Introducing Wikibase
Use cases, ecosystem and data model

Georgina Burnett, Mohammed Sadat Abdulai, Phil Stacey, Eugene Morozov
# Open Source KO and Reference Data Stack

<table>
<thead>
<tr>
<th>Taxonomy Management</th>
<th>Taxonomy Publication</th>
<th>Subject Indexing</th>
<th>Document Management</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protégé</td>
<td>Skosmos</td>
<td>annif</td>
<td>ejprints</td>
</tr>
<tr>
<td></td>
<td>jekyll</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ShowVoc</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Linked Data Storage**

- rdf4j
- Yago
By the end of this workshop you’ll be familiar with features of Wikibase and have some hands-on experience with it.
Wikibase introduction
by Georgina Burnett, Mohammed Sadat Abdulai
Wikibase hands on
by Eugene Morozov and Phil Stacey
Use case of collaborative threat modelling for data risks.
How are we going to collaborate?

- Define things related to data risk management and build taxonomy of data quality dimensions
- Define threats and link them to data quality dimensions
- List threats that we have defined
And in Wikibase terms...

1. **Create base set of Items (Principle, Policy, Threat, Vulnerability, etc.) and Properties**
2. **Create Threat Items and specify Property values for them**
3. **Define threats and link them to data quality dimensions**
4. **List threats that we have defined**
5. **Show items linked to Threat Item**

Define things related to data risk management and build taxonomy of data quality dimensions.
Picking CPG 235 Managing Data Risk

- Principle
- Policy
- Standard
- Guideline
- Procedure
- Control
- Threat
- Vulnerability

Items marked green are already created in our Wikibase instance.

This is what we are interested in most for this exercise.

Source: Australian Prudential Regulation Authority CPG 235 – Managing Data Risk. September 2013. Used under Creative Commons Attribution 3.0 Australia Licence (CCBY 3.0). APRA does not endorse this derivative work.
This taxonomy is already created in our Wikibase instance.

Source: Australian Prudential Regulation Authority CPG 235 – Managing Data Risk. September 2013. Used under Creative Commons Attribution 3.0 Australia Licence (CCBY 3.0). APRA does not endorse this derivative work.
Wikibase instance available for the workshop
https://iskouk-workshop.wiki.opencura.com/wiki/Main_Page

You can register with your email address to contribute.

No login is required to just follow along.

Data will be deleted after the workshop.
Create account

Username
Enter your username

Password
Enter a password

It is recommended to use a unique password that you are not using on any other website.

Confirm password
Enter password again

Email address
Enter your email address

Real name (optional)

Real name is optional. If provided, it may be used to give you attribution for your work.

To protect the wiki against automated account creation, we kindly ask you to solve the following CAPTCHA:

I'm not a robot

Create your account

isko.uk November 2021 is made by people like you.

22 edits
10 pages
0 recent contributors
Create a new Item

Make sure to check if the Item already exists!
You should create a label and a description for all new items.

By clicking "Create", you agree to the terms of use, and you irrevocably agree to release your contribution under the [None yet set].

Create a new Item

Language:

en

Label:

Data Quality Dimension

Description:

Data quality can be assessed using a range of dimensions. The relevance of each of these dimensions will vary

Aliases, pipe-separated:

enter some aliases in English

Create
Data Quality Dimension (Q13)

Dimensions typically considered in the assessment of data quality.

<table>
<thead>
<tr>
<th>Language</th>
<th>Label</th>
<th>Description</th>
<th>Also known as</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>Data Quality Dimension</td>
<td>Dimensions typically considered in the assessment of data quality.</td>
<td></td>
</tr>
<tr>
<td>British English</td>
<td>No label defined</td>
<td>No description defined</td>
<td></td>
</tr>
</tbody>
</table>

Statements

This page was last edited on 21 November 2021, at 15:44.
Editing Item talk:Q13

Data quality can be assessed using a range of dimensions. The relevance of each of these dimensions will vary depending upon the nature of the data. [[Item:Q7|Accuracy]], [[Item:Q8|Completeness]], [[Item:Q9|Consistency]], [[Item:Q10|Timeliness]], [[Item:Q11|Availability]] and [[Item:Q12|Fitness for Use]].

Item talk:Q13

Data quality can be assessed using a range of dimensions. The relevance of each of these dimensions will vary depending upon the nature of the data. Accuracy, Completeness, Consistency, Timeliness, Availability and Fitness for Use.

You need to define properties from scratch (but can map them to Wikidata such as instance of mapping to wdt:P31)
List of Properties

Get a list of Properties by data type

Data type:

- All data types

Find

Showing below up to 2 results starting with #.

View [previous 50] [next 50] [20] [50] [100] [250] [500]

1. instance of (P1)
2. broader (P2)

View [previous 50] [next 50] [20] [50] [100] [250] [500]
Accuracy (Q7)

The degree to which data is error free and aligns with what it represents.

<table>
<thead>
<tr>
<th>Language</th>
<th>Label</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>Accuracy</td>
<td>The degree to which data is error free and aligns with what it represents.</td>
</tr>
<tr>
<td>British English</td>
<td>No label defined</td>
<td>No description defined</td>
</tr>
</tbody>
</table>

Statements

broader

- Data Quality Dimension
  - Data Quality Dimension
    - Dimensions typically considered in the assessment of data quality.
      - 0 references
  - add qualifier
  - add reference
  - add value
  - add statement

This page was last edited on 21 November 2021, at 16:08.
Create a new Item that is an instance of (P1) Threat (Q14) and related to (P3) Availability (Q11)
Pages that link to "Item:Q13"

- Data Quality Dimension (Q13)

What links here

| Page: Item:Q13 | Namespace: all | Invert selection | Go |

Filters

- Hide transclusions
- Hide links
- Hide redirects

The following pages link to Data Quality Dimension (Q13):

Displayed 6 items.

View (previous 50) | next 50) (20 | 50 | 100 | 250 | 500)
- Accuracy (20) (~ links)
- Completeness (20) (~ links)
- Consistency (20) (~ links)
- Timeliness (20) (~ links)
- Availability (20) (~ links)
- Fitness for Use (20) (~ links)

View (previous 50) | next 50) (20 | 50 | 100 | 250 | 500)
Pages that link to "Item:Q14"

What links here... Page: Item:Q14... Namespace: all... Invert selection: Go...

Filters... Hide transclusions | Hide links | Hide redirects

The following pages link to Threat (Q14):

Displayed 4 items.

View (previous 50 | next 50) (20 | 50 | 100 | 250 | 500)

* Data store is down (Q21) (— links)
* Data got accidentally erased by buggy software (Q17) (— links)
* Hacker (Q22) (— links)
* Data kaBOOM oh no! (Q21) (— links)

View (previous 50 | next 50) (20 | 50 | 100 | 250 | 500)
References

- Wikibase: https://wikiba.se/
- Wikibase data model primer: https://www.mediawiki.org/wiki/Wikibase/DataModel/Primer
- Workshop Wikibase instance: https://iskouk-workshop.wiki.opencura.com/wiki/Main_Page
- ISKO UK: https://www.iskouk.org/