Informationsfreiheit by Design

Eine Einführung

Benjamin Walczak

Sommerakademie 2019

„Verbraucher im Fokus“

am 9. September 2019 in Kiel

ULD

Unabhängiges Landeszentrum für
Datenschutz Schleswig-Holstein

Überblick

1. Motivation / IFK-Position
2. Verständnis von „by Design“
3. Ansätze der Informationsfreiheit
4. Zugänglichkeit
5. Einschränkungen
6. Kategorisierung im Daten-Lebenszyklus
   Exkurs: Differential Privacy
7. Methoden(-bedarf) für Informationsfreiheit by Design
8. Versuch einer Systematisierung

Motivation

Warum Informationsfreiheit?

„ Jeder hat das Recht … sich aus allgemein zugänglichen Quellen ungehindert zu unterrichten.“
aus Art. 5 GG

BVerfG
03.10.1969

„Die Informationsfreiheit steht in der grundgesetzlichen Ordnung gleichwertig neben der Meinungs- und Pressefreiheit.“

„Die Behörden des Landes, der Gemeinden und Gemeindeverbände stellen amtliche Informationen zur Verfügung, soweit nicht entgegen-stehende öffentliche oderschutzwürdige private Interessen über-wiegen. Das Nähere regelt ein Gesetz.“
Artikel 53 Landesverfassung SH

Motivation / IFK-Position

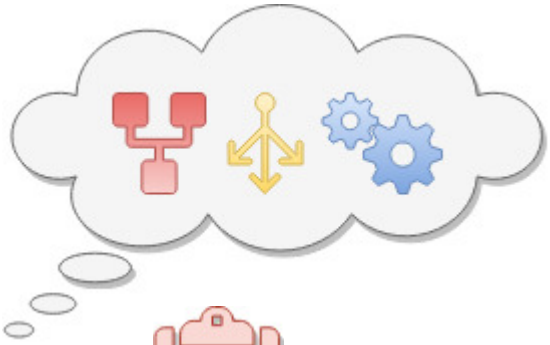
Hintergrund:

- Digitaler Wandel
  - E-Government-Gesetze
  - Online-Zugangsgesetz
- Mehr Interesse an Transparenz im Verwaltungshandeln
- Mehr Auskünfte
- OpenData

- Beschluss: 12. Juni 2019
- Definition
- Rechtlicher Rahmen
- mögliche Maßnahmen
- adressiert an Verantwortliche und Entwickelnde

Verständnis von „by Design“

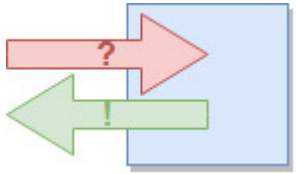
- Abstraktion
  - Generalisierung
  - Kategorisierung
  - Systematisierung
- Anforderungen
- Muster für wiederkehrende Fragestellungen
- Erkennen und ggf. Lösen von Zielkonflikten
- Entwicklungen erkennen



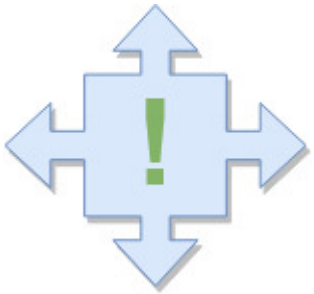
- Einzelfallprüfungen weiterhin nötig
- Verschiedene Designs für verschiedene Szenarien

Ansätze der Informationsfreiheit

1. Anlassbezogener Informationszugang
   z.B. Anfrage an eine Behörde



2. Proaktive Veröffentlichung (Transparenz)
   z.B. auf der Homepage oder in einem Transparenz-Portal



Zugänglichkeit

Barrierefreiheit

z.B. vorlesbar, vergrößerbar, ...

Maschinenlesbarkeit

z.B. als XML

ggf. bei IZG-Anfragen nicht gewollt

Schnittstelle zu Transparenzportalen

staatliche und nicht-staatliche

Auskunftsmöglichkeiten beim Einsatz von künstlicher Intelligenz

Einschränkungen

- Ausschlussgründe
  - Öffentliche Interessen (§ 9 IZG SH)
  - Private Interessen (§ 10 IZG SH), bspw.
    - Geschäftsgeheimnis
    - Schutz personenbezogener Daten

Wie kann vermieden werden, dass Ausschlussgründe für nur einen Teil der Unterlagen auf die gesamten Unterlagen ausgedehnt werden?

