

Präsentation Metadaten

Open Data Support wird von der Europäischen Kommission finanziert, gemäß SMART 2012/0107 'Lot 2: Provision of services for the Publication, Access and Reuse of Open Public Data across the European Union, through existing open data portals' (Vertrag No. 30-CE-0530965/00-17).

© 2013 European Commission



**OPEN
DATA
SUPPORT**

Trainingsmodul 2.2

Die Qualität von offenen Daten & Metadaten

Lernziele

Am Ende dieses Trainingsmoduls sollten Sie:

- wissen, was (offene) Datenqualität bedeutet.
- die Faktoren und Kritiken für die Qualität von Open Data kennen.
- Good Practices für die Veröffentlichung von hochwertigen (Linked) Open Data kennen.

Inhalt

Dies Modul enthält...

- Eine Definition von Datenqualität;
- Eine Übersicht über die Dimensionen von Daten- und Metadaten-Qualität;
- Eine Auswahl von Best-Practice-Beispielen für die Veröffentlichung von hochwertigen Daten und Metadaten.

Was ist Daten (und Metadaten) Qualität?

Daten sind dann von hoher Qualität, “wenn Sie für ihren vorgesehenen Gebrauch im operativen Geschäft, bei Entscheidungen und bei der Planung geeignet sind.”

Oder genauer gesagt:

“Daten von hoher Qualität sind genau, verfügbar, vollständig, konform, konsistent, glaubwürdig, verarbeitbar, relevant und zeitgemäß.”

Metadaten sind Daten über Daten...

“Metadaten sind strukturierte Informationen, die beschreiben, erklären, lokalisieren, oder es sonstwie einfacher machen, eine Informationsquelle abzurufen, zu verwenden, oder zu verwalten. Metadaten werden oft Daten über Daten genannt.”

-- National Information Standards Organization

- Wir beobachten, dass Metadaten eine Art von Daten sind.
- Die gleichen Überlegungen gelten gleichermaßen für die Qualität von Daten und von Metadaten.

Die Dimensionen der Datenqualität

Was sind die wichtigsten Dimensionen, die in Betracht gezogen werden müssen, um eine gute Qualität von (Meta-)Daten zu liefern?

Die Dimensionen der Datenqualität

- **Genauigkeit:** Stellen Sie die Daten Objekte oder Ereignisse aus der realen Welt richtig dar?
- **Konsistenz:** Enthalten die Daten keine Widersprüche?
- **Verfügbarkeit:** Können die Daten jetzt und auf Dauer abgerufen werden?
- **Vollständigkeit:** Umfassen die Daten alle Datenelemente, die ein Objekt oder Ereignis darstellen?
- **Konformität:** Folgen die Daten anerkannten Standards?
- **Glaubwürdigkeit:** Basieren die Daten auf vertrauenswürdigen Quellen?
- **Verarbeitung:** Sind die Daten maschinenlesbar?
- **Relevanz :** Umfassen die Daten eine angemessene Menge an Daten?
- **Rechtzeitigkeit:** Stellen die Daten die derzeitige Situation dar und werden sie früh genug veröffentlicht?

Genauigkeit

Die Genauigkeit von Daten ist das Maß, in dem diese die Eigenschaften der realen Welt Objekt, Situation oder Ereignis richtig darstellen.

Zum Beispiel:

- Richtige Messung von Witterungsverhältnissen (Temperatur, Niederschlag).
- Richtiger Hinweis auf die Bedingungen zur Wiederverwendung des Datensatzes.

Empfehlungen:

- **Wägen** Sie die **Genauigkeit** Ihrer Daten gegen die **Kosten** im Kontext der Anwendung ab; sie müssen **für die vorgesehenen Verwendung gut genug** sein.
- Stellen Sie sicher, dass es ein **organisatorisches Engagement** und eine **Investition in Verfahren und Werkzeugen** gibt, damit die Genauigkeit Ihrer Daten aufrecht erhalten werden kann.

Ein Beispiel für Genauigkeit

Hohe Genauigkeit

Weniger Genauigkeit

OpenStreetMap, Stadt von Utrecht,
Niederlande (2011 vs. 2007)



Konsistenz

Die Konsistenz der Daten ist das Ausmaß, in dem sie keine Widersprüche enthalten, die ihre Verwendung schwierig oder unmöglich machen würde.

Zum Beispiel:

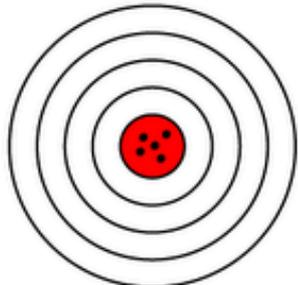
- Ein Datensatz, der Daten aus verschiedenen Quellen verbindet, die verarbeitet wurden, um widersprüchliche Aussagen aufzudecken, die gelöst worden sind.
- Eine Beschreibung eines Datensatzes, der keine unterschiedlichen Lizaussagen enthält, oder bei dem die Daten der letzten Änderung nicht vor dem Erstellungsdatum liegen.

