

Presentation metadata

Open Data Support is funded by the European Commission under 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' (Contract No. 30-CE-0530965/00-17).

© 2013 European Commission



OPEN DATA SUPPORT

Training Module 2.2

Atvirų duomenų ir metaduomenų kokybė

Turinys

Pristatymą sudaro...

- Duomenų kokybės apibrėžimas.
- Duomenų bei metaduomenų kokybės aspektų apžvalga.
- Gerųjų praktikų, publikuojant kokybiškus duomenis bei metaduomenis, apžvalga.

Kas yra duomenų (ir metaduomenų) kokybė?

Aukštos kokybės duomenys tai duomenys kurie „atitinka numatytą paskirtį veiklos vykdymo, sprendimų priėmimo ir planavimo srityse“.

Arba:

„Aukštos kokybės duomenys yra tikslūs, prieinami, pilni, nuoseklūs, patikimi, apdorojami, tinkami ir nepasenę“.

Kas yra metaduomenys?

„Metaduomenys tai struktūrizuota informacija, kuri apibūdina, paaiškina, nurodo vietą ir kitais būdais palengvina informacijos išteklių valdymą. Dažnai metaduomenys vadinami duomenimis apie duomenis.“

-- Nacionalinė informacijos standartų organizacija

- Tai duomenų tipas.
- Tokie patys kokybės principai taikomi tiek duomenims, tiek metaduomenims.

Duomenų kokybės kriterijai

Į kokius pagrindinius kriterijus reikia atsižvelgti, teikiant geros kokybės (meta)duomenis?

Duomenų kokybės kriterijai

- **Tikslumas:** Ar duomenys tiksliai parodo realaus pasaulio subjektus bei įvykius?
- **Prieinamumas:** Ar duomenys yra pasiekiami dabar ir bus pasiekiami ilguoju laikotarpiu?
- **Pilnumas:** Ar duomenys pilnai parodo subjekto ar įvykio sudedamąsias dalis?
- **Atitikimas:** Ar duomenys atitinka pripažintus standartus?
- **Nuoseklumas:** Ar duomenyse nėra prieštaravimų?
- **Patikimumas:** Ar duomenys pagrįsti patikimais šaltiniais?
- **Apdorojimas:** Ar duomenys yra pateikiami kompiuterio skaitomu formatu?
- **Tinkamumas:** Ar tinkamas duomenų turinys?
- **Savalaikiškumas:** Ar duomenys parodo aktualią situaciją ir ar jie bus publikuojami pakankamai greitai?

Tikslumas

Duomenų tikslumas parodo, kiek teisingai konkretūs duomenys reprezentuoja realaus pasaulio objektus, situacijas bei įvykius.

Pavyzdžiui:

- Teisingi oro sąlygų matavimai (temperatūra, krituliai).
- Tiksliai įvardytos esamos sąlygos dėl pakartotino duomenų rinkinio naudojimo.

Rekomendacijos:

- Duomenų **tikslumas** vertinant iš jų parengimui reikalingų kaštų turi būti toks, kad juos būtų galima panaudoti norimam tikslui.
- Turi būti užtikrintas **procedūrų ir įrankių palaikymas** reikiamam duomenų tikslumo užtikrinimui.

Tikslumo pavyzdys



Paveiksluko licencija= [Creative Commons Attribution-Share Alike 3.0 Unported](#) license.

Priskyrimas: [Garytx at en.wikipedia](#)

Didelis tikslumas

Debesuotumas = 100%

Mažas tikslumas

Debesuota

Prieinamumas

Duomenų prieinamumas reiškia jų pasiekiamumą t.y. kokia apimtimi duomenys gali būti pasiekiami taip pat įvertinant jų pasiekiamumą ilguoju laikotarpiu.