Kategorisierung im Daten-Lebenszyklus

Anlegen von Akten

- Trennung verschiedener Verfahren in mehrere Akten
- Durchsuchbar
- Strukturiert für eine Freigabe (Abtrennbarkeit)
- Barrierefreiheit / Verständlichkeit
- Sachlichkeit

Aktenführung / Auskünfte

- Speichern von Anfragen und Antworten
- Verkettbarkeit von Auskünften prüfen
- Meta-Daten bereinigen
  (z.B. Dateinamen, Versionshistorie, Foto-Daten)

Kategorisierung im Daten-Lebenszyklus

Datenbanken

- Abfragemöglichkeiten
- Anonymisierung (→ Differential Privacy)

Löschen / Korrekturen

- ggf. Weitergabe an Portale
- ggf. auch in Antworten auf Anfragen

Archiv

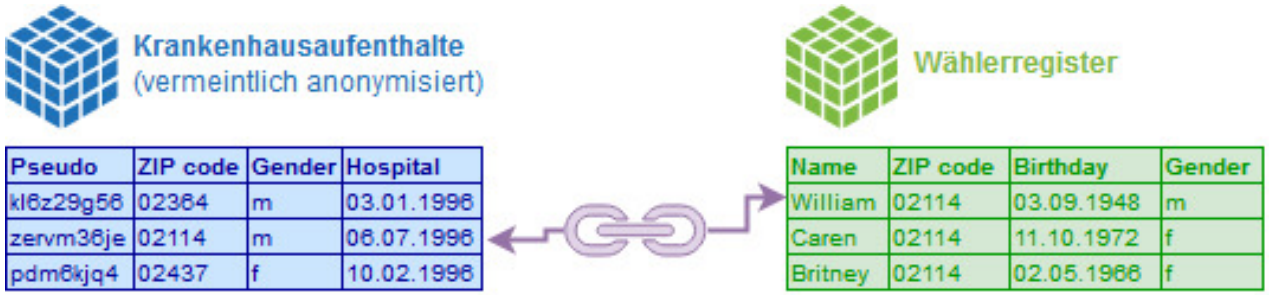
- Auffindbarkeit / Strukturierung
- Erhalt der Auskunftmöglichkeit

Differential Privacy

Anonymisieren ist nicht immer anonym ...

Beispiel Massachussetts:

- Veröffentlichung von Krankenhausaufenthalten (anonym?)
- Gouverneur William konnte identifiziert werden
- 87% der US-amerikanischen Bevölkerung sind eindeutig identifizierbar anhand:
  - Geburtsdatum
  - ZIP Code
  - Geschlecht



Pseudo	ZIP code	Gender	Hospital
kl6z29g56	02364	m	03.01.1996
zervm36je	02114	m	06.07.1996
pdm6kjq4	02437	f	10.02.1996

Name	ZIP code	Birthday	Gender
William	02114	03.09.1948	m
Caren	02114	11.10.1972	f
Britney	02114	02.05.1966	f

Differential Privacy

Jede Freigabe von Daten aus einer statistischen Datenbank kann ein Datenschutz-Risiko sein.

Lösungsansätze:

„Kaugummi unter den Tisch geklebt?“

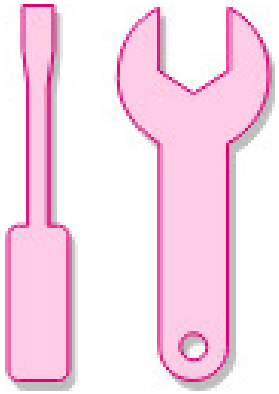
← Randomisierte Antworten

- Zufallsrauschen und Dummy-Einträge
- Limit für Datenfreigaben

Ziel:

Es ist nicht ermittelbar, ob der Datensatz einer bestimmten Person enthalten ist oder nicht.

Methoden(-bedarf)



- Organisatorische Vorgaben ☐
- Barrierefreiheit/Maschinenlesbarkeit (☐)
- Anforderungen an Künstliche Intelligenz
- Abtrennbarkeit gewährleisten
- Trennen von personenbezogenen Daten (z.B. Formular)
- Speichern von Anfragen/Antworten
- Anonymisierung/Verkettbarkeit von Antworten prüfen

Versuch einer Systematisierung

Zugang klären

- Transparenz oder auf Anfrage?
- Barrierefreiheit/ Maschinenlesbarkeit
- Weitergabe an Portale

Auswahl treffen

- Trennen
- Ausschlussgründe?
- Risiko abschätzen
- Anonymisierung?

Grundsätzliche Überlegungen ersetzen nicht Einzelfallprüfung!

Vorbereitung | Praxis

Management

- Durchsetzen von Standards
- Abfragemöglichkeiten aktualisieren
- Maßnahmen treffen, um Risiko zu minimieren
- (Fremd-)Archivierung organisieren

Kontrolle

- Gleiche Anfrage, gleiche Antwort
- Anforderungen an KI-Einsatz
- Verkettbarkeit prüfen
- Veröffentlichung über Portale kontrollieren

Vielen Dank für die Aufmerksamkeit

… und wir freuen uns auf die Diskussion!