



**OPEN
DATA
SUPPORT**

Õpimoodul 2.2

Avaandmete &
metaandmete
kvaliteet

pwc

PwC firms help organisations and individuals create the value they're looking for. We're a network of firms in 158 countries with close to 180,000 people who are committed to delivering quality in assurance, tax and advisory services. Tell us what matters to you and find out more by visiting us at www.pwc.com. PwC refers to the PwC network and/or one or more of its member firms, each of which is a separate legal entity. Please see www.pwc.com/structure for further details.

Esitluse metaandmed

Avaandmete tugi on loodud Euroopa Komisjoni poolt SMART 2012/0107 'Lot 2' alusel: Teenuste osutamine andmete avaldamiseks, juurdepääsu loomiseks, avalikele andmetele juurdepääsuks ja taaskasutamiseks Euroopa Liidu lõikes läbi olemasolevate avaandmete portaalide (Leping nr. 30-CE-0530965/00-17).

© 2013 Euroopa Komisjon

See esitus on loodud PwC poolt

Autorid:

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

Täpsustused

1. Esitluses välja toodud seisukohad kuuluvad ainult esitluse koostanud autoritele ja neid ei või mitte mingil juhul interpreteerida Euroopa Komisjoni ametlike seisukohtadena.

Euroopa Komisjon ei taga presentatsioonis välja toodud informatsiooni täpsust, ega võta vastutust selle kasutamise eest.

Iga viide presentatsioonis spetsiifilisele tootele, täpsustusele, protsessile, teenusele, ärinimele, kaubamärgile või tootjale ei viita Euroopa Komisjoni heakskiidule, soovitudele, ega eelistustele.

Autorite poolt on hangitud kõik vajalikud load käsikirjade, illustratsioonide, kaartide ja graafikute kasutamiseks, kui just autor või tema seaduslik esindaja ei oma juba antud intellektuaalse vara kasutusõigust.

2. Esitus on hoolikalt koostatud PwC poolt, kuid puudub garantii esitluses esitatud informatsiooni täpsuse ja täielikkuse osas. PwC ei vastuta esitluses esitatud informatsiooni põhjal tehtud otsuste tagajärgede, ega muu kahju eest. Presentatsioonis esitatud informatsioon on üldist laadi ja mõeldud ainult üldise huvi kohta juhtnööride andmiseks. Esitus ei asenda mitte ühegi küsimuse korral antavat professionaalset nõu. Ükski esitluse lugeja ei tohiks langetada otsuseid esitluses sisalduvate andmete põhjal vastava professionaalse nõu küsimiseta.

Õpieesmärgid

Selle õpimooduli lõpuks peaks teil olema arusaam järgnevast:

- Mida avaandmete kvaliteet tähendab.
- Avaandmete kvaliteeti määravad tegurid ja kriteeriumid.
- Kvaliteetsete avaandmete avaldamise head tavad.

Sisu

See moodul sisaldab...

- Andmekvaliteedi definitsiooni;
- Ülevaadet andmete ja metaandmete kvaliteedi dimensioonidest;
- Valikut parimatest praktikatest kvaliteetsete andmete ja metaandmete avaldamiseks.

Mis on andmete (ja metaandmete) kvaliteet?

Andmed on kvaliteetsed, kui nad on kõlblikud nende kavandatud kasutuseks protsessis, otsuste tegemises ja planeerimises.

Või veel spetsiifilisemalt:

“Kvaliteetsed andmed on täpsed, kättesaadavad, täielikud, reeglipärased, järjepidevad, usaldusväärsed, töödeldavad, olulised ja ajakohased“

Mis on metaandmed?

“Metaandmed on struktureeritud informatsioon, mis kirjeldab, selgitab, lokaliseerib või teeb muul viisil kergemaks teabeallika välja otsimise, kasutamise või haldamise. Metaandmeid nimetatakse tihti andmeteks andmete kohta.” -- National Information Standards Organization

- Me täheldame, et metaandmed on üks andmete liik.
- Samad kvaliteedimäärajad rakenduvad andmetele ja metaandmetele ühtviisi.

Andmekvaliteedi dimensioonid

Mis on peamised dimensioonid, mida peab arvesse võtma, et edastada kvaliteetseid (meta)andmeid?

