

Predstavitev metapodatki

Open Data Support podpira Evropska komisija v okviru projekta SMART 2012/0107 'Lot 2: Zagotavljanje storitev objave, dostopa in vnovične uporabe javnih podatkov v Evropski uniji preko obstoječih portalov odprtih podatkov (Pogodba št. 30-CE-0530965/00-17).

© 2014 Evropska komisija



OPEN DATA SUPPORT

Training Module 2.2

Odprti podatki & kakovost metapodatkov

Učni cilji

Po končanem modulu udeleženci pridobijo znanje o:

- o tem, kaj pomeni kakovost odprtih podatkov;
- determinantah in kriterijih kakovosti odprtih podatkov;
- dobrih praksah objavljanja zelo kakovostnih (povezanih) odprtih podatkov.

Vsebina

Ta modul vsebuje ...

- definicijo kakovosti podatkov;
- pregled obsega kakovosti podatkov in metapodatkov;
- izbor najboljših praks za objavo podatkov in metapodatkov dobre kakovosti.

Kaj je kakovost podatkov (in metadodatkov)?

Podatki so dobre kakovosti, “če dobro služijo svojemu namenu v operacijah, pri odločanju in načrtovanju.”

Natančneje:

“Visokokakovostni podatki so točni, razpoložljivi, popolni, skladni, konsistentni, verodostojni, omogočajo obdelavo, so relevantni in pravočasni.”

Metapodatki so podatki o podatkih ...

“Metapodatki so strukturirane informacije, ki opisujejo, pojasnjujejo, opredeljujejo ali kako drugače omogočajo lažje pridobivanje, uporabo ali upravljanje z viri informacij. Metapodatkom pogosto pravimo podatki o podatkih ali informacije o informacijah.”

-- National Information Standards Organization

- Metapodatki so torej posebna vrsta podatkov.
- Za metapodatke in podatke veljajo enake zahteve glede kakovosti.

Lastnosti kakovostnih podatkov

Katere glavne lastnosti je treba upoštevati za zagotovitev metapodatkov dobre kakovosti?

Lastnosti kakovostnih podatkov

- **Točnost:** ali podatek pravilno predstavlja realni svet ali dogodek?
- **Konsistentnost:** ali podatek morda ne vsebuje nasprotij?
- **Razpoložljivost:** ali je do podatka mogoče dostopati v tem trenutku in tudi čez nekaj časa?
- **Popolnost:** ali podatek vsebuje vse dele podatka, ki predstavljajo en subjekt ali dogodek?
- **Skladnost:** ali je podatek v skladu s sprejetimi standardi?
- **Verodostojnost:** ali podatek temelji na verodostojnih virih?
- **Zmožnost obdelave:** ali je podatek strojno berljiv?
- **Relevantnost:** ali podatki vsebujejo ustrezno količino podatkov?
- **Pravočasnost:** ali podatek predstavlja dejansko situacijo in ali je objavljen dovolj hitro?

Točnost

Točnost podatkov je stopnja, do katere podatki točno predstavijo lastnosti objekta, situacije ali dogodka iz realnega sveta.

Primer:

- pravilne meritve vremenskih pojavov (temperatura, padavine),
- pravilna navedba pogojev o vnovični uporabi nabora podatkov.

Priporočila:

- **Uravnotežite točnost** svojih podatkov v razmerju do **stroškov** v kontekstu uporabe; podatek mora biti **dovolj dober za nameravano uporabo**.
- Poskrbite, da bo **organizacija predana** in da bo **investirala v postopke in orodja**, s pomočjo katerih se ohranja točnost.