Empfehlungen:

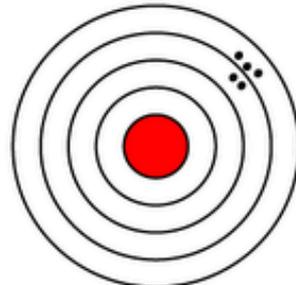
- **Bearbeiten Sie alle Daten vor ihrer Veröffentlichung**, um widersprüchliche Aussagen und andere Fehler zu entdecken (insbesondere, wenn Daten aus verschiedenen Quellen zusammengefügt wurden).

Konsistenz : Beispiel

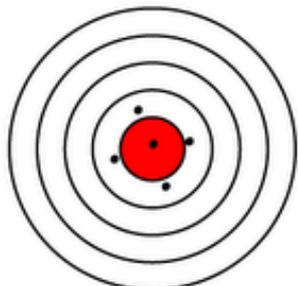
Hohe Konsistenz



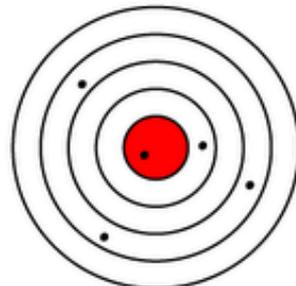
Accurate & Consistent



Not Accurate
but Consistent



Accurate but
Not Consistent



Not Accurate &
Not Consistent

```
:weather1-7 a dcat:Dataset ;
dct:title "Measurements from weather stations 1-7" ;
dct:description "Data from seven weather stations
                showing temparture, humidity,
                wind direction and wind speed" ;
dct:issued "2013-01-01T00:00:00+01:00" ;
dct:modified "2013-07-01T19:20:30+01:00" ;
dct:publisher <http://myweather.com/id/myweather> ;
dcat:keyword "weather" ;
dcat:landingpage <http://myweather.com/stations1-7.html> ;
dcat:distribution :weatherdata-xlsx
.

:weatherdata1-7-xlsx a dcat:Distribution ;
dct:format <http://publications.europa.eu/resource/authority/file-type/XLSX> ;
dct:licence <http://creativecommons.org/licenses/CC0> ;|
dcat:downloadURL <http://myweather.com/stations1-7.xlsx>
.
```

Geringere Konsistenz

```
:weather1-7 a dcat:Dataset ;
dct:title "Measurements from weather stations 1-7" ;|
dct:description "Data from seven weather stations
                showing temparture, humidity,
                wind direction and wind speed" ;
dct:issued "2014-01-01T00:00:00+01:00" ;
dct:modified "2013-07-01T19:20:30+01:00" ;
dct:publisher <http://myweather.com/id/myweather> ;
dcat:keyword "weather" ;
dcat:landingpage <http://myweather.com/stations1-7.html> ;
dcat:distribution :weatherdata-xlsx
.
```

ERROR INCONSISTENT DATA: Issue date is after modification date

```
:weatherdata1-7-xlsx a dcat:Distribution ;
dct:format <http://publications.europa.eu/resource/authority/file-type/XLSX> ;
dct:licence <http://creativecommons.org/licenses/CC0> ;
dct:licence <http://creativecommons.org/licenses/BY/3.0> ;
dcat:downloadURL <http://myweather.com/stations1-7.xlsx>
.
```

ERROR INCONSISTENT DATA: Licence element repeated

Verfügbarkeit

Die Verfügbarkeit von Daten ist das Maß, in dem auf sie zugegriffen werden kann; dies schließt auch die Langlebigkeit von Daten ein.

Zum Beispiel:

- Ein Datensatz, der von einem http identifiziert wird: URI, die kontinuierlich auf die richtige Ressource hinweist (und nicht 404 Not Found zurückgibt),.
- Eine Beschreibung der Datensätze, die in der Suchmaschine von einem Datenportal enthalten ist.

Empfehlungen:

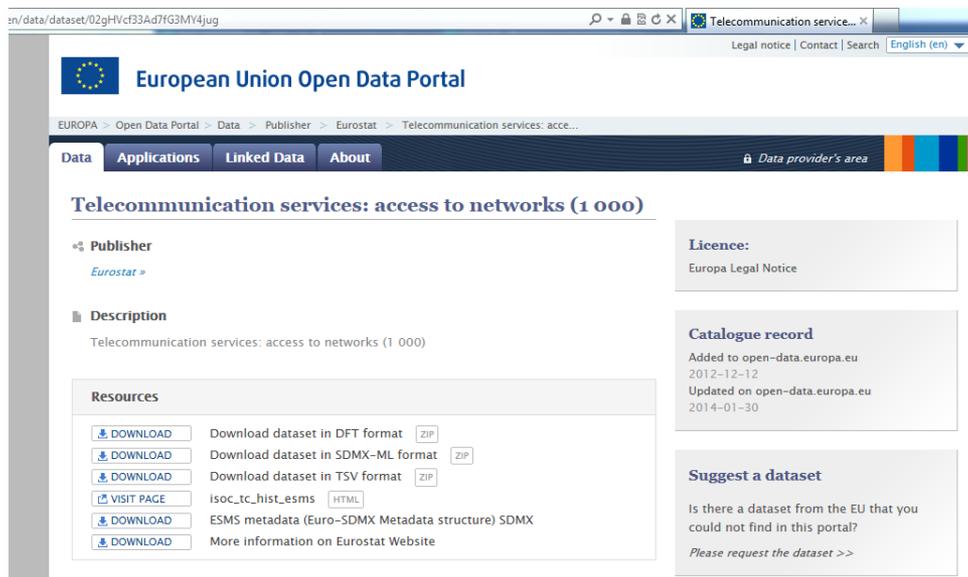
- Orientieren Sie sich bei der Zuordnung und Pflege von URIs an Best Practices.
- Stellen Sie sicher, dass in der Organisation **die Verantwortung** für die Pflege von Daten **eindeutig zugewissen** wird.

