Benefits of Publishing the Norwegian Petroleum Directorate’s FactPages as Linked Open Data

Martin G. Skjæveland
martige@ifi.uio.no

NIK 2013
19 November 2013
Outline

NPD and NPD FactPages

Linked Open Data

Converting NPD’s FactPages to Semantic Web Data

Examples
  Map Visualization
  Query Formulation
NPD and NPD FactPages

Linked Open Data

Converting NPD’s FactPages to Semantic Web Data

Examples
  Map Visualization
  Query Formulation
Norwegian Petroleum Directorate (NPD)

What is it?

- A governmental specialist directorate and administrative body
- Reports to the Ministry of Petroleum and Energy
- Main objective is to look after the national interests in the energy sector
- One of four functions:

  *The NPD has a national responsibility for data from the Norwegian continental shelf. The NPD’s data, overview and analyses constitute a crucial factual basis on which the activities are founded*
The FactPages contain data about petroleum activities on the NCS

Data is collected from companies that operate on the NCS

Important functions:
- forms the basis for the authorities’ planning of future activity and their judgement of existing activity
- makes companies share information
- provides information to the general public

Acts as a national reference data library for the activities on the NCS
Demonstration

Selecting Field...
- lists available views and subviews.
- Selecting Page view → All...
- lists recorded members for the view. Selecting TROLL...
- displays a (page view) report, i.e., all data about the field Troll.

Download CSV, XML, Excel

Shortcomings to the NPD FactPages:
- No query interface
- Arbitrary queries not possible
- Only predefined categories and views
- Poor identifiers
- Difficult information retrieval
- Integration "impossible"
Selecting Field ... lists available views and subviews. Selecting Page view → All ... lists recorded members for the view. Selecting TROLL ... displays a (page view) report, i.e., all data about the field Troll.

6 / 20
NPD FactPages
Demonstration

Selecting Field ...

11 main categories

Selecting Page view → All ...
... lists available views and subviews.
Selecting Page view → TROLL ...
... displays a (page view) report, i.e., all data, about the field Troll.

Shortcomings to the NPD FactPages:
• No query interface
• Arbitrary queries not possible,
• only predefined categories and views
• Poor identifiers
• Difficult information retrieval
• Integration “impossible”
NPD FactPages
Demonstration

Selecting Field ...

11 main categories

... lists available views and subviews.
Selecting Page view → All ...

Shortcomings to the NPD FactPages:
• No query interface
• Arbitrary queries not possible,
• only predefined categories and views
• Poor identifiers
• Difficult information retrieval
• Integration "impossible"
Selecting Field ... lists available views and subviews. Selecting Page view → All ...

... lists recorded members for the view. Selecting TROLL ...

11 main categories

- Attributes
- Page view
- All
- PDO approved
- Producing
- Shut down
- Ordered by...
- Table view
- Overview
- Status
- Operators
- Owners
- Licensees
- Production
- Reserves
- Investments
- Description
- Geography

... lists available views and subviews. Selecting Page view → All ...

... lists recorded members for the view. Selecting TROLL ...

Shortcomings to the NPD FactPages:
- No query interface
- Arbitrary queries not possible,
- only predefined categories and views
- Poor identifiers
- Difficult information retrieval
- Integration “impossible”
NPD FactPages
Demonstration

Selecting Field ...

11 main categories

- Attributes
- Page view
- All
- PDO approved
- Producing
- Shut down
- Ordered by...
- Table view
- Overview
- Status
- Operators
- Owners
- Licensees
- Production
- Reserves
- Investments
- Description
- Geography

... lists available views and subviews.
Selecting Page view → All ...

... displays a (page view) report, i.e., all data about the field Troll

Download CSV, XML, Excel

Shortcomings to the NPD FactPages:
• No query interface
• Arbitrary queries not possible,
• only predefined categories and views
• Poor identifiers
• Difficult information retrieval
• Integration "impossible"
NPD FactPages
Demonstration

11 main categories

Selecting Field ...

11 main categories

Download/Export

Download CSV, XML, Excel

Shortcomings to the NPD FactPages:
• No query interface
• Arbitrary queries not possible,
• only predefined categories and views
• Poor identifiers
• Difficult information retrieval
• Integration “impossible”
Shortcomings to the NPD FactPages:

- No query interface
- Arbitrary queries not possible,
- only predefined categories and views
- Poor identifiers
- Difficult information retrieval
- Integration “impossible”
Outline

NPD and NPD FactPages

Linked Open Data

Converting NPD’s FactPages to Semantic Web Data

Examples
  Map Visualization
  Query Formulation
5-star Linked Open Data

Tim Berners-Lee’s 5 Linked Open Data stars
Tim Berners-Lee’s 5 Linked Open Data stars:

