Open Data und Open Government

Sven Thomsen

ULD
Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein

Übersicht

- Open Data und Open Government
- IFG / UIG
- Thesen zum Umsetzungsstand
- Was ist zu tun? (aus Datenschutzsicht)
- Fazit

Open (Government) Data

- 8 Prinzipien, 2007 erstmalig zusammengestellt im „Sebastopol-RFC“ (Sebastopol, Kalifornien)
- Anforderungen, denen „offene Daten“ genügen sollten:
  - **Complete** (vollständige Daten)
  - **Primary** (aus Primärquellen, Ursprungsdaten)
  - **Timely** (zeitnah und aktuell)
  - **Accessible** (zugänglich)
  - **Machine processable** („maschinenlesbar“)
  - **Non-discriminatory** (Zugriff für jeden ohne Hindernisse)
  - **Non-proprietary** (offene Dateiformate)
  - **License-free** (Urheberrecht, Patente, NDAs etc.)

1: Complete

- Vollständigkeit
  - Alle öffentlichen Daten sollen veröffentlicht werden
  - Öffentliche Daten sind alle Daten der öffentlichen Verwaltung
  - Ausnahmen durch gesetzliche Einschränkungen
    - Datenschutzgesetze
    - Sicherheitserwägungen
  - Begründung: Daten der öffentlichen Verwaltung sind durch Abgaben und Entgelte durch Bürger bereits finanziert.
  - Begründung: Daten sind vorhanden. Geringe Kosten für zusätzliche Veröffentlichung bei potentiell hohem Nutzen.

2: Primary

- Primärdaten, Daten aus Primärquellen, Ursprungsdaten
- Daten sollen im Ursprungsformat/Quellformat bereitgestellt werden
- Möglichst hoher Detaillierungsgrad, keine aggregierten, modifizierten oder abgeleiteten Daten
  - Unkritisches Beispiel: Videoaufzeichnungen in HD vs. „Pixelkino“ im Netz
  - Diskussion:
    - Rohdaten zur Luftverschmutzung, geokodiert, Zeitreihen
    - Art und Anzahl von HIV-Erkrankungen, geokodiert auf 1km*1km, Zeitreihen (lesenswert: Infektionsschutzgesetz)

3: Timely

- Zeitnahes Bereitstellen
- Wert der Daten hängt häufig direkt vom Alter ab
- Werterhaltung bedeutet zeitnahe Veröffentlichung
- Alter der Daten muss (neben anderen Qualitätsmerkmalen) erkennbar sein

4: Accessible

- Zugänglich, zugreifbar
- Möglichst viele Nutzungsszenarien
- Möglichst viele Nutzerinnen und Nutzer
- Automatisiert abrufbar

5: Machine processable

- Maschinenlesbar, automatisiert verarbeitbar
- Strukturierte Datenhaltung

6: Non-discriminatory

- Zugriff für jeden ohne Hindernisse
- Keine Segmentierung der Nutzer
- Keine Beschränkung auf einzelne Gruppen
- Möglichst keine Barrieren wie Registrierung, geschlossene Nutzergruppen, Zahlsysteme

7: Non-proprietary

- Offene Dateiformate
- „De-facto-Standards“, „Industriestandards“ sind häufig keine offenen Dateiformate in diesem Sinn
- Definition: „Formate, über die keine Entität die alleinige Kontrolle hat“

8: License-free

- Lizenzfrei
- Keine Einschränkungen durch Urheberrechte, Patente, Markenrechte, Non-Disclosure-Agreements, ...

Open Data

- Die Übereinstimmung mit den 8 Prinzipien muss kontrollierbar sein
  - Benennung einer Kontaktperson
  - Beschwerdemanagement / Vorschlagswesen
  - Fachaufsicht über Anwendung der Open-Data-Prinzipien

Open Government

- Unterschiedliche Begrifflichkeit:
  - (Bekannte) Ableitung aus Informationsfreiheit
  - (Neue) Ableitung aus Open-Data-Prinzipien
- (Neue Ableitung) 10 Prinzipien
  - Completeness
  - Primacy
  - Timeliness
  - Ease of Physical and Electronic Access
  - Machine Readability
  - Non-discrimination
  - Use of Commonly Owned Standards
  - Licensing
  - Permanence
  - Usage Costs

IFG – Informationsfreiheitsgesetz

- Auf Bundesebene: seit 2005
- Grundsatz: „Jeder hat nach Maßgabe dieses Gesetzes gegenüber Behörden des Bundes einen Anspruch auf Zugang zu amtlichen Informationen.“
- „amtliche Information: jede amtlichen Zwecken dienende Aufzeichnung, unabhängig von derart ihrer Speicherung. Entwürfe oder Notizen, die nicht Bestandteil eines Vorgangs werden sollen, gehören nicht dazu.“

IFG – Informationsfreiheitsgesetz

- Detaillierte Aufzählung mit Einschränkungen/Ausnahmen
  - Gefahr für internationale Beziehungen
  - Auskunft über militärische/sonstige Belange der Bundeswehr
  - Belange der inneren oder äußeren Sicherheit
  - ...
  - ...
  - Verschlusssachen
  - Durchführung laufender Gerichtsverfahren
  - .... usw.