Andmekvaliteedi dimensioonid

- **Täpsus:** kas andmed kajastavad reaalsed objekti või nähtust õigesti?
- **Metoodiline järjepidevus:** kas andmetes ei ole vastuolusid?
- **Kättesaadavus:** kas andmetele pääseb ligi praegu ja ka mõne aja pärast?
- **Täielikkus:** kas andmetes sisalduvad kõik elemendid, mis kirjeldavad seda objekti või nähtust?
- **Reeglipärasus:** kas andmed järgivad üldtunnustatud standardeid?
- **Usaldusväärsus:** kas andmed põhinevad usaldusväärsetel allikatel?
- **Töödeldavus:** kas andmed on masinloetavad?
- **Olulisus:** kas andmete hulgas on vajalik kogus andmeid?
- **Ajakohasus:** kas andmed kajastavad tegelikku situatsiooni ja avaldatakse võimalikult kiiresti?

Täpsus

Andmete täpsus näitab seda, kui suurel määral kajastavad nad õigesti reaalselt objekti, situatsiooni või nähtust.

Näiteks:

- Ilmastikutingimuste õige hindamine (temperatuur, sademed).
- Andmekogumi taaskasutamise tingimuste õige esitlus.

Soovitused:

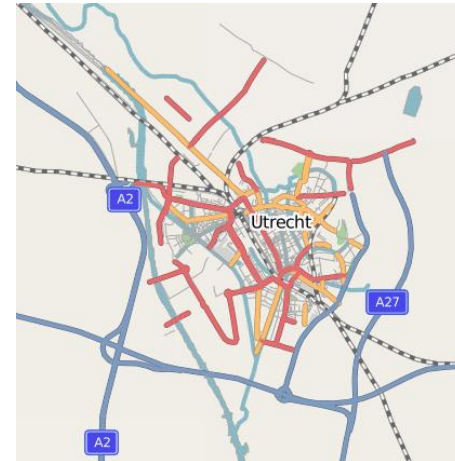
- Tasakaalusta oma andmete täpsus vs kulu infosüsteemi kontekstist lähtuvalt; see peab olema sihtotstarbeliseks kasutamiseks piisavalt hea.
- Täpsuse hoidmiseks kindlusta organisatsiooni pühendumus ning protseduuride investeeringud.

Täpsuse näited

Täpsem

Ebatäpsem

OpenStreetMap, Utrecht, Holland
(2011 vs. 2007)



Metoodiline järjepidevus

Andmed on järjepidevad, kui neis ei sisaldu vastuolusid, mis võivad muuta andmete kasutamise keeruliseks või võimatuks.

Näiteks:

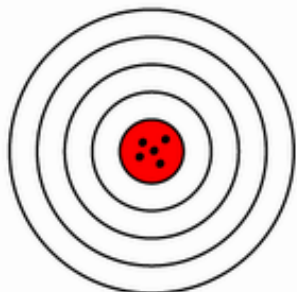
- Andmekogum, mis kombineerib andmeid erinevatest allikatest, on kontrollfunktsioonidega töödeldud, eesmärgiga elimineerida vastuolulised andmed.
- Andmekogum, mis ei sisalda andmete kohta erinevaid litsentse või kus andmete viimane muutmisaeg ei ole varasem kui nende loomisaeg.

Soovitused:

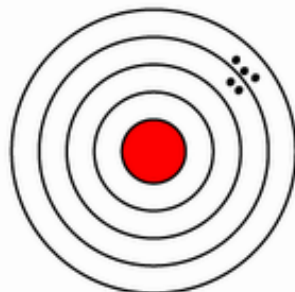
- **Töötle kõik andmed enne avalikustamist**, leidmaks vastuolulised andmed ning teised esinevad vead (eriti kui andmed on koondatud erinevatest allikatest).

