



# OPEN DATA SUPPORT

## Trainingsmodul 2.2

# Die Qualität von Open Data & Metadaten



*Die Mitglieder des PwC Netzwerks unterstützen Organisationen und Individuen dabei, die Werte zu schaffen, nach denen sie suchen. Wir sind ein Netzwerk von Unternehmen mit nahezu 180.000 Mitarbeitern in 158 Ländern, die sich dazu verpflichtet fühlen Qualität in den Bereichen Assurance, Tax & Legal sowie Advisory zu liefern. Sagen Sie uns, was Ihnen wichtig ist und besuchen Sie für weitere Informationen unsere Webseite [www.pwc.com](http://www.pwc.com)  
PwC bezieht sich auf das PwC Netzwerk und/oder eine oder mehrere Mitgliedsfirmen, von denen jede ein rechtlich selbstständiges Unternehmen ist. Besuchen Sie unsere Webseite [www.pwc.com/structure](http://www.pwc.com/structure) für weitere Details.*

# Metadaten der Präsentation

Open Data Support wird von der Europäischen Kommission, 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) finanziert.

© 2014 European Commission

---

***Diese Präsentation wurde von PwC erstellt***

***Autoren:***

**Makx Dekkers, Michiel De Keyzer, Nikolaos Loutas and Stijn Goedertier**

## ***Disclaimer***

1. Die Ansichten, die in dieser Präsentation vertreten werden, spiegeln einzig die Meinung des Autors wider und dürfen unter keinen Umständen als offizielle Position der Europäischen Kommission interpretiert werden.

Die Europäische Kommission übernimmt weder eine Garantie für die Genauigkeit der Informationen, die Inhalt dieser Präsentation sind, noch akzeptiert sie jegliche Verantwortung für die Nutzung der selbigen.

Referenzen innerhalb dieser Präsentation zu spezifischen Produkten, Spezifikationen, Prozessen oder Services durch Handelsnamen, Markenzeichen, Hersteller oder ähnliches, implizieren nicht unbedingt deren Unterstützung oder Bevorzugung durch die Europäische Kommission.

Der Autor hat sämtliche Anstrengungen unternommen, um sicherzustellen, dass er/sie, wo nötig, die Erlaubnis erhalten hat, die einzelnen Teile der in dieser Präsentation genutzten Manuskripte zu benutzen. Dies beinhaltet das Einholen einer Nutzungserlaubnis von den Lizenzinhabern oder deren gesetzlichen Vertretern zur Nutzung von Illustrationen, Karten, Schaubildern, die durch Rechte des geistigen Eigentums geschützt sind.

2. Diese Präsentation wurde sorgfältig von PwC zusammengestellt, jedoch gibt PwC keine Erklärung darüber ab und übernimmt keine Garantie dafür (weder ausdrücklich noch implizit), dass die Informationen in dieser Präsentation vollständig oder akkurat sind. PwC ist nicht für die Informationen in dieser Präsentation oder jedwede Entscheidung oder Konsequenz, die aus ihr resultieren, haftbar zu machen. PwC haftet nicht für etwaige Schäden, die durch die Nutzung der Informationen in dieser Präsentation entstehen sollten. Die Informationen in dieser Präsentation sind genereller Natur und einzig dafür bestimmt, eine Orientierungshilfe für Themen allgemeinen Interesses zu sein. Diese Präsentation ist kein Ersatz für professionelle Beratung zu jedwedem speziellen Thema. Kein Leser sollte auf der Basis der Themendarstellungen in dieser Präsentation handeln, ohne vorher angemessene professionelle Beratung einzuholen.

---

# *Lernziele*

Am Ende dieses Trainingsmoduls sollten Sie:

- wissen, was Qualität in Bezug auf Open Data bedeutet;
- die Faktoren und Kriterien kennen, die die Qualität von Open Data bestimmen;
- Good Practices für die Veröffentlichung von hochwertigen (Linked) Open Data kennen.

---

# *Inhalt*

Dieses Modul enthält...

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

---

## ***Was ist (Meta-)Datenqualität?***

*Daten sind dann von hoher Qualität, „wenn Sie für ihre vorgesehene Verwendung im operativen Geschäft, bei Entscheidungen oder der Planung geeignet sind.“*

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 anderweitig einfacher machen, eine Informationsquelle abzurufen, zu verwenden oder zu verwalten. Metadaten werden oft ‘Daten über Daten’ genannt.”*

*-- National Information Standards Organization*

- Wir stellen fest, dass Metadaten eine Art von Daten sind.
- Für Daten und Metadaten gelten die gleichen Überlegungen zur Qualitätssicherung.

