Responsible Data Science

Session 4: 22.05.2023, 13.30 – 16.45 h
MA Seminar, SoSe 2024, Hasso-Plattner Institute

Dr. Simon David Hirsbrunner

Today

topic	time
Orga and warm-up	13h30
Input from lecturer: concepts in privacy research	13h40
Input from students: group privacy and predictive privacy	13h45
Discussion	14h15
Break	15h00
Input: identifying stakeholders and risks	15h15
Exercise in groups: mapping stakeholders and risks	15h45
Sharing and discussing insights from groups	16h15
Assignments for next week and preview	16h40
End	16h45

When did you last experience an infliction of privacy?

-- Illustrating picture removed by the author for copyright reasons ---
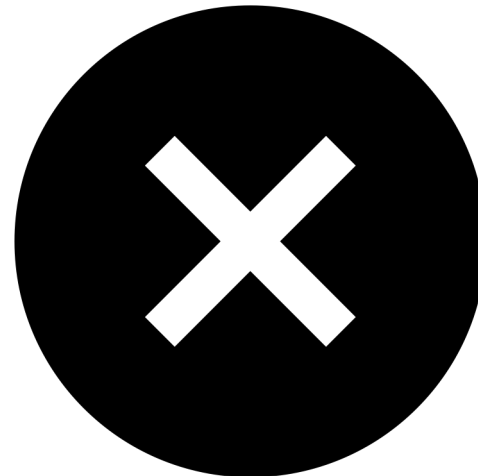
Privacy as a heterogenous concept

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)

Hasso Plattner Institut
Digital Engineering · Universität Potsdam

INTERNATIONALES ZENTRUM
FÜR ETHIK IN DEN
WISSENSCHAFTEN (IZEW)

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

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

Group privacy

Taylor L, Floridi L and van der Sloot B (2017) Introduction: A New Perspective on Privacy. In: Taylor L, Floridi L, and van der Sloot B (eds) Group Privacy: New Challenges of Data Technologies. Philosophical Studies Series. Cham: Springer International Publishing, pp. 1–12.

Predictive privacy

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

Questions

- What are algorithmic groups and why/how do they matter?
- How can we effectively ensure that users provide truly informed consent in the context of predictive analytics?
- How can we balance the interests of individuals in maintaining their privacy with the benefits of data-driven innovations for society?

Internationales Zentrum
für Ethik in den
Wissenschaften (IZEW)

Hasso
Plattner
Institut
Digital Engineering · Universität Potsdam

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

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

Internationales Zentrum
für Ethik in den
Wissenschaften (IZEW)

Eberhard Karls
UNIVERSITÄT
TÜBINGEN

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)
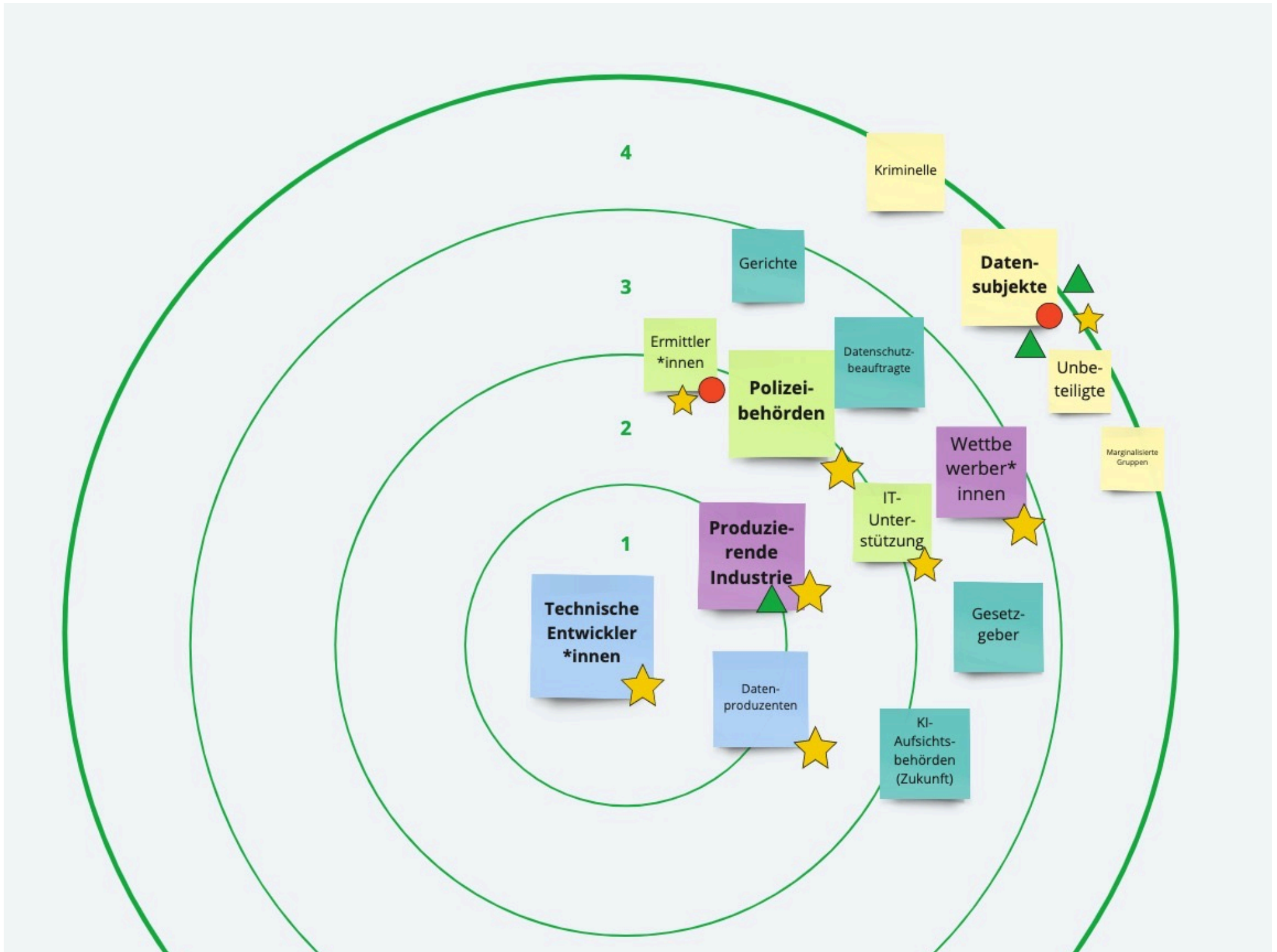
Focus on roles, not entities