- Available on the web with an open licence
5-star Linked Open Data

Tim Berners-Lee’s 5 Linked Open Data stars:

⭐ Available on the web with an open licence

⭐⭐ + machine-readable structured data
Tim Berners-Lee’s 5 Linked Open Data stars:

★ Available on the web with an open licence
★★ + machine-readable structured data
★★★ + non-proprietary format
5-star Linked Open Data

Tim Berners-Lee’s 5 Linked Open Data stars:

★   Available on the web with an open licence
★★  +  machine-readable structured data
★★★ +  non-proprietary format
★★★★ +  use W3C standards (URI, RDF) to identify and serve
★★★★★ +  provide query interface (SPARQL)
★★★★★★ +  define vocabulary semantics (OWL)
5-star Linked Open Data

Tim Berners-Lee’s 5 Linked Open Data stars:

⭐️ Available on the web with an open licence
⭐⭐ + machine-readable structured data
⭐⭐⭐ + non-proprietary format
⭐⭐⭐⭐ + use W3C standards (URI, RDF) to identify and serve
⭐⭐⭐⭐⭐ + link to other data
Tim Berners-Lee’s 5 Linked Open Data stars:

- ★ Available on the web with an open licence
- ★★ + machine-readable structured data
- ★★★ + non-proprietary format
- ★★★★ + use W3C standards (URI, RDF) to identify and serve
- ★★★★★ + link to other data
- ★★★★★★ + provide query interface (SPARQL)
- ★★★★★★★ + define vocabulary semantics (OWL)
Tim Berners-Lee’s 5 Linked Open Data stars:

- ★ Available on the web with an open licence
- ★★ + machine-readable structured data
- ★★★ + non-proprietary format
- ★★★★ + use W3C standards (URI, RDF) to identify and serve
- ★★★★★ + link to other data

Even better:

- ★★★★★★ + provide query interface (SPARQL)
5-star Linked Open Data

Tim Berners-Lee’s 5 Linked Open Data stars:

★  Available on the web with an open licence
★★ + machine-readable structured data
★★★ + non-proprietary format
★★★★ + use W3C standards (URI, RDF) to identify and serve
★★★★★ + link to other data

Even better:

★★★★★★ + provide query interface (SPARQL)
★★★★★★★ + define vocabulary semantics (OWL)
5-star Linked Open Data

Tim Berners-Lee’s 5 Linked Open Data stars:

★  Available on the web with an open licence
★★ + machine-readable structured data
★★★ + non-proprietary format
★★★★ + use W3C standards (URI, RDF) to identify and serve
★★★★★ + link to other data

Even better:

★★★★★★ + provide query interface (SPARQL)
★★★★★★★ + define vocabulary semantics (OWL)
5-star Linked Open Data

Tim Berners-Lee’s 5 Linked Open Data stars:

★  Available on the web with an open licence
★★ + machine-readable structured data (Excel)
★★★ + non-proprietary format (CSV, XML)
★★★★ + use W3C standards (URI, RDF) to identify and serve
★★★★★ + link to other data

Even better:

★★★★★★ + provide query interface (SPARQL)
★★★★★★★ + define vocabulary semantics (OWL)
5-star Linked Open Data

Tim Berners-Lee’s 5 Linked Open Data stars:

★ Available on the web with an open licence
★★ + machine-readable structured data (Excel)
★★★ + non-proprietary format (CSV, XML)
★★★★ + use W3C standards (URI, RDF) to identify and serve
★★★★★ + link to other data

Even better:

★★★★★★ + provide query interface (SPARQL)
★★★★★★★ + define vocabulary semantics (OWL)
Linked Open Data

Uniform Resource Identifier (URI)

- Schema for global identifiers
Linked Open Data
Uniform Resource Identifier (URI)

- Schema for global identifiers
  - Example: Core sample no. 3 of wellbore 1/6-A-7
    
    http://sws.ifi.uio.no/data/npd-v2/wellbore/903/core/3
    
    http://factpages.npd.no/factpages/default.aspx?culture=en&nav1=wellbore&nav2=PageView|Development|With|WithCores&nav3=903
Linked Open Data
Uniform Resource Identifier (URI)

- Schema for global identifiers
  - Example: Core sample no. 3 of wellbore 1/6-A-7
    - http://sws.ifi.uio.no/data/npd-v2/wellbore/903/core/3
  - Crucial for integration
Linked Open Data
Uniform Resource Identifier (URI)

- Schema for global identifiers
  - Example: Core sample no. 3 of wellbore 1/6-A-7
    - http://sws.ifi.uio.no/data/npd-v2/wellbore/903/core/3
  - Crucial for integration
  - Makes retrieval of information easy
Linked Open Data
Uniform Resource Identifier (URI)