Siehe auch:

<http://www.slideshare.net/OpenDataSupport/de/sign-and-manage-persitent-uris>

Verfügbarkeit: Beispiel

Hohe Verfügbarkeit



The screenshot shows a web browser window displaying the European Union Open Data Portal. The page title is "Telecommunication services: access to networks (1 000)". The page is structured with a header, a navigation menu, and several content sections:

- Publisher:** Eurostat
- Description:** Telecommunication services: access to networks (1 000)
- Resources:** A list of download options for the dataset in various formats (DFT, SDMX-ML, TSV) and a link to the dataset's metadata on the Eurostat website.
- Licence:** Europa Legal Notice
- Catalogue record:** Added to open-data.europa.eu on 2012-12-12 and updated on 2014-01-30.
- Suggest a dataset:** A section asking if there is a dataset from the EU that could not be found in this portal, with a link to request the dataset.

Geringe Verfügbarkeit

The page cannot be found

The page you are looking for might have been removed, had its name changed, or is temporarily unavailable.

Please try the following:

- If you typed the page address in the Address bar, make sure that it is spelled correctly.
- Open the www.shawnandrews.ca home page, and then look for links to the information you want.
- Click the [Back](#) button to try another link.

HTTP 404 - File not found
Internet Information Services

Technical Information (for support personnel)

- More information:
[Microsoft Support](#)

Vollständigkeit

Die Vollständigkeit der Daten ist der Umfang, der die Datenelemente oder Datenpunkte umfasst, die notwendig sind, um die Anwendung, für die sie bestimmt sind, zu unterstützen.

Zum Beispiel:

- Ein Datensatz, der alle Investitionsdaten für alle Ministerien beinhaltet, ermöglicht einen vollständigen Überblick über die Regierungsausgaben.
- Eine Beschreibung von Daten, die in Echtzeit generiert sind und die das Datum und die Uhrzeit der letzten Änderung einschließt.

Empfehlungen:

- **Entwerfen Sie den Erfassungs- und Veröffentlichungsprozess**, um die notwendigen Datenpunkte aufzunehmen.
- **Überwachen Sie** kontinuierlich die Mechanismen der Aktualisierung.

Vollständigkeit: Beispiel

Hohe Vollständigkeit

```
:weather1-7 a dcat:Dataset ;
  dct:title "Measurements from weather stations 1-7" ;
  dct:description "Data from seven weather stations
    showing temparture, humidity,
    wind direction and wind speed" ;
  dct:modified "2013-07-01T19:20:30+01:00" ;
  dct:publisher <http://myweather.com/id/myweather> ;
  dcat:keyword "weather" ;
  dcat:landingpage <http://myweather.com/stations1-7.html> ;
  dcat:distribution :weatherdata-xlsx
.

:weatherdata1-7-xlsx a dcat:Distribution ;
  dct:format <http://publications.europa.eu/resource/authority/file-type/XLSX> ;
  dct:licence <http://creativecommons.org/licenses/CC0> ;
  dcat:downloadURL <http://myweather.com/stations1-7.xlsx>
.
```

Geringere Vollständigkeit

```
:weather1-7 a dcat:Dataset ;
  dct:title "Measurements from weather stations 1-7" ;
  dct:description "Data from seven weather stations
    showing temparture, humidity,
    wind direction and wind speed" ;
  dct:publisher <http://myweather.com/id/myweather> ;
  dcat:keyword "weather" ;
  dcat:landingpage <http://myweather.com/stations1-7.html> ;
  dcat:distribution :weatherdata-xlsx
.

:weatherdata1-7-xlsx a dcat:Distribution ;
  dct:format <http://publications.europa.eu/resource/authority/file-type/XLSX> ;
  dct:licence <http://creativecommons.org/licenses/CC0> ;
  dcat:downloadURL <http://myweather.com/stations1-7.xlsx>
.
```

ERROR: MISSING DATA dct:modified

Konformität

Die Konformität der Daten ist das Ausmaß, in dem sie einer Reihe von expliziten Regeln oder Standards für die Erfassung, Veröffentlichung und Beschreibung folgen.