- 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

- …

Stakeholder-Map (konzentrische Kreise 1–4):

Kreis 1:
- Technische Entwickler*innen ⭐
- Produzierende Industrie 🔺 ⭐

Kreis 2:
- Daten-produzenten ⭐
- IT-Unterstützung ⭐
- Wettbewerber*innen ⭐
- KI-Aufsichtsbehörden (Zukunft)
- Gesetzgeber
- Polizeibehörden ⭐ 🔴

Kreis 3:
- Ermittler*innen
- Datenschutzbeauftragte
- Gerichte
- Datensubjekte 🔺 🔴 ⭐ 🔺
- Unbeteiligte
- Marginalisierte Gruppen

Kreis 4:
- Kriminelle

2 Exercise: tasks

- Form groups.

- 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.

INTERNATIONALES ZENTRUM
FÜR ETHIK IN DEN
WISSENSCHAFTEN (IZEW)

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

2 Exercise: scenario

After two years of parental leave, Jana Kusturica applies for a job advertisement for a commercial assistant via an online application tool on a company website. In her opinion, the position matches her experience and qualifications very well.

Jana receives a rejection 2 minutes later, stating that other candidates were even better suited than her. She is not given any other reasons for the rejection. Due to the very fast response time of the recruiters and the generic reasons, Jana suspects that the pre-selection of incoming applications was automated by an AI. When Jana emails the recruiters, they confirm that an automated system based on AI was used.

Jana wonders on the basis of which variables this AI classified her profile as unsuitable for the position or did not even recommend a human assessment of the application documents or an invitation to a selection interview. When she asks the HR staff again, she receives the answer that unfortunately she cannot be given any information about the reasons for the rejection, but that the system called Personoptikum "has already been used successfully" and "makes objective decisions".

Jana now researches the software and finds the website of the company that sells Personoptikum. The advertising text for the software states that "the AI has been trained with countless CVs from job applicants in the USA" and is now "constantly being improved with new data". The software saves recruiters 70% of the normal processing time for finding a suitable candidate for a position, regardless of the field of application.

Jana writes an email to the company and asks them to give her an assessment of why the software has classified her as unsuitable for the position despite her qualifications. She explicitly asks whether the poor rating could be related to her Serbian name or her time off after the birth of her child. She is told that the software only calculates an overall rating and that individual decisions are "complex" and therefore cannot be easily broken down into one factor (such as parental leave). The software is also "blind to sensitive variables such as name, gender and origin". Jana asks whether she could recreate the automated selection process by loading her CV into the software herself. She is told that she would have to buy the software (which she cannot afford).

Jana is also rejected in countless other application processes. These involve very different job descriptions and qualification requirements. However, the software that rejects her so categorically is increasingly the same - Personoptikon.

Quelle: https://www.kitqar.com/de/szenarien

HPI Hasso Plattner Institut
Digital Engineering · Universität Potsdam

INTERNATIONALES ZENTRUM
FÜR ETHIK IN DEN
WISSENSCHAFTEN (IZEW)

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Reading for next week

Godulla A, Hoffmann C and Seibert D (2021) Dealing with deepfakes - An interdisciplinary examination of the state of research and implications for communication studies. Studies in Communication and Media 10. https://web.archive.org/web/20210402232954id_/https://www.nomos-elibrary.de/10.5771/2192-4007-2021-1-72.pdf

Sources

The entire bibilography for the course can be found on Github here:
https://github.com/simonsimson/responsible-data-science/blob/main/slides/Bibliography-of-the-entire-course.pdf

Image sources
Most sources are cited on the relevant slide. Slide 1: © Adobe Stock / kras99