Pavyzdžiui:

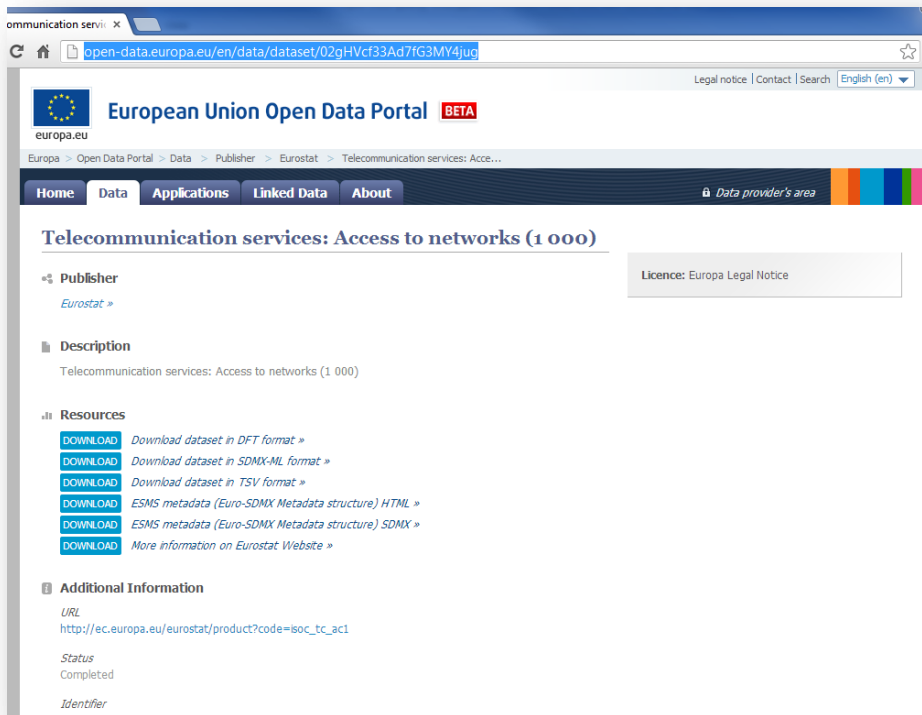
- Duomenų rinkinys, identifikuojamas pagal HTTP URI, kuris nukreipia į tinkamą resursą (ir negražina klaidos „puslapio nerasta“ ar angl. „404 Not found“).
- Duomenų įrankio aprašas, kuris yra įtrauktas į duomenų portalo paieškos mechanizmą.

Rekomendacijos:

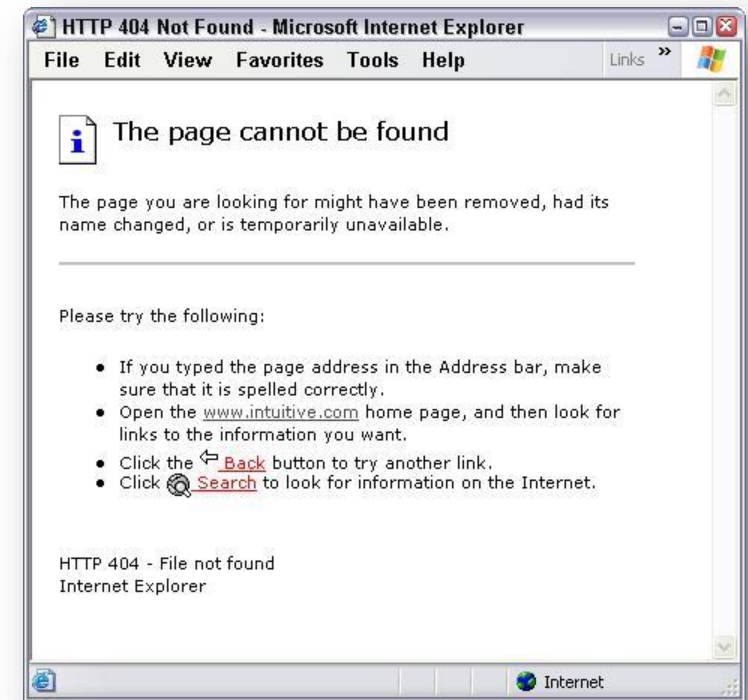
- Vadovaukitės **gerosiomis praktikomis**, skirtoms vieningų išteklių identifikatorių (angl. Uniform Resource Identifier – URI) sukūrimui bei palaikymui.
- Pasirūpinkite, kad organizacijoje būtų **aiškiai nustatyta atsakomybė** už duomenų priežiūrą.