Zum Beispiel:

- Ein Datensatz, der die Koordinaten in WGS84 und Statistiken in SDMX wiedergibt.
- Eine Beschreibung eines Datensatzes nach dem DCAT Anwendungsprofil.

Empfehlungen:

- **Gebrauchen Sie die am häufigsten verwendeten Standards** in der Domain, die für die Daten oder Metadaten am relevantesten ist.
- **Definieren Sie lokale Vokabulare, wenn kein Standard verfügbar ist**, aber veröffentlichen Sie Ihre Vokabulare entsprechend der Best Practices (z.B. dereferenzierbare URIs).

Konformität : Beispiel

Hohe Konformität

```
:weather1-7 a dcat:Dataset ;
  dct:title "Measurements from weather stations 1-7" ;
  dct:description "Data from seven weather stations
    showing temparture, humidity,
    wind direction and wind speed" ;
  dct:modified "2013-07-01T19:20:30+01:00" ;
  dct:publisher <http://myweather.com/id/myweather> ;
  dcat:keyword "weather" ;
  dcat:landingpage <http://myweather.com/stations1-7.html> ;
  dcat:distribution :weatherdata-xlsx
.

:weatherdata1-7-xlsx a dcat:Distribution ;
  dct:format <http://publications.europa.eu/resource/authority/file-type/XLSX> ;
  dct:licence <http://creativecommons.org/licenses/CC0> ;
  dcat:downloadURL <http://myweather.com/stations1-7.xlsx>
.
```

Geringere Konformität

```
:weather1-7 a dcat:Dataset ;
  dct:description "Data from seven weather stations
    showing temparture, humidity,
    wind direction and wind speed" ;
  dct:modified "2013-07-01T19:20:30+01:00" ;
  dct:publisher <http://myweather.com/id/myweather> ;
  dcat:keyword "weather" ;
  dcat:landingpage <http://myweather.com/stations1-7.html> ;
  dcat:distribution :weatherdata-xlsx
.

:weatherdata1-7-xlsx a dcat:Distribution ;
  dct:format <http://publications.europa.eu/resource/authority/file-type/XLSX> ;
  dct:licence <http://creativecommons.org/licenses/CC0> ;
  dcat:downloadURL <http://myweather.com/stations1-7.xlsx>
.
```

ERROR MISSING MANDATORY ELEMENT dct:title

Siehe auch:

https://joinup.ec.europa.eu/asset/adms_foss/news/just-released-admssw-validator-verify-and-visualise-rdf-software-metadata

Glaubwürdigkeit

Die Glaubwürdigkeit der Daten ist das Maß, in dem sie auf vertrauenswürdigen Quellen basieren oder von vertrauenswürdigen Organisationen geliefert werden.

Zum Beispiel:

- Ein Datensatz, der Daten von Prozessen enthält, die unabhängig überprüft werden können, z.B. Wahlergebnisse oder parlamentarischen Verfahren.
- Eine Beschreibung eines Datensatzes, die von einer Regierungsbehörde veröffentlicht wird.

Empfehlungen:

- **Basieren Sie Ihre Daten auf Quellen, denen vertraut werden kann** oder, dort, wo es möglich und sinnvoll ist, auf expliziten Dienstleistungsvereinbarungen.
- **Machen Sie entsprechende Zuordnungen**, so dass Wiederverwender entscheiden können, ob sie den Daten vertrauen können oder nicht.

Glaubwürdigkeit : Beispiel

Hohe Glaubwürdigkeit

Daten aus dem Amt für
Veröffentlichungen der EU:

```
<skos:ConceptScheme at:table.version.number="2013-05-29 14:01:09" at:table.id="language"
rdf:about="http://publications.europa.eu/resource/authority/language">
  <rdfs:label>Languages Authority Table</rdfs:label>
  <at:prefLabel xml:lang="en">Languages Authority Table</at:prefLabel>
</skos:ConceptScheme>
<skos:Concept rdf:about="http://publications.europa.eu/resource/authority/language/ENG" at:pr
skos:inScheme rdf:resource="http://publications.europa.eu/resource/authority/language">
  <at:authority-code>ENG</at:authority-code>
  <at:op-code>ENG</at:op-code>
  <atold:op-code>ENG</atold:op-code>
  <dc:identifier>ENG</dc:identifier>
  <at:start.use>1950-05-09</at:start.use>
  <skos:prefLabel xml:lang="bg">английски</skos:prefLabel>
  <skos:prefLabel xml:lang="cs">angličtina</skos:prefLabel>
  <skos:prefLabel xml:lang="da">engelsk</skos:prefLabel>
  <skos:prefLabel xml:lang="de">Englisch</skos:prefLabel>
  <skos:prefLabel xml:lang="el">αγγλικά</skos:prefLabel>
```

The Metadata Registry is maintained by the [Publications Office of the EU](#).

