HPI Hasso Plattner Institut
Digital Engineering · Universität Potsdam

Internationales Zentrum für Ethik in den Wissenschaften (IZEW)

Eberhard Karls Universität Tübingen

Responsible Data Science

Session 4: 17.05.2023, 15.15 – 19.30 h
MA Seminar, SoSe 2023, Hasso-Plattner Institut

Dr. Simon David Hirsbrunner

Today

topic	time
Orga and warm-up	15h15
Input: concepts in privacy research	15h30
Literature discussion: predictive privacy	15h45
—— Break ——	17h00
Input: value scenarios	17h15
Exercise: value scenarios	17h45
Input: identifying stakeholders and risks	18h15
Exercise: mapping stakeholders and risks	18h30
Sharing and discussing insights from groups	19h00
End	19h30

Orga

Next sessions

24.5. Data quality
25.5. People and planet + ?
1.6. Transparency and accountability

Paper preparation

Form groups until next week

When did you last experience an infliction of privacy?

-- Illustrating picture removed for copyright reasons ---

Privacy as a heterogenous concept

Privacy as "the right to be left alone" (Warren/Brandeis 1890).

Differentiation of various forms of privacy:

- Local privacy (≈ private sphere)
- Decisional privacy (≈ autonomy)
- Informational privacy (≈ data protection)

(see Rössler 2001)

Privacy as a heterogenous concept

Privacy as "the right to be left alone" (Warren/Brandeis 1890).

Differentiation of various forms of privacy:

- Local privacy (≈ private sphere)
- Decisional privacy (≈ autonomy)
- Informational privacy (≈ data protection)

(see Rössler 2001)

+ group / inferential / predictive privacy?
(related to data analytics and AI)

Literature discussion

Mühlhoff, R. (2023). Predictive privacy: Collective data protection in the context of artificial intelligence and big data. *Big Data & Society*, *10*(1). https://doi.org/10.1177/20539517231166886

Predictive privacy

- What is the article about?
- How is predictive privacy defined?
- How is it different from related concepts such as *group privacy* and *inferential privacy?*
- What are the major innovations of the concept?
- Do you think this concept has potential in informing the production of legal rules?
- Where might be the limits of a translation into legal rules? What weaknesses to do you see?

1 Exercice

Think about a situation where group belonging is co-constructed through technology.

2 Exercice

Think about a situation in which members of a group are not aware of pertaining to the group, while belonging to the group has real consequences for them.

2 Exercice

Think about a situation in which members of a group are not aware of pertaining to the group, while belonging to the group has real consequences for them.

"If men define situations as real, they are real in their consequences."
(Thomas and Thomas 1928)

Scenario-based design (I): definition

Scenario-based design is a family of techniques in which the **use of a future system is concretely described at an early point in the development process.** Narrative descriptions of envisioned usage episodes are then employed in a variety of ways to guide the development of the system that will enable these use experiences.

(Rosson and Carroll 2007: 1)
image: https://unitid.nl/2014/02/scenario-based-design/

Scenario-Based Design (II): example

S Martin is a police officer at the organized crime unit of the federal police. He currently investigates the selling of fake COVID-19 vaccination passports by an alleged criminal organization named *The Medics*. The Medics offer the counterfeit certificates to their S customers via the Telegram messenger. Unknown to Martin yet, S Chris, S Carlos, and S Eggert are Medics members, also communicating with their colleagues and suppliers via group Telegram channels while using pseudonyms, sometimes coded language, and images. In their free time, they also communicate with several friends, including their girlfriends, S Sarah and S Marta, who are unaware of their business. Martin's police unit gathers much information about The Medics using traditional investigative methods. This information leads to the identification of the suspect, Chris, who seems to be a low-level member of The Medics. On one evening, Chris is found with blank vaccination certificates during a traffic stop. He is arrested, and his phone is seized by investigator Martin, who aims at using the information on the phone to track down the individuals pulling the strings. After calling judge S Robert to get a search warrant, which is granted, he then searches Chris's unprotected phone, finds the Telegram communication, and extracts it. He recalls that his superior, S Dr. D, asked him to try out the new AutoCommAnalyzer software, which was recently purchased from the multinational company AI-Tech Corp. The software purchase was part of a strong push by the government to digitally optimize work processes at the police forces. Martin looks at the training notes by the head developer S Molly, trying to remember how the machine learning-driven software — trained with texts by S Alf and S Bert — is supposed to direct him to the relevant communication. The software presents him with the most frequent contacts, with Sarah on top. He reads through this communication, as the software has flagged several words like package and hospital, discovering some explicit images but finding that the

(Fischer MT, Hirsbrunner SD, Jentner W, et al. (2022) Promoting Ethical Awareness in Communication Analysis: Investigating Potentials and Limits of Visual Analytics for Intelligence Applications. In: *2022 ACM Conference on Fairness, Accountability, and Transparency*, New York, NY, USA, 20 June 2022, pp. 877–889. FAccT '22. Association for Computing Machinery. DOI: 10.1145/3531146.3533151)

Scenario-Based Design (III): qualities

- Scenarios are concrete but rough;
- Scenarios maintain an orientation to people and their needs;
- Scenarios are evocative, raising questions at many levels.