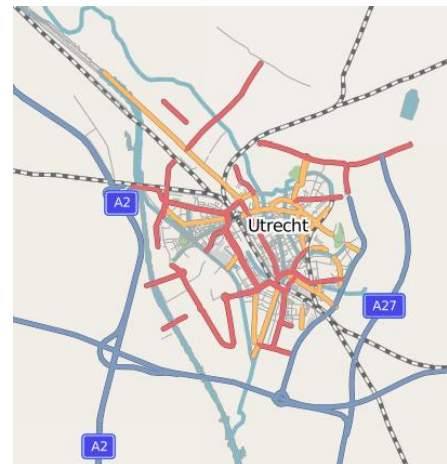
Točnost – primeri

Višja stopnja



Nižja stopnja

OpenStreetMap, Utrecht, Nizozemska
(2011 proti 2007)



Konsistentnost

Konsistentnost podatkov je stopnja, do katere podatek ne vsebuje nasprotij, ki bi njegovo uporabo otežila ali onemogočila.

Npr.:

- Nabor podatkov, ki je kombinacija podatkov iz različnih virov in je bil obdelan, zato da bi identificirali nasprotujoče si izjave, ki se nato odpravijo.
- Opis nabora podatkov, ki ne vsebuje več licenčnih izjav ali kjer datum zadnje spremembe ni pred datumom nastanka.

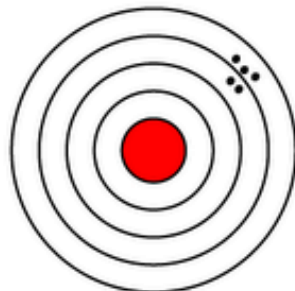
Priporočila:

- **Pred objavo obdelajte podatke**, da boste prepoznali nasprotujoče si izjave in druge napake (zlasti, če so podatki zbrani iz različnih virov).

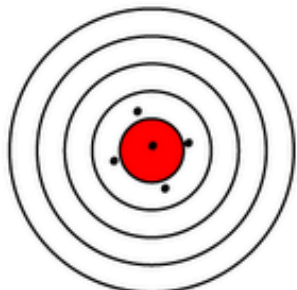
Konsistentnost (primeri)



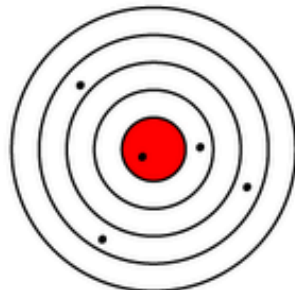
Accurate & Consistent



Not Accurate
but Consistent



Accurate but
Not Consistent



Not Accurate &
Not Consistent

veliko konsistence

```
: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>
.
```