Geringere Glaubwürdigkeit

Daten aus Lexvo:

```
- <rdf:Description rdf:about="http://lexvo.org/id/iso639-3/eng">
  <rdf:type rdf:resource="lvont:Language"/>
  <rdfs:comment xml:lang="en" rdf:datatype="xsd:string"> English is a West
  Germanic language that arose in the Anglo-Saxon kingdoms of England and
  spread into what was to become south-east Scotland under the influence of
  the Anglian medieval kingdom of Northumbria. Following the extensive
  influence of Great Britain and the United Kingdom from the 18th century, via
  the British Empire, and of the United States since the mid-20th century, it
  has been widely dispersed around the world, becoming the leading language
  of international discourse and the lingua franca in many regions. It is widely
  learned as a second language and used as an official language of the
  European Union and many Commonwealth countries, as well as in many
  world organisations. It is the third most natively spoken language in the
  world, after Mandarin Chinese and Spanish.</rdfs:comment>
  <rdfs:label xml:lang="aa" rdf:datatype="xsd:string">English</rdfs:label>
  <rdfs:label xml:lang="ace" rdf:datatype="xsd:string">Bahsa Inggreh</rdfs:label>
  <rdfs:label xml:lang="af" rdf:datatype="xsd:string">Engels</rdfs:label>
  <rdfs:label xml:lang="agq" rdf:datatype="xsd:string">Kingele</rdfs:label>
  <rdfs:label xml:lang="aii" rdf:datatype="xsd:string">ܐܢܓܠܝܫܐ</rdfs:label>
  <rdfs:label xml:lang="ak" rdf:datatype="xsd:string">Borfo</rdfs:label>
  <rdfs:label xml:lang="ak" rdf:datatype="xsd:string">English</rdfs:label>
```

Rights: Lexvo.org is Copyright © 2008-2012 [Gerard de Melo](#). All rights reserved.

Liability for Contents: We make every reasonable effort to ensure that the content of Lexvo.org is accurate and up-to-date. Nevertheless, the possibility of errors and inaccuracies cannot be ruled out. We do not give any warranty with respect to the information provided from Lexvo.org being accurate, up-to-date, or complete. We disclaim all liability for material or non-material loss or damage arising directly or indirectly from the use of our services.

Lingvoj/Lexvo Daten müssen nicht unbedingt von geringerer Qualität sein, als diejenigen des Amtes für Veröffentlichungen, aber das Amt ist eine maßgebliche Quelle, während Linvoj und Lexvo Initiativen von Einzelpersonen sind.

Verarbeitbarkeit

Die Verarbeitbarkeit der Daten ist das Maß, in dem diese verstanden und von automatisierten Prozessen behandelt werden können.

Zum Beispiel:

- Ein Datensatz, der codierte Informationen enthält, die auf öffentlich zugänglichen, kontrollierten Vokabularen und Code-Listen basieren.
- Ein Beschreibung des Datensatzes, der Daten in W3C Date und Time Format (z.B. 2013-06-01) anstatt als Text (z.B. 1 June 2013) ausdrückt.

Empfehlungen:

- **Identifizieren Sie die Quelle von Terminologie und Codes** der Daten auf maschinenlesbarer Weise.
- **Nutzen Sie Empfehlungen für die Syntax** von Daten, die in gemeinsamen Standards und Anwendungsprofilen gegeben werden.

Verarbeitbarkeit : Beispiel

Hohe Verarbeitung

```
▼<recipe>
  <script/>
  <script/>
  <title>Hippie Pancakes</title>
  ▼<recipeinfo>
    <blurb>Socially conscious breakfast food.</blurb>
    <author>David Horton</author>
    <yield>12 to 16 small pancakes, enough for two hippies</yield>
    <preptime>10 minutes</preptime>
  </recipeinfo>
  ▼<ingredientlist>
    ▼<ingredient>
      <quantity>1</quantity>
      <unit>C.</unit>
      <fooditem>unbleached wheat blend flour</fooditem>
    </ingredient>
    ▼<ingredient>
      <quantity>2</quantity>
      <unit>tsp.</unit>
      <fooditem>baking powder</fooditem>
    </ingredient>
    ▼<ingredient>
      <quantity>1</quantity>
      <unit>tsp.</unit>
      <fooditem>unrefined sugar</fooditem>
    </ingredient>
    ▼<ingredient>
      <quantity>1/4</quantity>
      <unit>tsp.</unit>
      <fooditem>coarse kosher salt</fooditem>
    </ingredient>
    ▼<ingredient>
      <quantity>1</quantity>
      free-range egg
    ..
  ..
```

Geringere Verarbeitung

Hippie Pancakes

Socially conscious breakfast food.
Recipe by: David Horton
Yield: 12 to 16 small pancakes, enough for two hippies
Preptime: 10 minutes

Ingredients

1 C. unbleached wheat blend flour
2 tsp. baking powder
1 tsp. unrefined sugar
1/4 tsp. coarse kosher salt
1 free-range egg
1 1/4 C. hormone-free milk
1 tsp. organic vegetable oil

Preparation Instructions

Pre-heat griddle over medium heat. Combine dry ingredients in a mixing bowl. Stir in egg, milk and oil. Use a large spoon or gravy ladle to transfer pancake batter to the griddle. Pancakes are ready to flip when large bubbles can be seen on top.

