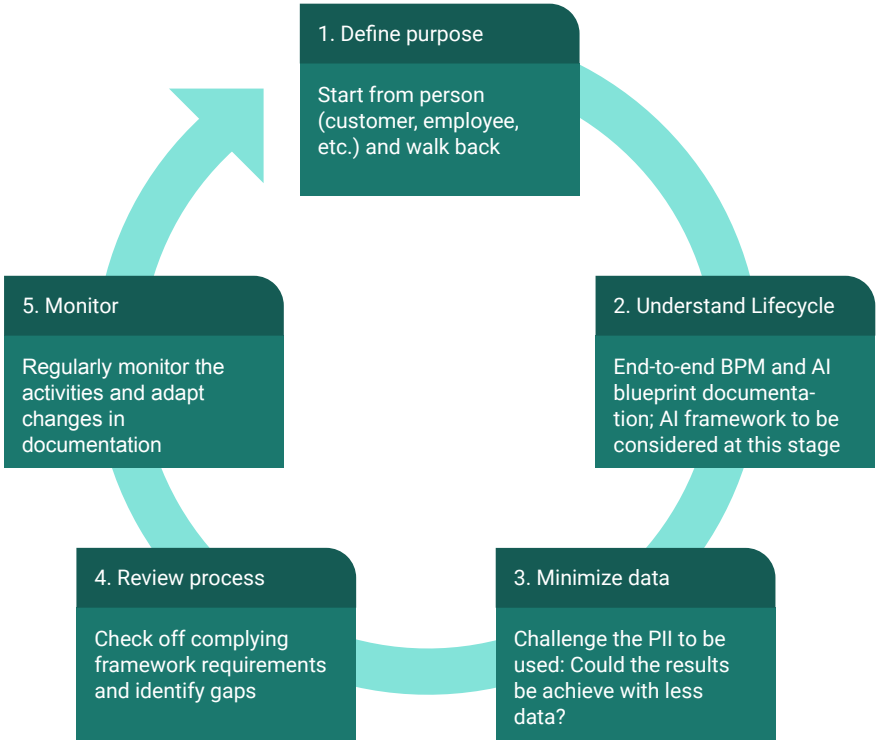
Key elements

- AI privacy design framework enhances an already existing DPMS
- Supporting privacy preserving design of AI models
- Foundation for guidelines and maturity assessments (audit function)
- Guarantees transparent processing throughout the data lifecycle, by design (Art. 25 GDPR)
- Each element is a standalone feature

Exemple: AI accountability principles

- Pursuant to Art. 5, 13, 14 and Recital 60 GDPR
- Fairness and transparency in profiling
- Accuracy (of used data)
- Data minimization and purpose limitation

5. Use Case: Fake Reviews

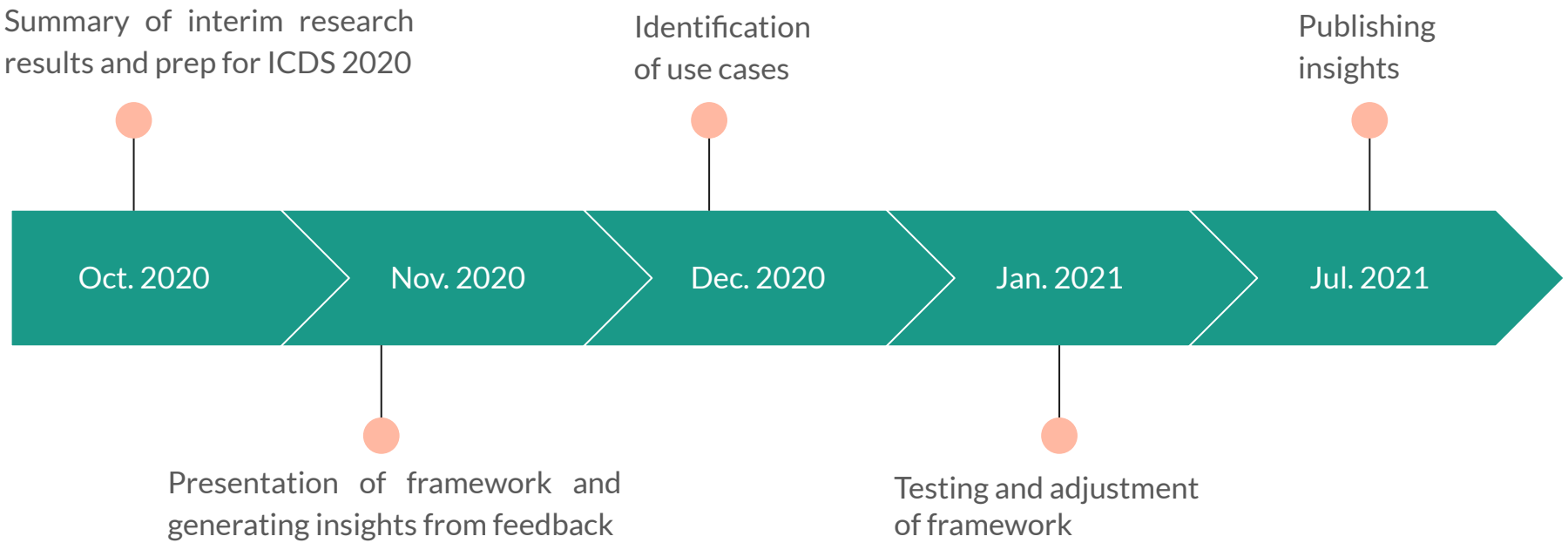


Takeaways:

Organizational challenges: Understanding and considering the entire data lifecycle in the AI compliance documentation goes beyond an isolated view on the algorithms functionality. Holistic view decreases the risk of non-compliance, as the entire data flow must be compliant.

Algorithmic pitfalls: Performance of an algorithm has immediate impact on the privacy. Overtraining or inherent discrimination, e.g. due to market-specific parameters, can lead to non-compliance with data protection requirements.

6. Project Milestones



Questions & Feedback: ICDS 2020

Submit your request **here**