(Rosson and Carroll 2007)

Scenario-Based Design (IV): ingredients

- **Actors** (direct and indirect stakeholders)
- **Setting** (application context, situation of use)
- **Tools and objects** (technologies, interfaces)

(Rosson and Carroll 2007)

Value scenario

- Stakeholders
- Pervasiveness
- Time
- Systemic effects
- **Value implications**

Value scenario = VSD + Scenario-Based Design (SBD)

(Nathan et al. 2007)

3 Exercice

- Watch the provided video representing a scenario of a comprehensive AI system.
- While watching, try to identify the value scenario that is promoted through the video.
- We discuss these value representations in the group.

Stakeholders

- Whose values should be taken into account? **Stakeholder's values**
- What are stakeholders?
  - "A stakeholder is anyone who will be affected, directly or indirectly, by the new system like the end users, the software staff, and the organization's clients." (Shneiderman and Rose 1996: 92)
  - Stakeholders "can be people, groups, neighborhoods, communities, organizations, institutions, or societies, and can also include past and future generations, nonhuman species, and other elements such as historic buildings or sacred mountaintops" (Friedman and Hendry 2019: 37)

Focus on roles, not entities

- Stakeholders are defined by and understood in relationship to their interaction with a technology or sociotechnical system.

- They are considered by role, rather than by “person” or other “entity.”
  A “role” pertains to a stakeholder’s duties, contextual identity, or particular circumstances.

(Friedman and Hendry 2019: 37)

Methods to identify stakeholders

- Semi-structured interviews
- Participant observation
- Document analysis
- Scenario-Based Design
- Integrated Technology Development
- …

4 Exercice

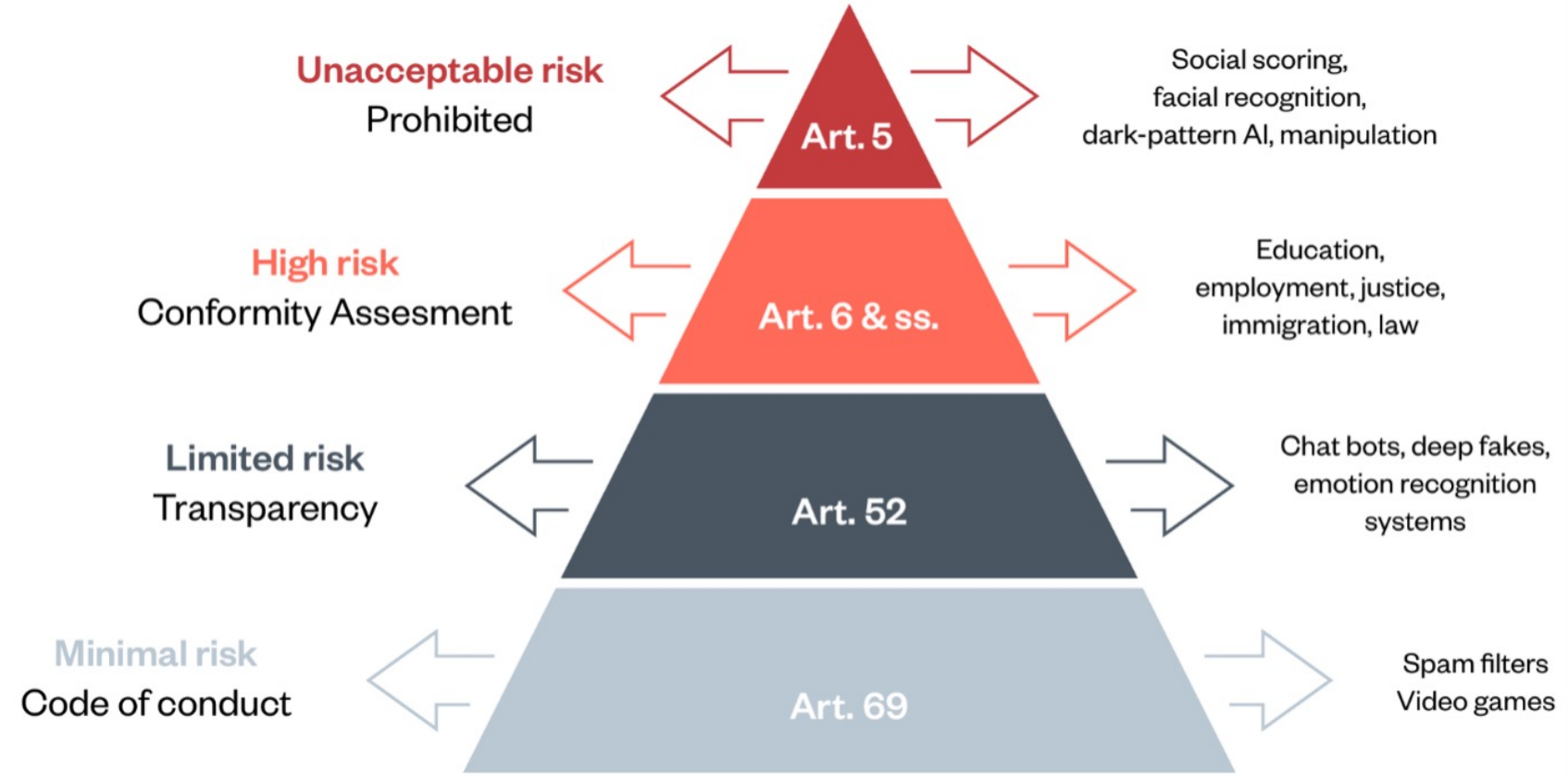
- Form groups of 2-3 people.
- Identify 5+ stakeholders (roles) in the provided scenario.
- Discuss at least one stakeholder with multiple roles.
- Create a stakeholder mapping, placing stakeholders according to their agency in the system.