Serving Instructions

Top with fruit and berries or serve with traditional maple syrup.

Relevanz

Die Relevanz der Daten ist das Maß, in dem sie die notwendigen Informationen enthalten, um die Anwendung zu unterstützen.

Zum Beispiel:

- Ein Datensatz, der für Klima-Berechnungen zu Grad Celsius gerundete Temperaturmessungen enthält; Ein Datensatz mit der Genauigkeit von einem tausendstel Grad für chemische Reaktionen.
- Eine Beschreibung eines Datensatzes, der Daten nur in einem bestimmten zeitlichen Rahmen, in dem sie für die Verarbeitung notwendig sind, enthält.

Empfehlungen:

- **Stimmen Sie die Abdeckung und Granularität** von Daten auf ihre beabsichtigte Verwendung im Rahmen der verfügbaren Zeit und Geld ab.
- **Berücksichtigen Sie jedoch auch potenzielle zukünftige Verwendungen** der Daten.

Relevanz : Beispiel

Hohe Relevanz

	Engine (cm3)	Fuel type	CO2 (g/km)	Tax (%)
Car Type 1	900	Gasoline	90	0
Car Type 2	1.100	Gasoline	120	5
Car Type 3	1.300	Gasoline	125	5
Car Type 4	1.400	Gasoline	150	5
Car Type 5	1.800	Diesel	180	10
Car Type 6	2.200	Diesel	190	10
Car Type 7	2.500	Gasoline	210	15

Geringere Relevanz

	Engine (cm3)	Fuel type	CO2 (g/km)	Color	Tax (%)
Car Type 1	900	Gasoline	90	Red	0
Car Type 2	1.100	Gasoline	120	Silver	5
Car Type 3	1.300	Gasoline	125	Black	5
Car Type 4	1.400	Gasoline	150	White	5
Car Type 5	1.800	Diesel	180	Silver	10
Car Type 6	2.200	Diesel	190	Blue	10
Car Type 7	2.500	Gasoline	210	Black	15

Aktualität

Die Aktualität der Daten ist das Maß, in dem sie den gegenwärtigen Zustand des Objektes oder des Ereignisses richtig widerspiegeln, und das Maß, in dem die Daten (in ihre letzten Version) ohne unnötige Verzögerung verfügbar gemacht werden.

Zum Beispiel:

- Ein Datensatz, der Echtzeit-Verkehrsdaten enthält, die alle paar Minuten aktualisiert werden.
- Ein Beschreibung eines Datensatzes, der jährliche Kriminalstatistiken enthält, die innerhalb von Tagen der Veröffentlichung des Datensatzes verfügbar gemacht werden.

Empfehlungen:

- **Passen Sie die Update-Frequenz der Daten an die Natur der Daten und ihre Verwendungszwecke an.**
- Stellen Sie sicher, dass **Prozesse und Werkzeuge vorhanden sind**, um die Aktualisierung zu unterstützen.

Aktualität: Beispiel

Hohe Aktualität

NOAA's National Weather Service
Pacific Tsunami Warning Center

Home News Organization Search for: NWS All NOAA

DOC > NOAA > NWS > PTWC

Click on the tabs below to see tsunami messages relevant to each of PTWC's [areas of responsibility](#).

[All Regions](#) [Pacific Ocean](#) [Hawai'i](#) [Caribbean Sea](#) [Indian Ocean \(discontinued\)](#)

Page last loaded at: Tue, 03 Dec 2013 15:10:09 UTC.

 **No Current Warning, Watch, or Advisory in Effect**
low [Click here to read the latest tsunami message.](#)

Tsunami Messages for All Regions (Past 30 days)

Click on the map or table below for more information.

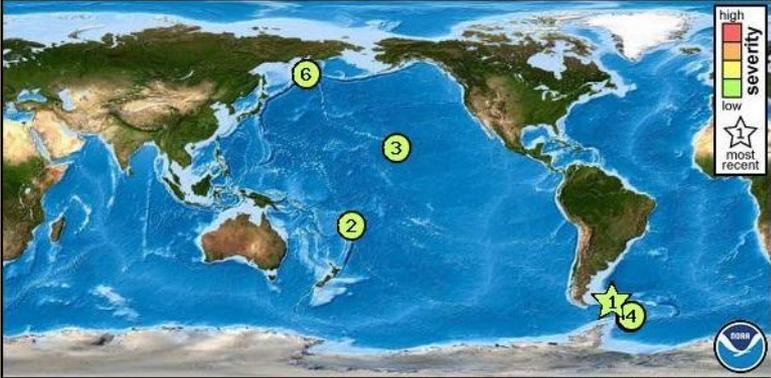


Image last created on Tue, 03 Dec 2013 15:09:23 UTC.

Geringere Aktualität



Best Practices

*Best Practices für die Veröffentlichung von hochwertigen Daten
und Metadaten*

W3C: Best Practices für die Veröffentlichung von Linked Open Data

IDENTIFY Identify data sets that other people may wish to re-use.

MODEL Model the data in an application-independent, objective way in terms of representation. Denormalize the data as necessary.