Prieinamumo pavyzdys

Didelis prieinamumas



Mažas prieinamumas



Pilnumas

Duomenų pilnumas - tai duomenų kiekis, reikalingas taikomųjų programų palaikymui, kad jos galėtų tinkamai vykdyti joms skirtas užduotis.

Pavyzdžiui:

- Duomenų rinkinys, kuriame yra visų ministerijų išlaidų duomenys, leidžia susidaryti pilną vaizdą apie valstybės išlaidas.
- Duomenų aprašas, kuriame realiu laiku yra įtraukiamos paskutinių padarytų duomenų pakitimų datos ir laikai.

Rekomendacijos:

- Apsibrėžkite **duomenų kaupimo ir publikavimo procesus taip**, kad jie apimtų visus reikalingus taškus.
- Periodiškai **stebėkite** duomenų atnaujinimo procesus.

Pilnumo pavyzdys

Aukštas pilnumo lygis

```
: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>
.
```

Žemas pilnumo lygis

```
: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

Atitikimas

Duomenų atitikimas parodo kiek duomenys atitinka nustatytas rinkimo, publikavimo bei aprašymo taisykles bei standartus.

Pavyzdžiui:

- Duomenų rinkinys, kuris pateikia koordinates WGS84, o statistiką SDMX standartu.
- Duomenų rinkinio aprašas, parengtas pagal DCAT specifikacijas.

Rekomendacijos:

- Duomenims ir metaduomenims **taikykite dažniausiai naudojamus standartus.**
- **Jei nėra tinkamo standarto, susikurkite savo žodynus** ir juos publikuokite remiantis gerosiomis praktikomis, pvz., naudojant URI.

Duomenų atitikimo pavyzdžiai

Aukštas atitikimo lygis

```
: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>
.
```

Žemas atitikimo lygis

```
: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

Taip pat žr.:

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

Nuoseklumas

Nuoseklūs duomenys yra tokie duomenys, kuriuose nėra prieštaravimų, dėl kurių būtų sudėtinga ar netgi neįmanoma jų panaudoti.

Pavyzdžiui:

- Duomenų rinkinys, apjungiantis skirtingų šaltinių duomenis, kurie buvo apdoroti išsprendžiant prieštaravimus.
- Duomenų rinkinio aprašas, kuris neturi keleto licencijų sąlygų arba jo paskutinis pakeitimo laikas nėra ankstesnis nei aprašo sukūrimo laikas.

Rekomendacijos:

- **Apdorokite duomenis prieš juos publikuojant**, kad būtų išspręsti prieštaravimai bei ištaisytos kitos klaidos (ypač, jei duomenys gauti iš skirtingų šaltinių).

Nuoseklumo pavyzdžiai

Aukštas nuoseklumo lygis

```
: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>
.
```

Žemas nuoseklumo lygis

```
: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
.

: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: Issue date is after modification date

ERROR INCONSISTENT DATA: Licence element repeated

Patikimumas

Duomenų patikimumas priklauso nuo duomenų šaltinių ar organizacijų patikimumo.

Pavyzdžiui:

- Duomenų rinkiniai kurių duomenų šaltiniai gali būti nepriklausomai patikrinami, pvz. rinkimų rezultatai.
- Duomenų rinkinio aprašas publikuojamas valstybinės institucijos.

Rekomendacijos:

- **Rinkite duomenis iš patikimų šaltinių** arba, jei tik įmanoma, susitarkite dėl teikiamų paslaugų lygmens (angl. SLA).
- **Priskirkite tinkamus atributus**, pagal kuriuos naudotojai galėtų įvertinti duomenų patikimumą.

Patikimumo pavyzdys

Didelis patikimumas

Duomenys pateikiami iš ES leidinių biuro:

```
<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](#).

Mažas patikimumas

Duomenys pateikiami iš 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.

Linvoj/Lexvo duomenys nebūtinai yra mažesnės kokybės nei ES leidinių biuro, tačiau ES leidinių biuras yra autorizuotas šaltinis, o Linvoj ir Lexvo veikimas grįstas atskirų asmenų iniciatyva.

Apdorojamumas

Duomenų apdorojamumas parodo kokių automatizuotumo lygiu duomenys gali būti tvarkomi.