Metoodilise järjepidevuse näide

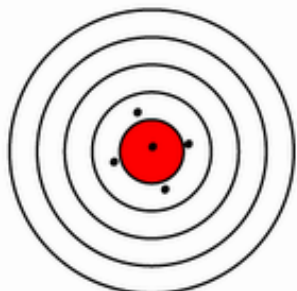
Suur järjepidevus



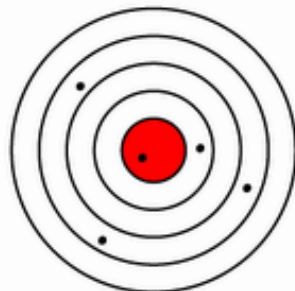
Accurate & Consistent



Not Accurate
but Consistent



Accurate but
Not Consistent



Not Accurate &
Not Consistent

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

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

Väiksem järjepidevus

```
: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

Kättesaadavus

Andmete kättesaadavus on seotud neile juurdepääsu saamise määraga; see hõlmab samuti pikaajalise järjekindlusega kogutud andmeid.

Näiteks:

- Andmekogum, mis on tuvastatud http poolt: URI, mis avab püsivalt paremad allikad (ning ei anna vastuseks „404 Lehekülge ei leitud“).
- Andmekogumi kirjeldus, mis on lisatud andmebaasi otsingumootorisse.

Soovitused:

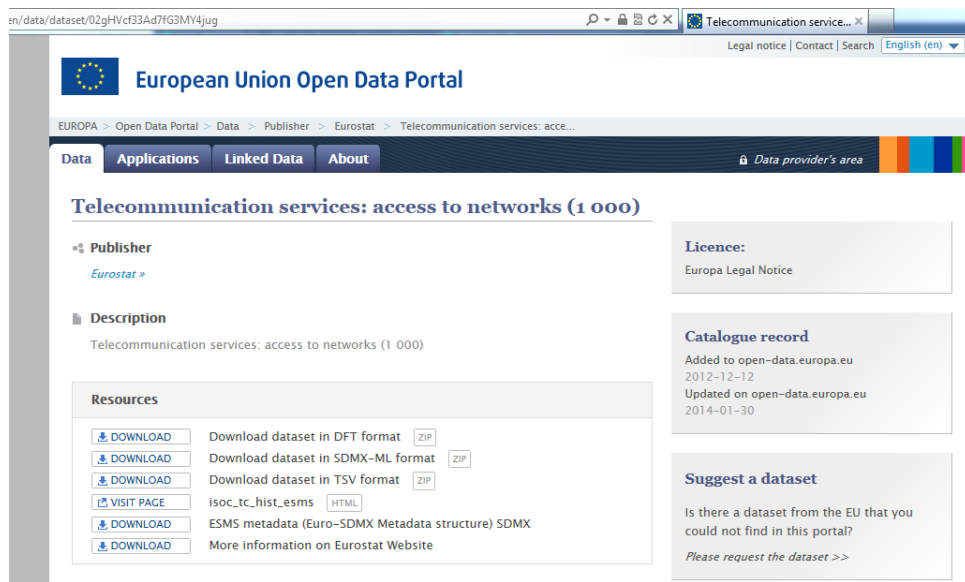
- Järgi URI määramise ja hooldamise **parimaid praktikaid**.
- Kanna hoolt, et andmete korrashoiu **kohustus** on organisatsioonis **selgelt delegeeritud**.

Vaata lisaks:

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

Kättesaadavuse näited

Kõrge kättesaadavus



The screenshot shows a web browser window displaying the European Union Open Data Portal. The page title is "Telecommunication services: access to networks (1 000)". The publisher is Eurostat. The description is "Telecommunication services: access to networks (1 000)". The resources section includes several download options: "Download dataset in DFT format" (ZIP), "Download dataset in SDMX-ML format" (ZIP), "Download dataset in TSV format" (ZIP), "isoc_tc_hist_esms" (HTML), "ESMS metadata (Euro-SDMX Metadata structure) SDMX", and "More information on Eurostat Website". The page also includes a "Licence" section (Europa Legal Notice), a "Catalogue record" (Added to open-data.europa.eu 2012-12-12, Updated on open-data.europa.eu 2014-01-30), and a "Suggest a dataset" section.

Madal kättesaadavus

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](#)

Täielikkus

Andmete täielikkus tähendab seda, mil määral see sisaldab andmeühikuid, mida on vaja protsessi ja seda toetava tarkvara eesmärkide täitmiseks.

Näiteks:

- Andmekogum, mis sisaldab ministeeriumite kulude andmeid, võimaldab anda täieliku ülevaate valitsuse kuludest.
- Andmete kirjeldus, mis on loodud reaalajas ning sisaldab viimaste muudatuste kuupäeva ja kellaaega.

Soovitused:

- **Kavanda hõive ja väljastamise protsess** vajalike andmeühikute lisamiseks.
- **Jälgi** pidevalt andmete uuendusprotseduure.

Täielikkuse näited

Kõrge täielikkus

```
: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>
.
```

Madal täielikkus

```
: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

Vastavus

Andmete vastavus tähendab seda, mil määral see jälgib reeglite ja standardite hulka andmehõiveks, avalikustamiseks ja kirjeldamiseks.