- Schema for global identifiers
  - Example: Core sample no. 3 of wellbore 1/6-A-7
    
    http://sws.ifi.uio.no/data/npd-v2/wellbore/903/core/3
    
    http://factpages.npd.no/factpages/default.aspx?culture=en&nav1=wellbore&nav2=PageView|Development|With|WithCores&nav3=903

- Crucial for integration

- Makes retrieval of information easy
  
  http://sws.ifi.uio.no/data/npd-v2/page/wellbore/903/core/3
  http://sws.ifi.uio.no/data/npd-v2/data/wellbore/903/core/3
  http://sws.ifi.uio.no/data/npd-v2/wellbore/903/core/3
Linked Open Data
Uniform Resource Identifier (URI)

- Schema for global identifiers
  - Example: Core sample no. 3 of wellbore 1/6-A-7
    http://sws.ifi.uio.no/data/npm-v2/wellbore/903/core/3
    http://factspages.npd.no/factpages/default.aspx?culture=en&nav1=wellbore&nav2=PageView|Development|With|WithCores&nav3=903
  - Crucial for integration
  - Makes retrieval of information easy
    http://sws.ifi.uio.no/data/npm-v2/page/wellbore/903/core/3
    http://sws.ifi.uio.no/data/npm-v2/data/wellbore/903/core/3
    http://sws.ifi.uio.no/data/npm-v2/wellbore/903/core/3
  - Using existing "web machinery"
Linked Open Data
Resource Description Framework (RDF)

• Simple, generic data model
  • It’s all triples!

<table>
<thead>
<tr>
<th>subject</th>
<th>predicate</th>
<th>object</th>
</tr>
</thead>
<tbody>
<tr>
<td>npd:wellbore/903</td>
<td>rdf:type</td>
<td>npdv:Wellbore</td>
</tr>
<tr>
<td>npd:wellbore/903</td>
<td>npdv:hasCoreSample</td>
<td>npd:wellbore/903/core/3</td>
</tr>
</tbody>
</table>
Linked Open Data
Resource Description Framework (RDF)

- **Simple, generic data model**
  - It’s all triples!

<table>
<thead>
<tr>
<th>subject</th>
<th>predicate</th>
<th>object</th>
</tr>
</thead>
<tbody>
<tr>
<td>npd:wellbore/903</td>
<td>rdf:type</td>
<td>npdv:Wellbore</td>
</tr>
<tr>
<td>npd:wellbore/903</td>
<td>npdv:hasCoreSample</td>
<td>npd:wellbore/903/core/3</td>
</tr>
</tbody>
</table>

- **No/One schema/schema independent, so (at a technical level):**
  - Easy to extend
  - Easy to exchange
  - Easy to consume
  - Easy to query
Linked Open Data
Resource Description Framework (RDF)

- Simple, generic data model
  - It’s all triples!
    
    | subject                  | predicate          | object                     |
    |--------------------------|--------------------|----------------------------|
    | npd:wellbore/903         | rdf:type           | npdv:Wellbore              |
    | npd:wellbore/903         | npdv:hasCoreSample | npd:wellbore/903/core/3    |

- No/One schema/schema independent, so (at a technical level):
  - Easy to extend
  - Easy to exchange
  - Easy to consume
  - Easy to query

- Use with “any” RDF tool
  - tools to browse, visualise, edit
Linked Open Data
SPARQL Protocol and RDF Query Language (SPARQL)

- RDF query language, similar to SQL

Not possible to answer with existing NPD FactPages
Linked Open Data
SPARQL Protocol and RDF Query Language (SPARQL)

- RDF query language, similar to SQL
- Protocol for
  - sending queries
  - receiving results
  - result format(s)
  - using existing web infrastructure

Not possible to answer with existing NPD FactPages
Linked Open Data
SPARQL Protocol and RDF Query Language (SPARQL)

- RDF query language, similar to SQL
- Protocol for
  - sending queries
  - receiving results
  - result format(s)
  - using existing web infrastructure
- Safely expose data to querying over Internet

Not possible to answer with existing NPD FactPages
Linked Open Data
SPARQL Protocol and RDF Query Language (SPARQL)

- RDF query language, similar to SQL
- Protocol for
  - sending queries
  - receiving results
  - result format(s)
  - using existing web infrastructure
- Safely expose data to querying over Internet
- Examples
  - Total production of oil and gas jan–jun 2010 for Statoil operated fields
  - Find things close to Svalbard
  - Ekofisk oil and gas production

Not possible to answer with existing NPD FactPages
Linked Open Data
Web Ontology Language (OWL)

- Language for formally defining vocabulary semantics
  - what the data *means*, not how to store it
• Language for formally defining vocabulary semantics
  • what the data *means*, not how to store it

• Domain-centric vocabulary
  • Data is accessible to the “right people”
  • Easier to formulate queries
Linked Open Data
Web Ontology Language (OWL)