Pavyzdžiui:

- Duomenų rinkinys, kuriame koduota informacija pagrįsta viešai prieinamais žodynais bei kodų sąrašais.
- Duomenų rinkinio aprašas, kuriame datos pateikiamos W3C formatu (pvz., 2014-03-11), o ne tekstiniu pavidalu (pvz., 2014 m. kovo 11 d.).

Rekomendacijos:

- **Nustatykite duomenims terminų ir kodų šaltinius naudojamus kompiuteriu skaitomam formatui.**
- **Duomenų sintaksei taikykite rekomendacijas** pateikiamas bendruosiuose standartuose bei specifikacijose.

Duomenų apdorojimo pavyzdžiai

Aukštas apdorojimo lygis

```
: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"  
  .
```

Žemas apdorojimo lygis

```
: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 "January 1, 2013"  
  .
```

Aktualumas

Duomenų aktualumas parodo, kiek jais pateikiama informacija atitinka panaudojimo srities poreikį.

Pavyzdžiui:

- Duomenų rinkinys, kuriame temperatūrų informacija pateikta sveiko skaičiaus tikslumu, kadangi bus naudojama tolimesniems klimato skaičiavimams.
- Duomenų rinkinys kur temperatūrų informacija pateikiama tūkstantinės skaičiaus dalies tikslumu, kadangi bus naudojama cheminėms reakcijoms tirti.

Rekomendacijos:

- **Suderinti** duomenų **apimtį bei detalumą** pagal jų paskirtį, atsižvelgiant į esamus laiko, valiutos ir kitų veiksnių apribojimus.
- **Apsvarstyti galimas** duomenų **panaudojimo** galimybes ateityje.

Aktualumas pavyzdys

Aktualūs duomenys

Table to determine special tax on motor vehicles based on emission CO₂

	Engine displacement (cm ³)	Fuel type	CO ₂ (g/km)	Tax (%)
Car type 1	900	Gasoline	90	0
Car type 2	1100	Gasoline	120	5
Car type 3	1300	Gasoline	125	5
Car type 4	1400	Gasoline	150	5
Car type 5	1800	Diesel	180	10
Car type 6	2200	Diesel	190	10
Car type 7	2500	Gasoline	210	15

Mažiau aktualūs duomenys

Table to determine special tax on motor vehicles based on emission CO₂

	Engine displacement (cm ³)	Fuel type	CO ₂ (g/km)	Weight (kg)	Tax (%)
Car type 1	900	Gasoline	90	750	0
Car type 2	1100	Gasoline	120	1000	5
Car type 3	1300	Gasoline	125	1200	5
Car type 4	1400	Gasoline	150	1200	5
Car type 5	1800	Diesel	180	1700	10
Car type 6	2200	Diesel	190	1600	10
Car type 7	2500	Gasoline	210	1900	15

Weight is not a relevant data item

Savalaikiškumas

Duomenų savalaikiškumas parodo, kiek duomenys tinkamai atspindi dabartinę situaciją ir įvykius, ir kaip greitai yra pasiekiamos tokių duomenų naujausios versijos.

Pavyzdžiui:

- Duomenų rinkinys, kuriame realiu laiku kaupiami ir kas kelias minutes atnaujinami kelių eismo sąlygų duomenys.
- Duomenų rinkinio, kuriame pateikta metinė nusikalstamumo statistika, aprašas, kuris pateikiamas po kelių dienų nuo duomenų rinkinio publikavimo.

Rekomendacijos:

- **Nustatykite** duomenų **atnaujinimo dažnumą**, atsižvelgdami į duomenų prigimtį ir jų panaudojimo paskirtį.
- Įsitikinkite, kad egzistuoja **tinkami įrankiai bei procesai** užtikrinantys duomenų atnaujinimą.