Näiteks:

- Andmekogum, mis väljastab koordinaadid WGS84 ja statistika SDMX vormingus.
- Andmekogumi kirjeldus vastavalt DCAT sõnastikule.

Soovitused:

- Rakenda andmetele ja metaandmetele **enim kasutatud standardeid** teatud valdkonnas.
- **Kui ükski standard ei ole kättesaadav, määratle kohalik sõnavara**, kuid avalikusta enda sõnavara vastavalt parimatele tavadele (nt viidatavad URId).

Vastavuse näited

Suure vastavusega

```
: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>
.
```

Madala vastavusega

```
: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

Vaata lisaks:

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

Usaldusväärsus

Andmete usaldusväärsus tähendab seda, mil määral see põhineb usaldusväärsetel allikatel või mil määral on see edastatud usaldusväärsete organisatsioonide poolt.

Näited:

- Andmekogum, mis pärineb sõltumatult auditeeritud protsessidest, nt valimistulemused või parlamentaarsed toimingud.
- Andmekogumi kirjeldus, mis on väljastatud valitsusasutuse poolt.

Soovitused:

- Kus võimalik ja kohane, **loo andmed allikatest, mida saab usaldada** või mis on kaetud teenustaseme lepingutega (*Service Level Agreements*).
- **Lisa vajalikud atribuudid**, et kasutajad saaksid otsustada, kas neid andmeid saab või ei saa usaldada.

Usaldusväärsuse näide

Suur usaldusväärsus

Andmed on võetud EL-i Väljaannete osakonnast:

```
<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](#).

Madal usaldusväärsus

Andmed on võetud Lexvo-st:

```
- <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 Inggris</rdfs:label>
  <rdfs:label xml:lang="af" rdf:datatype="xsd:string">Engels</rdfs:label>
  <rdfs:label xml:lang="agg" rdf:datatype="xsd:string">Kingele</rdfs:label>
  <rdfs:label xml:lang="aii" rdf:datatype="xsd:string">ܐܢܓܠܝܩܝܬܝܢ</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">ግብጽያዊት ቋንቋ</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 andmed ei sa olla väiksema kvaliteediga kui Väljaannete osakonna andmed, kuid Väljaannete osakond on autoriteetne allikas, samal ajal kui Linvoj ja Lexvo on individuaalsed algatused.

Töödeldavus

Andmete töödeldavus tähendab seda, mil määral see võib olla mõistetav ja töödeldav automaatsete protsesside poolt.

Näiteks:

- Andmekogum, mis sisaldab klassifitseeritud informatsiooni, mis põhineb avalikult kättesaadavatel kontrollitud sõnastikel ja klassifikaatoritel.
- Andmekogumi kirjeldus, mis väljastab kuupäevad W3C kuupäeva ja kellaaja formaadis ning aja formaadis (nt 2013-06-01), mitte vabateksti formaadis (nt 1.juuni 2013).

Soovitused:

- **Tuvastada terminoloogia ja klassifikaatorid**, mida on kasutatud valdkonna andmete masinloetavas vormingus.
- **Rakenda soovitused** andmete **süntaksis** üldistes standardites.

Töödeldavuse näide

Kõrge töödeldavus