---

# Dimensionen der Datenqualität

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

# *Dimensionen der Datenqualität*

- **Genauigkeit:** Stellen die Daten Objekte und Ereignisse aus der realen Welt richtig dar?
- **Konsistenz:** Enthalten die Daten Widersprüche?
- **Verfügbarkeit:** Können die Daten sowohl jetzt als auch in Zukunft abgerufen werden?
- **Vollständigkeit:** Umfassen die Daten alle Datenelemente, die das 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?
- **Aktualität:** 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 des realen Objekts, der Situation oder des Ereignisses richtig darstellen.*

*Zum Beispiel:*

- Korrektes Messen von Witterungsverhältnissen (Temperatur, Niederschlag).
- Korrekter Verweis auf die Wiederverwendungsbestimmungen des Datensatzes.

## **Empfehlungen:**

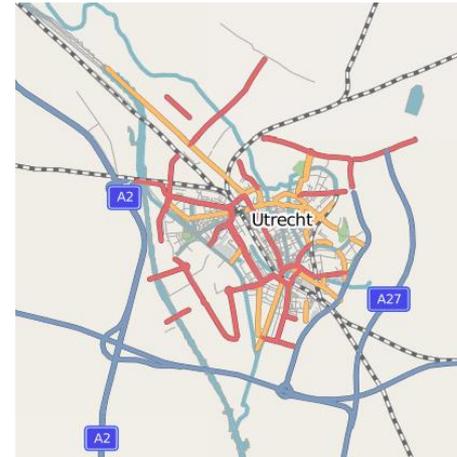
- **Wägen** Sie die **Genauigkeit** Ihrer Daten gegen die entstehenden **Kosten** ab; die Daten müssen **für ihre vorgesehene Verwendung gut genug** sein.
- Stellen Sie sicher, dass eine **organisatorische Selbstverpflichtung** besteht und in **Verfahren und Werkzeuge investiert** wird, um die Genauigkeit Ihrer Daten beizubehalten.

# *Genauigkeit: Beispiel*

## Hohe Genauigkeit



## Wenig Genauigkeit



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

# Konsistenz

*Die Konsistenz von Daten ist das Ausmaß, in dem diese keine Widersprüche enthalten, die ihre Verwendung schwierig oder unmöglich machen würden.*

*Zum Beispiel:*

- Ein Datensatz, der Daten aus verschiedenen Quellen verbindet, die so angepasst wurden, dass sie widersprüchliche Aussagen aufdecken, die gelöst wurden.
- Eine Beschreibung eines Datensatzes, der nicht mehrere Lizenzaussagen enthält oder bei der die Daten der letzten Änderung nicht vor dem Erstellungsdatum liegen.