manj konsistence

```
: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

Razpoložljivost

Razpoložljivost podatkov je stopnja dostopnosti podatkov; to vključuje tudi dolgoročno obstojnost podatkov.

Npr.:

- Nabor podatkov, ki ga prepozna http: URI in ki vedno usmeri na pravi vir (pri čemer se ne odzove z izjavo “404 Not found”).
- Opis nabora podatkov, ki je vključen v iskalnik podatkovnega portala.

Priporočila:

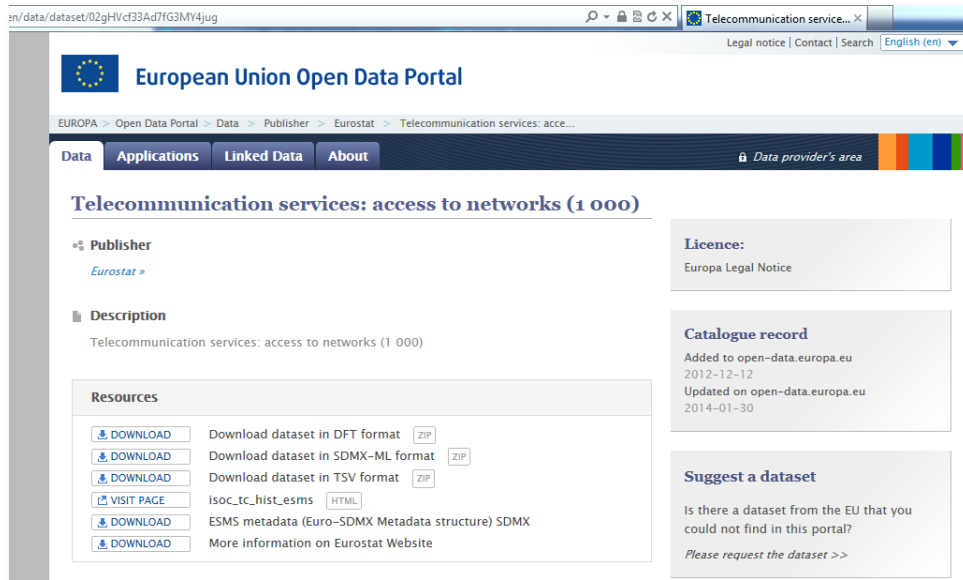
- Sledite **najboljši praksi** za določanje in vzdrževanje URI-jev.
- Zagotovite, da je **odgovornost** za vzdrževanje podatkov v organizaciji **jasno določena**.

See also:

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

Razpoložljivost (primer)

Dobra razpoložljivost



European Union Open Data Portal

EUROPA > Open Data Portal > Data > Publisher > Eurostat > Telecommunication services: acce...

Data Applications Linked Data About 🔒 Data provider's area

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 SDMXML format
- Download dataset in TSV format
- isoc_tc_hist_esms
- ESMS metadata (Euro-SDMX Metadata structure) SDMX
- More information on Eurostat Website

License:
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 >>](#)

Slaba razpoložljivost

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](#)

Popolnost

Popolnost podatkov je stopnja zagotovljenosti delov podatkov ali podatkovnih točk, ki so nujne za podporo aplikaciji, ki so ji namenjene.

Npr.:

- Nabor podatkov, ki vključuje podatke o izdatkih vseh ministrstev, omogoča popoln pregled nad izdatki vlade.
- Opis podatkov, ustvarjen v realnem času, ki vključuje datum in čas zadnje spremembe.

Priporočila:

- **Oblikujte postopek zajemanja in objave**, ki vključuje potrebne podatkovne točke.
- **Redno spremljajte** mehanizme posodabljanja.

Popolnost (primeri)

Zelo popolni

```
: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>
.
```

popolni

```
: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

Skladnost

Skladnost podatkov je stopnja prilagojenosti naboru izrecno določenih pravil ali standardov zajemanja, objave in deskripcije.

Npr.:

- Podatkovni niz, ki izraža koordinate v WGS84 in statistiko v SDMX.
- Opis nabora podatkov po DCAT Application Profile.

Priporočila:

- **Uporabite najbolj uporabljane standarde** v domeni, ki je najbolj relevantna za podatke in metapodatke.
- **Če standard ni na voljo, opredelite lokalne slovarje**, vendar objavite svoje slovarje v skladu z najboljšo prakso (npr. (angl.) *dereferenceable URIs*).

Primeri skladnosti

visoka skladnost

```
: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>
.
```