METADATA Provide basic metadata, including MIME type, publishing organization and/or agency, creation date, modification date, version, frequency of updates, contact email for the data steward(s).

PII Do not Publish Personally Identifiable Information as Open Data on the Web Data on the public Web can be potentially misused. Examples of personally identifiable data include: individual names, national identification number, phone number, credit card number and driver license number.

NAME Use HTTP URIs as names for your objects. Give careful consideration to the URI naming strategy. Consider how the data will change over time and name as necessary.

STANDARD_VOCABULARIES Describe objects with standard vocabularies whenever possible.

VOCABULARY_USE Use vocabularies as loosely coupled modular components.

REPRESENTATION Convert the source data into a Linked Data representation, also called an RDF serialization including Turtle, Notation-3 (N3), N-Triples, XHTML with embedded RDFa, and RDF/XML.

HUMAN READABLE Provide human readable descriptions with your Linked Data.

MACHINE ACCESSIBLE Provide access to the data representation via RESTful API, SPARQL endpoint(s) and RDF download.

SPECIFY_LICENSE Specify an appropriate license.

HOST Deliver open government data on authoritative domain to increase perceived trust.

ANNOUNCE Announce open government data, have a feedback mechanism and be prepared to be responsive to feedback.

SOCIAL_CONTRACT Maintenance is critical. Without a permanent identifier scheme, if you move or remove data that is published to the Web, you may break third party applications or mashups which is clearly undesirable. URI strategy and implementation are critical.

Siehe auch:

<http://www.slideshare.net/OpenDataSupport/the-linked-open-government-data-lifecycle>

Opquast: 72 Good Practices für Open Data

Einige Beispiele

Metadata	1	23	Each dataset is accompanied by a descriptive record
Metadata	1	24	Each dataset includes at least a title and a description
Metadata	1	25	A creation date is given for each dataset
Metadata	1	26	A last-updated date is given for each dataset
Metadata	2	27	The datasets are categorised
Format	1	33	Each dataset includes a reference to the charset used
Format	1	34	The format of downloadable files is indicated
Format	1	35	Dates are given in documented formats
Format	2	36	Dates are available in a standard format
Format	2	37	Data is provided in at least one open format

License	1	47	The datasets are accompanied by a licence
License	1	48	The licence sets out the conditions of attribution, reuse, redistribution and commercialisation
License	2	49	Usage rights are provided for an unlimited period
License	2	50	Data producers declare their policy on releasing data
License	3	51	The datasets are accompanied by a summary and a link to the full version of the licence
Linkeddata	2	52	Any vocabularies used within the dataset are identified and documented
Linkeddata	3	53	Data adheres to the defined syntax of any specified vocabularies
Linkeddata	3	54	It is possible to query data and metadata in accordance with standards of the web of data (Linked Open Data)

Siehe auch:

<http://checklists.opquast.com/en/opendata>

Was sind die gemeinsamen Elemente der Best Practices?

- Sie **liefern die entsprechenden Beschreibungen** von Daten (d.h. Metadaten).
- Sie **verwenden Standard-Vokabulare** für Metadaten und Daten, wann immer solche Vokabulare existieren.
- Sie **geben die Lizenz an**, unter der die Daten wiederverwendet werden können.
- Sie **halten an gesetzlichen Anforderungen fest**, die den Schutz von persönlichen und anderen sensible Daten betreffen.
- Sie **stellen Metadaten und Daten gemäß den Prinzipien der Linked Data mit persistenten URIs** für das Identifizieren von Dingen **dar**.
- Sie **liefern Informationen über die Quelle** der Daten.

Die Pflege von Metadaten und Daten ist entscheidend!

Siehe auch:

<http://www.slideshare.net/OpenDataSupport/introduction-to-metadata-management>

Schlussfolgerungen

- Die Qualität der Daten wird von ihre Eignung für die (Wieder-) Verwendung durch Datenverbraucher bestimmt.
- Metadaten sind “Daten über Daten”, d.h. Metadaten sind eine Art von Daten.
 - Die Qualitätsaspekte gelten gleichermaßen für Daten und Metadaten.
- Datenqualität hat mehrere Dimensionen und geht über die Richtigkeit von Daten hinaus.
 - Genauigkeit, Verfügbarkeit, Vollständigkeit, Konformität, Konsistenz, Glaubwürdigkeit, Verarbeitbarkeit, Relevanz, Aktualität.

Gruppenfragen



<http://www.visualpharm.com>

Welche Faktoren tragen Ihrer Meinung nach am meisten zur Qualität von Metadaten bei?



<http://www.visualpharm.com>

Die Verbesserung der Qualität kann Zeit und Ressourcen beanspruchen. In welchem Maße wäre Ihre Organisation bereit, in die Qualität von Metadaten zu investieren?



<http://www.visualpharm.com>

Könnten Sie ein Beispiel für eine hochwertige Metadaten-Beschreibung eines Datensatzes nennen, bei dem das DCAT Anwendungsprofil genutzt wird?

Nehmen Sie auch den Online-Test!