```
▼<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>
```

Madal töödeldavus

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.

Olulisus

Andmete olulisus tähendab seda, mil määral see sisaldab vajalikku informatsiooni protsessi eesmärkide toetamiseks.

Näiteks:

- Andmekogum, mis sisaldab ühe kraadini ümardatud temperatuuri mõõdikut Celsiuse skaalal kliimaandmete arvutamiseks; andmekogum tuhandiku täpsusega samal skaalal keemiliste reaktsioonide mõõtmiseks.
- Andmekogum, mis sisaldab ainult ajutisi andmeid sel hetkel kui neid on vaja töödelda.

Soovitused:

- **Sobita omavahel** andmete **ulatus ja detailsus** selle kavatsatud kasutamise jooksul, võttes arvesse aja ja raha piirangud.
- Siiski, **kaalu võimalikke tulevasi** andmete kasutamisi.

Olulisuse näide

Kõrge olulisusega

	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

Madala olulisusega

	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

Õigeaegsus

Andmete õigeaegsus tähendab seda, mil määral see kajastab õigesti objekti või sündmuse hetkeseisu ning mil määral, millised andmed (viimane versioon) on kättesaadavad ilma viivitusega.

Näiteks:

- Andmekogum, mis sisaldab liikluse andmeid reaalsajas ning mida uuendatakse iga paari minuti tagant.
- Andmekogumi kirjeldus, mis sisaldab iga-aastast kuritegevuse statistikat, mis on kättesaadav andmekogumi väljastamisest mõne päeva jooksul.

Soovitused:

- **Kohanda** andmete **uuendamissagedusi vastavalt andmete laadile ja selle kasutusotstarbele.**
- Veendu selles, et **meetod ja tööriistad on olemas**, et uuendusi saaks toetada.

Õigeaegsuse näide

Kõrge õigeaegsusega

NOAA's National Weather Service 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.

Madala õigeaegsusega



Parim praktika

Kõrge kvaliteediga andmete ja metaandmete väljastamise parimad praktikad

W3C: Lingitud avaandmete väljastamise praktika

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.

Vaata lisaks:

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

Opquast: 72 avaandmete praktikat

Mõned näited

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)

Vaata lisaks:

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

Millised on parima praktika ühisjooned?

- **Tagada sobiv** andmete **kirjeldus** (nt metaandmed).
- **Kasutada standardseid sõnastikke** andmetel ja metaandmetel, millal iganes selline sõnavara eksisteerib.
- **Määrata** litsents vastavalt sellele, milliseid andmeid võidakse korduvkasutada.
- **Jälgida juriidilisi nõudeid** isikuandmete kaitse ja muude tundlike andmete suhtes.
- **Esitada** metaandmeid ja andmeid **vastavalt linkandmete põhimõtetele** kasutades **püsivaid URI-sid** asjade tuvastamiseks. **Esitleda andmete allikad.**

Metaandmete ja andmete hooldus on kriitiline!

Vaata lisaks:

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

Kokkuvõte

- Andmete kvaliteedi määrab selle sobivus korduvkasutamiseks.
- Metaandmed on “andmed andmete kohta”, metaandmed on andmete tüüp.
 - Andmetele ja metaandmetele kehtivad samad kvaliteedinõuded.
- Andmete kvaliteedil on mitmeid dimensioone ja seda on rohkem kui andmed on korrektsed.
 - Täpsus, kättesaadavus, täielikkus, vastavus, järjepidevus, usaldusväärsus, hea töödeldavus, asjakohasus, õigeaegsus.

Diskussioon



<http://www.visualpharm.com>

Millised tegurid soodustavad andmete kvaliteeti?



<http://www.visualpharm.com>

Kvaliteedi parandamine võib nõuda aega ja ressursse. Mil määral oleks Teie organisatsioon nõus investeerima andmete kvaliteeti?

Internetiküsitluse leiad [siit!](#)

Tänään kuulamast!
...kysymusi?

Viited

Lk 5:

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

Lk 6:

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

Lk 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>

Lk 10:

- [Garytx](#) at [en.wikipedia](#)

Lk 11:

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

Lk 13:

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

Lk 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>

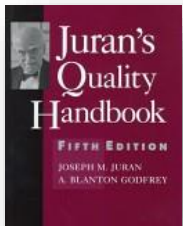
Lk 29:

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

Lisalugemist



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



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

Seotud projektid ja algatused



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

Ole osa meie meeskonnast...

Leia meid



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

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



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

<http://goo.gl/y9ZZI>

Järgi meid



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

Liitu meiega



joinup

<http://www.opendatasupport.eu>

Kontakt

contact@opendatasupport.eu