manj skladnosti

```
: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

Glej še:

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

Verodostojnost

Verodostojnost podatkov je stopnja, do katere podatki temeljijo na zaupanja vrednih virih ali pa jih zagotavljajo zaupanja vredne organizacije.

Npr.:

- Nabor podatkov, ki vsebuje podatke iz procesov, ki jih je mogoče neodvisno preveriti, npr. rezultati volitev ali parlamentarnih postopkov.
- Opis nabora podatkov, ki ga objavi vladna agencija.

Priporočila:

- **Vaši podatki naj temeljijo na virih, ki jim je mogoče zaupati**, ali pa, kjer je to mogoče, na izrecnih dogovorih o zagotavljanju storitev.
- **Dodajte ustrezne pripise**, da se lahko vnovični uporabniki odločijo, ali podatkom lahko zaupajo ali ne.

Verodostojnost (primeri)

Visoka stopnja verodostojnosti Podatki iz Urada za publikacije 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](#).

Manj verodostojnosti Podatki iz 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 Inggréh</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">ܐܢܓܪܝܬܐ</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>
  <rdfs:label xml:lang="am" rdf:datatype="xsd:string">ግብጽኛ</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.

Podatki Lingvoj/Lexvo morda niso slabše kakovost od tistih, ki jih zagotavlja Urad za publikacije, vendar je urad vir z avtoriteto, medtem ko sta Linvoj in Lexvo pobudi posameznikov.

Zmožnost obdelave

Zmožnost obdelave podatkov je stopnja, do katere je podatke mogoče razumeti in obvladovati v avtomatiziranem procesu.

Npr.:

- Nabor podatkov, ki vsebuje kodirane informacije na podlagi javno dostopnih kontroliranih slovarjem in seznamov kod.
- Opis nabora podatkov, ki datume izrazi v formatu W3C Date and Time (npr. 2013-06-01), ne pa v besedilu (npr. 1. junij 2013).

Priporočila:

- **Identificirajte vir terminologije in kod**, ki so uporabljene v podatkih na strojno berljiv način.
- **Držite se priporočil za sintakso** podatkov, ki jih zagotavljajo skupni standardi ali profili aplikacij.

Zmožnost obdelave

Boljša