Vielen Dank!

...und jetzt IHRE Fragen?

Diese Präsentation wurde von Open Data Support erstellt

Disclaimers

1. The views expressed in this presentation are purely those of the authors and may not, in any circumstances, be interpreted as stating an official position of the European Commission. The European Commission does not guarantee the accuracy of the information included in this presentation, nor does it accept any responsibility for any use thereof.

Reference herein to any specific products, specifications, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favouring by the European Commission.

All care has been taken by the author to ensure that s/he has obtained, where necessary, permission to use any parts of manuscripts including illustrations, maps, and graphs, on which intellectual property rights already exist from the titular holder(s) of such rights or from her/his or their legal representative.

2. This presentation has been carefully compiled by PwC, but no representation is made or warranty given (either express or implied) as to the completeness or accuracy of the information it contains. PwC is not liable for the information in this presentation or any decision or consequence based on the use of it. PwC will not be liable for any damages arising from the use of the information contained in this presentation. The information contained in this presentation is of a general nature and is solely for guidance on matters of general interest. This presentation is not a substitute for professional advice on any particular matter. No reader should act on the basis of any matter contained in this publication without considering appropriate professional advice.

Autoren:

Maxx Dekkers, Michiel De Keyzer, Nikolaos Loutas and Stijn Goedertier

Referenzen

Folie 5:

- Juran, Joseph M. and A. Blanton Godfrey, Juran's Quality Handbook, Fifth Edition, p. 2.2, McGraw-Hill, 1999

Folie 6:

- National Information Standards Organization, <http://www.niso.org/publications/press/UnderstandingMetadata.pdf>

Folien 8-26:

- Mark David Hansen. Zero Defect Data: Tackling the Corporate Data Quality Problem. 1991. <http://dspace.mit.edu/handle/1721.1/13812>
- Kevin Roebuck. Data Quality: High-impact Strategies - What You Need to Know: Definitions, Adoptions, Impact, Benefits, Maturity, Vendors. Emereo Pty Limited, 2011. <http://bit.ly/19Qb6Ov>
- Thomas R. Bruce, Diane Hillmann. The Continuum of Metadata Quality: Defining, Expressing, Exploiting. ALA Editions, 2004. <http://www.ecommons.cornell.edu/handle/1813/7895>
- Sharon Dawes. Open data quality: a practical view. Open Data Roundtable. October 2012. <http://www.slideshare.net/cityhub/sharon-dawes-ctg>
- Joshua Tauberer. Open Government Data. Section 5.2 Data Quality: Precision, Accuracy, and Cost. June 2012. <http://opengovdata.io/2012-02/page/5-2/data-quality-precision-accuracy-and-cost>
- Stefan Urbanek. Data Quality: What is It? January 2011. <http://ckan.org/2011/01/20/data-quality-what-is-it/>
- Amrapali Zaveri, Anisa Rula, Andrea Maurino, Ricardo Pietrobon, Jens Lehmann, Sören Auer. Quality Assessment Methodologies for Linked Open Data. Semantic Web Journal (unpublished), 2012. <http://www.semantic-web-journal.net/content/quality-assessment-methodologies-linked-open-data>

Folie 13:

- ISA Programme. 10 Rules for Persistent URIs. <https://joinup.ec.europa.eu/community/semic/document/10-rules-persistent-uris>

Folie 14:

- European Commission. Telecommunication services: Access to networks (1 000). <http://open-data.europa.eu/>

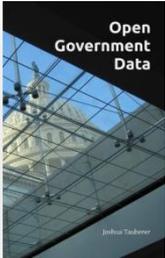
Folie 28:

- W3C. Best Practices for Publishing Linked Data. W3C Note 06 June 2013. <https://dvcs.w3.org/hg/gld/raw-file/default/bp/index.html>

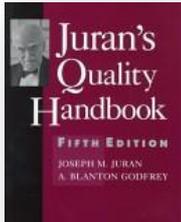
Folie 29:

- OPQUAST. 72 Open data good practices. <http://checklists.opquast.com/en/opendata>

Weiter lesen



Joshua Tauberer. Open Government Data. <http://opengovdata.io/>



Juran, Joseph M. and A. Blanton Godfrey, Juran's Quality Handbook

Verwandte Projekte und Initiativen



Best Practices for Publishing Linked Data.

<https://dvcs.w3.org/hg/gld/raw-file/default/bp/index.html>



OPQUAST. Open data good practices.

<http://checklists.opquast.com/en/opendata>



Eurostat. European Statistical System

http://epp.eurostat.ec.europa.eu/portal/page/portal/ess_eurostat/introduction

Werden Sie Teil unseres Teams...

*Finden Sie
uns auf*



Open Data Support

<http://www.slideshare.net/OpenDataSupport>



Open Data Support

<http://goo.gl/y9ZZI>

*Folgen Sie
uns auf*



@OpenDataSupport

*Begleiten Sie
uns auf*



joinup

<http://www.opendatasupport.eu>

*Kontaktieren
Sie uns unter*

contact@opendatasupport.eu