- Language for formally defining vocabulary semantics
  - what the data *means*, not how to store it
- Domain-centric vocabulary
  - Data is accessible to the “right people”
  - Easier to formulate queries
- Sophisticated integration using relations between ontologies
Linked Open Data
Web Ontology Language (OWL)

- Language for formally defining vocabulary semantics
  - what the data *means*, not how to store it
- Domain-centric vocabulary
  - Data is accessible to the “right people”
  - Easier to formulate queries
- Sophisticated integration using relations between ontologies
- Formal semantics allows reasoning
  - infer new facts
  - check for inconsistencies
  - explain consequences and query answers
Outline

NPD and NPD FactPages

Linked Open Data

Converting NPD’s FactPages to Semantic Web Data

Examples
  Map Visualization
  Query Formulation
Conversion Process

1a. download

1b. gen.

Transformation specification

1c. create

1d. load

DB schema

D2RQ map

SQL

2a. generate

2b. dump

Rel. DB

2c. load

RDF

Triple store

3a. generate

3b. load

OWL

SPARQL

Triple store
Conversion Process

1. Create relational database representation
1. Create relational database representation
   a) Download CSV files from NPD FactPages website
1. Create relational database representation

b) Generate database schema—based on CSV files and transformation spec.
1. Create relational database representation
   
   c) Create and

   d) load the database
2. Create RDF/LOD representation
2. Create RDF/LOD representation

a) Generate D2RQ map—based on transformation spec. and data
2. Create RDF/LOD representation
   
b) Dump database to RDF using map
2. Create RDF/LOD representation
   c) and load into triple store
3. Create ontology representation
3. Create ontology representation

a) Generate OWL ontology—based on transformation spec. and data—and extend manually
Conversion Process

3. Create ontology representation
   
b) Load RDF + OWL into triple store
Conversion Process

We test the different representations in our query evaluation experiment.
Results

- Availability of data is better
  - Each thing has an identifier
  - Find information using identifier
- Easier to exchange and integrate data
- More data is available
  - Arbitrary SPARQL queries
  - New facts added by reasoning with semantics
- Semantics are more explicit
- All using standardised languages
  - application independent
  - general-purpose tools available
Outline

NPD and NPD FactPages

Linked Open Data

Converting NPD’s FactPages to Semantic Web Data

Examples
   Map Visualization
   Query Formulation
Example: Map Visualization

NPD FactMap vs. “our” SPARQL map visualiser.

- Official map app.

- Server side
  - Predefined queries only
  - Map results only

- Generic map app
  - Client side, SPARQL based, open source
  - Predefined generic ontology driven queries and any SPARQL query
  - Results link to LOD
Example: Map Visualization

NPD FactMap vs. “our” SPARQL map visualiser.

- Official map app.
- Server side
NPD FactMap vs. “our” SPARQL map visualiser.

- Official map app.
- Server side
- Predefined queries only
Example: Map Visualization

NPD FactMap vs. “our” SPARQL map visualiser.

- Official map app.
- Server side
- Predefined queries only
- Map results only
Example: Map Visualization

NPD FactMap vs. “our” SPARQL map visualiser.

- Official map app.
- Server side
- Predefined queries only
- Map results only

- Generic map app
Example: Map Visualization

NPD FactMap vs. “our” SPARQL map visualiser.

- Official map app.
- Server side
- Predefined queries only
- Map results only

- Generic map app
- Client side, SPARQL based, open source
Example: Map Visualization

NPD FactMap vs. “our” SPARQL map visualiser.

- Official map app.
- Server side
- Predefined queries only
- Map results only

- Generic map app
- Client side, SPARQL based, open source
- Predefined generic ontology driven queries and any SPARQL query
Example: Map Visualization

NPD FactMap vs. “our” SPARQL map visualiser.

- Official map app.
- Server side
- Predefined queries only
- Map results only

- Generic map app
- Client side, SPARQL based, open source
- Predefined generic ontology driven queries and any SPARQL query
- Results link to LOD
Example: Visual Query Formulation

- We want data to be available to domain experts
- Most domain experts don’t know query languages
- Developed query formulation tool
- Generic, ontology/data driven application
- Customized to dataset
Example: Visual Query Formulation

Query: *Fields operated by Statoil Petroleum AS with gas producing facilities*

- [http://sws.ifi.uio.no/project/optique/qf/interface/](http://sws.ifi.uio.no/project/optique/qf/interface/)
- Field
- Company
  - Name: Statoil Petroleum ASA
- Facility
  - facility function: GAS PRODUCER
- Show SPARQL
- Run query
- Select SNØHVIT
For more information:
http://sws.ifi.uio.no/project/npd-v2/

Available for download:
- CSV
- SQL
- Mappings
- RDF dumps
- OWL ontologies
- SPARQL endpoint
- LOD front-end
- Queries