```
▼<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
      ..
  </ingredientlist>
```

Slabša

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.

Relevantnost

Relevantnost podatkov je stopnja, do katere ti vsebujejo potrebne informacije za podporo aplikaciji.

Npr.:

- Nabor podatkov vsebuje meritve temperature, zaokrožene na stopinje Celzija za napoved vremena; nabor podatkov z natančnostjo tisočinke stopinje za kemijsko reakcijo.
- Opis nabora podatkov, ki vsebuje le podatke o časovnih lastnostih, če so potrebni za njihovo obdelavo.

Priporočila:

- **Uparite prekrivnost in granulacijo** podatkov za njihovo nameravano uporabo ob upoštevanju časa in denarja, ki ga imate na voljo.
- Vendar pa **razmislite tudi o potencialni prihodnji rabi** podatkov.

Relevantnost (primeri)

Visoka stopnja

	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

Nizka stopnja

	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

Pravočasnost

Pravočasnost podatkov je stopnja, do katere podatki odražajo trenutno stanje subjekta ali dogodka, in stopnja, do katere so podatki (v svoji najnovejši različici) razpoložljivi brez nepotrebne odlašanja.

Npr.:

- Nabor podatkov, ki vsebuje podatke o podatkovnem prometu iz realnega časa, ki se posodablja na nekaj minut.
- Opis nabora podatkov letne statistike o kaznivih dejanjih, ki je na voljo v nekaj dneh od objave nabora podatkov.

Priporočila:

- **Prilagodite gostoto posodabljanja podatkov naravi podatkov** in njeni nameravani uporabi.
- Zagotovite prisotnost **postopkov in orodij**, ki podpirajo posodobitev.

Pravočasnost: primeri

Visoka stopnja

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.

Nizka stopnja



Najboljše prakse

Najboljše prakse za objavo visokokakovostnih podatkov in metapodatkov.

W3C: najboljše prakse za objavo povezanih odprtih podatkov

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.

See also:

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

Opquast: 72 Dobre prakse – odprti podatki

Nekaj primerov

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)

See also:

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

Kateri so skupni elementi v najboljših praksah

- **Zagotovite ustrezne opise** podatkov (tj. metapodatkov).
- **Uporabljajte standardne slovarje** za metapodatke in podatke, če obstajajo.
- **Natančno opredelite licence**, pod katerimi se podatki lahko vnovič uporabijo.
- **Upoštevajte pravne zahteve** glede varstva osebnih in drugih občutljivih podatkov.
- Metapodatke in podatke predstavite **po načelih povezanih podatkov** ob uporabi **trajnih URI-jev** za identifikacijo stvari.
- **Zagotovite informacije o viru** podatkov.

Vzdrževanje metapodatkov in podatkov je bistvenega pomena!

See also:

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

Ugotovitve

- Kakovost podatkov določa njihova zmožnost, da jih uporabniki podatkov lahko (vnovič) uporabijo.
- Metapodatki so “podatki o podatkih”, kar pomeni, da so metapodatki določena vrsta podatkov.
 - Za podatke in metapodatke naj veljajo enake zahteve glede kakovosti.
- Kakovost podatkov ima več razsežnosti in ne pomeni le to, da so podatki pravilni.
 - Točnost, razpoložljivost, popolnost, skladnost, konsistentnost, verodostojnost, zmožnost obdelave, relevantnost, pravočasnost.

Vprašanja za delo v skupinah



<http://www.visualpharm.com>

Kateri dejavniki po vašem mnenju najbolj prispevajo h kakovosti podatkov?



<http://www.visualpharm.com>

Za izboljšanje kakovosti so potrebni čas in viri. V kolikšni meri bi bila vaša organizacija pripravljena investirati v kakovost metapodatkov?



<http://www.visualpharm.com>

Navedite primer visokokakovostne metapodatkovne deskripcije za nabor podatkov, ki uporablja DCAT Application Profile.

Rešite spletni test tukaj!

Hvala lepa!
...vprašanja?

Predstavitev je pripravila organizacija Open Data Support

Zavrnitev odgovornosti

Pogledi, izraženi v tej predstavitvi, so izključno pogledi avtorjev in se v nobenih okoliščinah ne smejo tolmačiti kot uradno mnenje Evropske komisije.

Evropska komisija ne odgovarja za točnost podatkov v tej predstavitvi niti ne prevzema nobene odgovornosti za njihovo uporabo .

Sklicevanje na katere koli posebne produkte, specifikacije, postopke ali storitve blagovnih znamk, proizvajalcev ali drugih oseb ne pomenijo nujno potrditev, priporočilo ali prednostno izbiro Evropske komisije.

Avtor je z vso skrbnostjo poskrbel, da je po potrebi za uporabo pridobil vsa dovoljenja za dele besedil, vključujoč ilustracije, zemljevide in grafične prikaze, na katerih že obstajajo intelektualne pravice nosilcev teh pravic ali njihovih zakonitih zastopnikov.

Predstavitev je skrbno pripravila skupina PwC, ki pa ne odgovarja in ne jamči (niti izrecno niti implicitno) za popolnost ali točnost vsebovanih informacij. PwC ne odgovarja za informacije v tej predstavitvi ali odločitve ali posledice, ki bi nastale na njihovi podlagi. Prav tako ni odgovoren za škodo, ki bi nastala na podlagi informacij iz te predstavitve. Informacije v tej predstavitvi so splošne narave in namenjene zgolj kot napotek v zadevah splošnega interesa. Predstavitev ni nadomestek za strokovni nasvet v kakršni koli zadevi. Bralec naj pred ukrepanjem na podlagi informacij v tej publikaciji poišče strokovni nasvet.

Avtorji:

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

Reference

Slide 5:

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

Slide 6:

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

Slide 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>

Slide 13:

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

Slide 14:

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

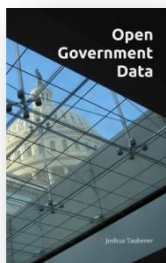
Slide 28:

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

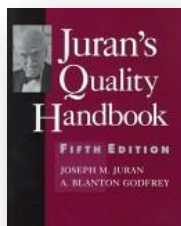
Slide 29:

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

Dodatno branje



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



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

Povezani projekti in pobude



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

Pridružite se nam ...

Poiščite nas na



Open Data Support

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



Open Data Support

<http://goo.gl/y9ZZI>

Sledite ščebetu



@OpenDataSupport

Pridružite se nam



<http://www.opendatasupport.eu>

Navežite stik

contact@opendatasupport.eu