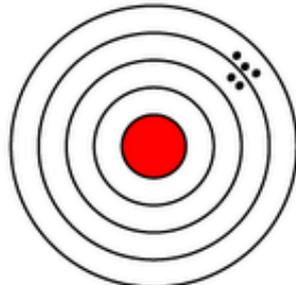
**Empfehlungen:**

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

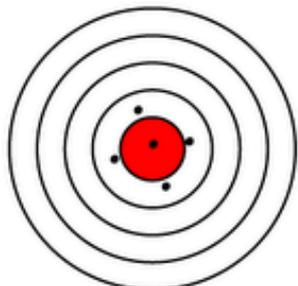
# Konsistenz : Beispiel



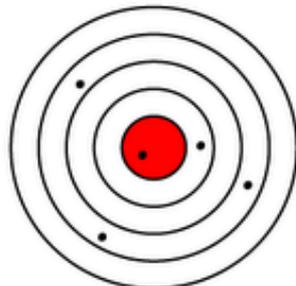
Accurate & Consistent



Not Accurate  
but Consistent



Accurate but  
Not Consistent



Not Accurate &  
Not Consistent

## Hohe 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 "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>
.
```

## Geringe 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 korrekte Ressource verweist.
- Eine Beschreibung der Datensätze, die in der Suchmaschine eines Datenportals enthalten ist.

## **Empfehlungen:**

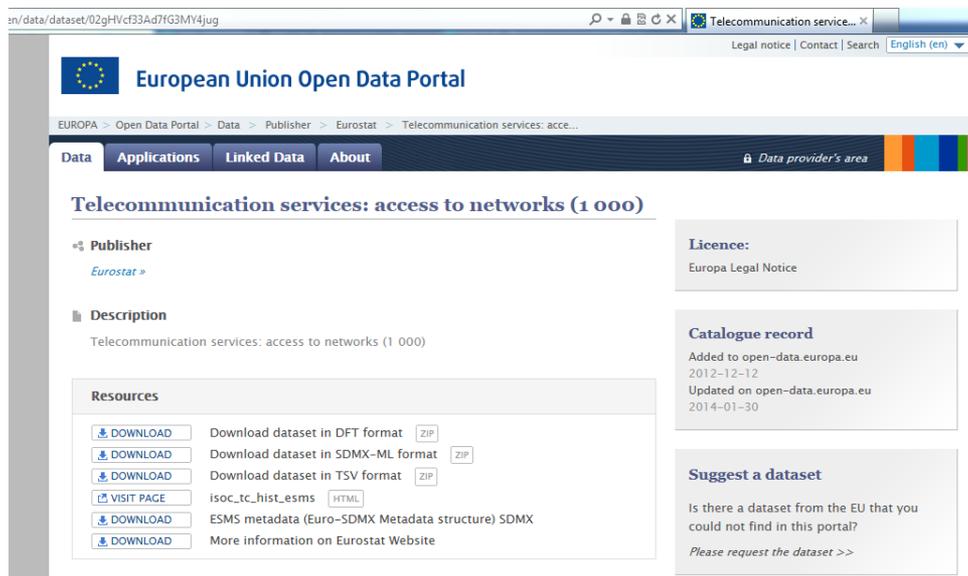
- Orientieren Sie sich bei der Zuordnung und Pflege von URIs an Best Practices.
- Stellen Sie sicher, dass **die Verantwortung** für die Pflege von Daten in der Organisation **eindeutig zugewiesen** ist.

**Siehe auch:**

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

# Verfügbarkeit: Beispiel

## Hohe Verfügbarkeit



European Union Open Data Portal

EUROPA > Open Data Portal > Data > Publisher > Eurostat > Telecommunication services: access to networks (1 000)

**Telecommunication services: access to networks (1 000)**

**Publisher**  
Eurostat >

**Description**  
Telecommunication services: access to networks (1 000)

**Resources**

- Download dataset in DFT format
- Download dataset in SDMX-ML format
- Download dataset in TSV format
- isoc\_tc\_hist\_esms
- ESMS metadata (Euro-SDMX Metadata structure) SDMX
- More information on Eurostat Website

**Licence:**  
Europa Legal Notice

**Catalogue record**  
Added to open-data.europa.eu 2012-12-12  
Updated on open-data.europa.eu 2014-01-30

**Suggest a dataset**  
Is there a dataset from the EU that you could not find in this portal?  
*Please 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](http://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 gegeben, wenn sie diejenigen Datenelemente oder Datenpunkte enthalten, die notwendig sind, um die bestimmungsgemäße Anwendung der Daten zu unterstützen.*

*Zum Beispiel:*

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

**Empfehlungen:**

- **Entwerfen Sie den Erfassungs- und Veröffentlichungsprozess** so, dass die notwendigen Datenpunkte mit aufgenommen werden.
- **Sorgen Sie für eine kontinuierliche Überwachung** der Aktualisierungsmechanismen.

# 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>
.
```

## Geringe 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 von Daten ist gegeben, wenn sie einer Reihe von expliziten Regeln oder Standards für die Erfassung, Veröffentlichung und Beschreibung entsprechen.*

*Zum Beispiel:*

- Ein Datensatz, der Koordinaten in WGS84 und Statistiken in SDMX wiedergibt.
- Eine Beschreibung eines Datensatzes gemäß des DCAT Anwendungsprofils.

***Empfehlungen:***

- **Gebrauchen Sie die am häufigsten verwendeten Standards** 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>
.
```

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

# Glaubwürdigkeit

*Die Glaubwürdigkeit von Daten ist gegeben, wenn sie auf vertrauenswürdigen Quellen basieren oder von vertrauenswürdigen Organisationen geliefert wurden.*

*Zum Beispiel:*

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

**Empfehlungen:**

- **Verknüpfen Sie Ihre Daten mit Quellen, denen man vertrauen kann** oder (wo möglich und sinnvoll) auf expliziten Dienstleistungsvereinbarungen.
- **Weisen Sie entsprechende Attribute zu**, 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:p
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](#).

## Geringe 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">ܐܢܓܠܝܬܝܢ</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. Jedoch ist das Amt eine maßgebliche Quelle, während Linvoj und Lexvo Initiativen von Einzelpersonen sind.

# Verarbeitbarkeit

*Die Verarbeitbarkeit von Daten ist gegeben, wenn diese von automatisierten Prozessen verstanden und verarbeitet werden können.*

*Zum Beispiel:*

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

***Empfehlungen:***

- **Weisen Sie die Quelle von Terminologie und Codes** der Daten auf maschinenlesbarer Weise aus.
- **Nutzen Sie Empfehlungen für die Syntax** von Daten, die auf gemeinsamen Standards und Anwendungsprofilen basieren.

# Verarbeitbarkeit: Beispiel

## Hohe Verarbeitbarkeit