AI risks for stakeholders

Definition of risk
„general probability of negative consequences to actions“ (see Cambridge Dictionary, 2022 in Lütge et al. 2022).

Algorithm-related risks for stakeholders
❏ Access to goods, benefits, or services ❏ Financial ❏ Property/material resources ❏ Reputation ❏ Emotional ❏ Life/security ❏ Privacy ❏ Liberty ❏ Rights/intellectual property (see City and County of San Francisco's Ethics and Algorithms Toolkit).

Risk-based approach of the (proposed) EU AI regulation



Source: https://www.adalovelaceinstitute.org/resource/eu-ai-act-explainer/

Tool
Stakeholder and risk mapping

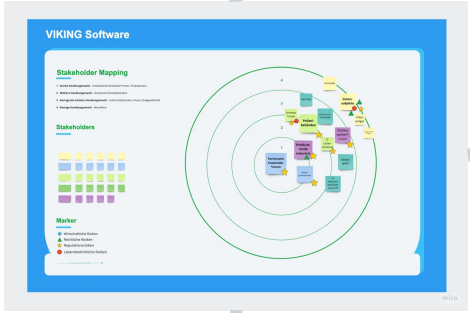
Grad der Handlungsmacht

1 - **Starke Handlungsmacht** - Entwickelnde (Entwickler*innen, Produzenten)

2 - **Mittlere Handlungsmacht** - Nutzende (Polizeibehörden)

3 - **Geringe bis mittlere Handlungsmacht** - Aufsichtsbehörden, Presse, Zivilgesellschaft

4 - **Geringe Handlungsmacht** - Betroffene



Risiken

- Wirtschaftliche Risiken
- Juristische Risiken
- Reputationsrisiken
- Lebensbedrohliche Risiken
- Grundrechtliche/ethische Risiken


4
3
2
1
Kriminelle
Gerichte
Daten-subjekte
Ermittler *innen
Datenschutz-beauftragte
Unbe-teiligte
Polizei-behörden
Wettbe werber* innen
Marginalisierte Gruppen
Polizeiliche IT-Unter-stützung
Produzie-rende Industrie
Zivil-gesellschaft
Technische Entwickler *innen
Gesetz-geber
Daten-produzenten
KI-Aufsichts-behörden (Zukunft)
miro

Risks for different stakeholder groups

In the event of faulty or otherwise problematic functioning,

Technical developers

- can lose their scientific / professional reputation;

Companies offering or selling the software or service,

- may be banned from a market;
- be sued and lose money and reputation

Users

- make mistakes that result in lost time, personal or material damage
- get sued or fired for using the software (incorrectly)

Data subjects

- are wrongly suspected of having committed a crime or intending to commit a crime in the future
- can be discriminated
- be subjected to threats and violence
- be unjustly persecuted, arrested, injured or killed

Institutional actors

- lose confidence in AI technology and stop supporting it (e.g., governments, trade associations, corporations, private and institutional funders, mass media)

(Civil) **society** actors are threatened in their exercise of fundamental rights.

Sources

See entire list of course references on Github: https://github.com/simonsimson/responsible-data-science/tree/main/slides