IFG – Informationsfreiheitsgesetz

- Auskünfte können mündlich, schriftlich oder **elektronisch** erteilt werden.
- Informationszugang in der Regel innerhalb eines Monats
- Gebühren können erhoben werden, Verwaltungsaufwand ist zu berücksichtigen

- Achtung: deutlich höhere Qualität für den Anfragenden als bei „Open Data“-Ansätzen
- Akteneinsicht, komplette Übersicht über einen Verwaltungsprozess, nicht nur Nutzung von Daten öffentlicher Stellen

UIG – Umweltinformationsgesetz

- Auf Bundesebene seit 2004
- „Zweck dieses Gesetzes ist es, den rechtlichen Rahmen für den freien Zugang zu Umweltinformationen bei informationspflichtigen Stellen sowie für die Verbreitung dieser Umweltinformationen zu schaffen.“
- Umfangreiche Darstellung, was Umweltinformationen sind:
  - Zustand von Luft und Atmosphäre, Wasser, Boden, Landschaft und natürlichen Lebensräumen ...
  - Faktoren wie Stoffe, Energie, Lärm und Strahlung, Abfälle aller Art sowie Emissionen ...
  - Maßnahmen oder Tätigkeiten ...
  - Kosten-Nutzen-Analysen oder sonstige wirtschaftliche Analysen ...

UIG – Umweltinformationsgesetz

- Detaillierte Liste von Ablehnungsgründen
- Aber auch: detaillierte Vorgaben zur Zugangserleichterung

§ 7 Unterstützung des Zugangs zu Umweltinformationen

(1) Die informationspflichtigen Stellen ergreifen Maßnahmen, um den Zugang zu den bei ihnen verfügbaren Umweltinformationen zu erleichtern. Zu diesem Zweck wirken sie darauf hin, dass Umweltinformationen, über die sie verfügen, zunehmend in elektronischen Datenbanken oder in sonstigen Formaten gespeichert werden, die über Mittel der elektronischen Kommunikation abrufbar sind.

(2) Die informationspflichtigen Stellen treffen praktische Vorkehrungen zur Erleichterung des Informationszugangs, beispielsweise durch

1. die Benennung von Auskunftspersonen oder Informationsstellen,
2. die Veröffentlichung von Verzeichnissen über verfügbare Umweltinformationen,
3. die Einrichtung öffentlich zugänglicher Informationsnetze und Datenbanken oder
4. die Veröffentlichung von Informationen über behördliche Zuständigkeiten.

(3) Soweit möglich, gewährleisten die informationspflichtigen Stellen, dass alle Umweltinformationen, die von ihnen oder für sie zusammengestellt werden, auf dem gegenwärtigen Stand, exakt und vergleichbar sind.

Sachstand

- IFG und UIG: Bereitstellung „auf Antrag“
  - Abwägung im Einzelfall
  - Informationszugang und -bereitstellung im Einzelfall
- Open Data und Open Government: Bereitstellung „generell“
  - Keine Einzelfallbetrachtung, sondern generelle Freigabe
- IFG und UIG bieten rechtliche Grundlage für Open-Data-Zugang
  - Abkehr vom Einzelfallprinzip nötig („schlanker Staat“, „Wirtschaftlichkeit“?)
  - Notwendig: Treffen von Grundsatzentscheidungen für Datenarten, bestimmte Datentypen, Szenarien, Kontexte, ...
- Warum gibt es nicht mehr „offene öffentliche Daten“?
  - Mangelndes Interesse bei Nachfragern?
  - Angst bei Anbietern?
  - Rechtliche Hemmnisse?

Zwei Thesen zur Diskussion ...

These 1

- Öffentliche Verwaltung hat keine ausreichende Erfahrung
  - zu Risikoanalysen („Was kann passieren?“)
  - zur Risikobewertung („Wie wahrscheinlich?“)
  - zur Risikobehandlung („Was machen wir?“)
  - zum Umgang mit und zur Übernahme von Restrisiken
  - zu kontinuierlichen Verbesserungsprozessen
- Wenn Risiken unklar sind, der Umgang ungeplant ist und keine dauerhafte Befassung mit Risiken stattfindet, dann wird Risikovermeidung durch „Nichts-Tun“ attraktiv ...
- Verwaltungen haben Angst, Informationen bereitzustellen

These 2

- Open Data ist fokussiert auf „Veröffentlichen“, nicht auf „Verarbeiten“

- Anwender haben aktuell keine Mechanismen ...
  - zum Qualitätsmanagement
  - zur automatisierten Prüfung der Qualität von „Mashups“
  - zur nachvollziehbaren Dokumentation von Datenverarbeitungsschritten

Was ist zu tun?