```
▼<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
    ..
  ..
```

## Geringe Verarbeitbarkeit

### 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 von Daten ist gegeben, wenn sie die notwendigen Informationen enthalten, um ihre vorgesehene Verwendung zu unterstützen.*

*Zum Beispiel:*

- Ein Datensatz für Klima-Berechnungen, der gerundete Temperaturmessungen auf ‚Grad Celsius‘ enthält; Ein Datensatz für chemische Reaktionen mit einer Genauigkeit von einem tausendstel Grad.
- Eine Beschreibung eines Datensatzes, die Daten nur für den zeitlichen Rahmen enthält, der für die Verarbeitung notwendig ist.

**Empfehlungen:**

- **Stimmen Sie die Reichweite und Beschaffenheit** von Daten auf deren beabsichtigte Verwendung mit Rücksicht auf die verfügbare 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

## Geringe 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 von Daten ist gegeben, wenn sie den gegenwärtigen Zustand des Objektes oder des Ereignisses richtig wiedergeben und wenn die Daten (in ihrer aktuellen 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, die jährliche Kriminalstatistiken enthält, die innerhalb von Tagen nach der Veröffentlichung des Datensatzes verfügbar gemacht werden.

**Empfehlungen:**

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

# Aktualität: Beispiel

## Hohe Aktualität

NOAA's National Weather Service [weather.gov](http://weather.gov)

### 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.

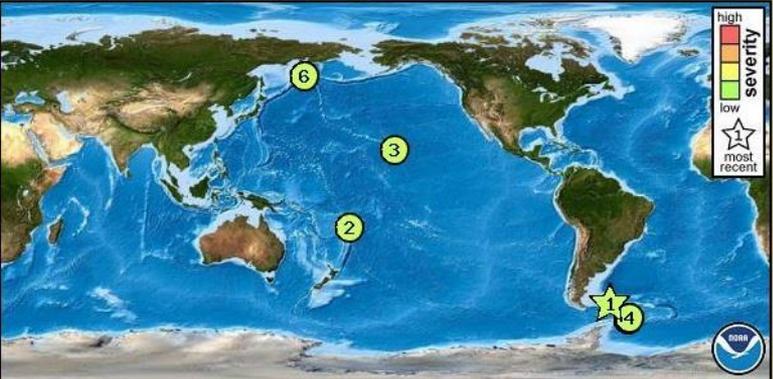


Image last created on Tue, 03 Dec 2013 15:09:23 UTC.

## Geringe 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 adäquate Beschreibungen** von Daten (d. h. Metadaten).
- Sie **verwenden Standardvokabulare** 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 sensiblen 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 ihrer Eignung zur (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 Kriterien 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?

Machen Sie auch den Online-Test!

---

Vielen Dank!

...und jetzt IHRE Fragen?

# 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/>

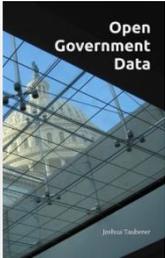
Folie 28:

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

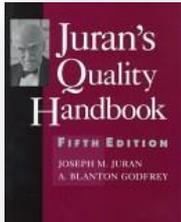
Folie 29:

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

# Weiterführende Informationen



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

# Werden Sie Teil unseres Teams...

*Besuchen Sie  
uns*



[Open Data Support](http://www.slideshare.net/OpenDataSupport)

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



[Open Data Support](http://goo.gl/y9ZZI)

<http://goo.gl/y9ZZI>

*Lernen Sie uns  
kennen*



joinup

<http://www.opendatasupport.eu>

*Folgen Sie uns*



[@OpenDataSupport](https://twitter.com/OpenDataSupport)

*Kontaktieren Sie  
uns*

[contact@opendatasupport.eu](mailto:contact@opendatasupport.eu)