Duomenų savalaikiškumo pavyzdys

Savalaikiški duomenys

Weather report of 1 July 2013, 12:00

	Temp. °C	Humidity %	Wind direction	Wind speed km/h
Station 1	18.1	60	WSW	18
Station 2	17.5	59	WSW	20
Station 3	18.2	55	SW	22
Station 4	19.0	62	SW	18
Station 5	18.0	65	WSW	19
Station 6	18.2	63	SSW	21
Station 7	17.9	61	SW	22

Published 1 July 2013, 12:05

Nesavalaikiški duomenys

Weather report of 1 July 2013, 12:00

	Temp. °C	Humidity %	Wind direction	Wind speed km/h
Station 1	18.1	60	WSW	18
Station 2	17.5	59	WSW	20
Station 3	18.2	55	SW	22
Station 4	19.0	62	SW	18
Station 5	18.0	65	WSW	19
Station 6	18.2	63	SSW	21
Station 7	17.9	61	SW	22

Published 1 July 2013, 20:00

Gerosios praktikos

Gerųjų praktikų, publikuojant aukštos kokybės duomenis ir metaduomenis, apžvalga.

W3C: Geroji praktika publikuojant susietuosius atvirus duomenis

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.

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

Opquast: 72 atvirų duomenų gerosios praktikos

Keli pavyzdžiai

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 license
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)

Taip pat žr.:
<http://checklists.opquast.com/en/opendata>

Gerųjų praktikų apibendrinimas

- **Teikti tinkamus** duomenų **aprašus** (t.y. metaduomenis).
- **Naudoti standartinius žodynus** metaduomenims ir duomenims, jei tik tokie žodynai egzistuoja.
- **Nurodyti licenciją**, pagal kurią duomenys gali būti pakartotinai naudojami.
- **Laikytis teisinių reikalavimų**, susijusių su asmeniniais bei slaptais duomenimis.
- Duomenis bei metaduomenis **teikti, remiantis susietų duomenų principais**, naudojant URI.
- **Teikti informaciją apie duomenų šaltinius**.

Metaduomenų bei duomenų palaikymas yra būtinas!

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

Išvados

- Duomenų kokybę apibrėžia naudotojų galimybė naudotis šiais duomenimis.
- Metaduomenys – tai „duomenys apie duomenis“, t.y. metaduomenys yra duomenų tipas.
 - Tie patys kokybės principai tiek duomenims, tiek metaduomenims taikomi vienodai.
- Duomenų kokybę aprėpia daug aspektų ir tai nėra vien tik duomenų teisingumas.
 - Tikslumas, prieinamumas, pilnumas, atitikimas, nuoseklumas, patikimumas, apdorojimas, aktualumas, naujumas.

Ačiū!
...ir dabar JŪSŲ klausimai

Klausimai



<http://www.visualpharm.com>

Kokie veiksniai Jūsų nuomone daro didžiausią įtaką metaduomenų kokybei?



<http://www.visualpharm.com>

Kokybės gerinimas gali pareikalauti laiko ir išteklių. Koku mastu Jūsų organizacija yra pasiruošusi investuoti į metaduomenų kokybę?



<http://www.visualpharm.com>

Ar galėtumėte įvardyti Jums žinomų aukštos kokybės metaduomenų aprašo pavyzdžių?

This presentation has been created by Open Data Support

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.

Authors:

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

Šaltiniai

5 skaidrė:

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

6 skaidrė:

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

8-26 skaidrė:

- 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>

10 skaidrė:

- [Garytx at en.wikipedia](#)

11 skaidrė:

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

13 skaidrė:

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

28 skaidrė:

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

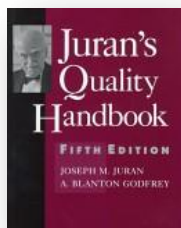
29 skaidrė:

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

Papildomam skaitymui



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



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

Susiję projektai bei iniciatyvos



Gerosios praktikos publikuojant susietuosius duomenis.
<https://dvcs.w3.org/hg/gld/raw-file/default/bp/index.html>



OPQUAST. Atvirų duomenų gerosios praktikos.
<http://checklists.opquast.com/en/opendata>

Prisijunkite prie mūsų...

Mus rasite



Open Data Support

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



Open Data Support

<http://goo.gl/y9ZZI>

Sekite mus



@OpenDataSupport

Prie mūsų prisijungsite



joinup

<http://www.opendatasupport.eu>

Susisiekite su mumis

contact@opendatasupport.eu