Was ist zu tun? Risikoanalyse vorab

- Blick auf Stand der IT-Sicherheit:
  - Abkehr von **maßnahmenorientierten** Vorgehensweisen (Kataloge, Checklisten)
  - Hinwendung zu **prozessorientierten** Vorgehensweisen (Risikoanalyse und -bewertung, integriertes Sicherheits- und Datenschutzmanagement)
- Open Data und Open Government bedingen eine **vorausgehende Risikoanalyse** und das Aufstellen eines Risikobehandlungsplans
  - Datenschutz: Nicht-Verkettbarkeit
  - Risikoanalyse: Verkettung von Daten mit eindeutigen Identifiern, Personenbezug durch Korrelation und Analyse

Was ist zu tun?
Kontinuierliche Risikoanalyse

- Bei Abkehr von Einzelfallprinzip: erhöhte Risiken
  - keine kontrollierte Evaluationssituation bei Freigabe
  - andere Stellen können Daten freigeben, die das Risiko der Verkettbarkeit steigern
- Veröffentlichende Stelle muss Prozesse einrichten zur ...
  - kontinuierlichen Beobachtung des „Open Data“-Markts
  - kontinuierlichen Risikobewertung in Bezug auf die veröffentlichten Daten
  - dauerhaften Simulation von „Angriffen“ auf die Nicht-Verkettbarkeit: Zeitreihen, Hinzufügen von Kontext, ...
  - kontinuierlichen Beobachtung der Verwendung der veröffentlichten Daten

Was ist zu tun?
„Management von Datenschutzvorfällen“

- Betroffenenrechte:
  - Aufklärung, Auskunft
  - Berichtigen
  - Löschen, Sperren
- Wie kontrolliert man veröffentlichte Daten?
  - Veröffentlichen vs. Erteilen von Nutzungsrechten?
- „Digitales Management von Betroffenenrechten“
- „DRM“ ist verbrannte Erde und ist in aktueller Ausprägung nicht akzeptabel (siehe Studie „Privacy4DRM“)
- Aber: Für risikoreiche Veröffentlichungen personenbezogener Daten wird ein Mechanismus zum „Remote-Wipe“ und „Remote-Fix“ benötigt

Was ist zu tun?
„Digitaler Verwendungsnachweis“

- Probleme:
  - Qualität und Aussagekraft der Ausgangsdaten bewerten
  - Qualität und Aussagekraft der Datenverarbeitung bewerten
- Beispiel: Fehler bei der Verwendung von Geodaten
  - Unterschiedliche Maßstäbe / Auflösungen
  - Überschneidung und Doppelerfassung durch unterschiedliche Referenzmodelle und Flächenschnitte
- Verarbeitung von Open Data muss nachvollziehbar werden
  - Welche Datenquellen?
  - Welche Schritte zur Aggregation, Korrelation, ...

Was ist zu tun?
„Digitaler Verwendungsnachweis“

- Anleihe:

Was ist zu tun? „Digitaler Verwendungsnachweis“

- „Open Data“ funktioniert nur mit „**Open Processing**“
  - Gleiche Transparenz- und Verfügbarkeits-Prinzipien nicht nur auf Daten, sondern auf **Datenverarbeitung** anwenden
  - Herkunft der Daten nachweisen
  - Verarbeitung revisionsfähig ausgestalten und vor allem: dokumentieren
- Dokumentation automatisiert auswertbar gestalten: „Zu jedem Ergebnis einen XML-Nachweis der Verarbeitung und Quellen“
- Empfehlung: neuer Codex in der Open Data Community:
  - **„Wer mit Hilfe von Open Data etwas nachweisen möchte, muss die Korrektheit seiner Datenverarbeitung automatisiert nachvollziehbar nachweisen.“**
- Vorerst: Schriftliche Darstellung der Datenverarbeitung zusammen mit den Daten veröffentlichen

Fazit, Ausblick und Prognose

Fazit

- Der Fokus von Open Data und Open Government liegt (zur Zeit) auf „**Veröffentlichen**“
- Datenschutz betrachtet „**Verarbeiten**“: gesamte Vearbeitungskette über längeren Zeitraum
- Teilweise Abkehr vom Einzelfallprinzip benötigt deutlich besseres **Know-How** der öffentlichen Verwaltung im **Risikomanagement**
- Die erhoffte Transparenz und Wertschöpfung von „Open Data“ und „Open Government“ ist nur möglich mit „**Open Processing**“

Ausblick und Prognose

- Benennung von **„Open Data“-Beauftragten**
  - Beispiel: UK, USA
  - Fortbildung im Bereich der Risikoanalyse insb. mit Fokus auf Datenschutz und Datensicherheit notwendig
- Trend zum **„generellen“ Veröffentlichen**
  - Hohe Prozesskosten bei Einzelfallentscheidungen
  - Stetig steigende Nachfrage nach Daten der öffentlichen Verwaltung

Diskussion

Vielen Dank!

Sven Thomsen
Unabhängiges Landeszentrum für Datenschutz
Holstenstraße 98
24103 Kiel

Tel: 0431 – 988 1211
Mail: ULD3@datenschutzzentrum.de