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

Outline

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

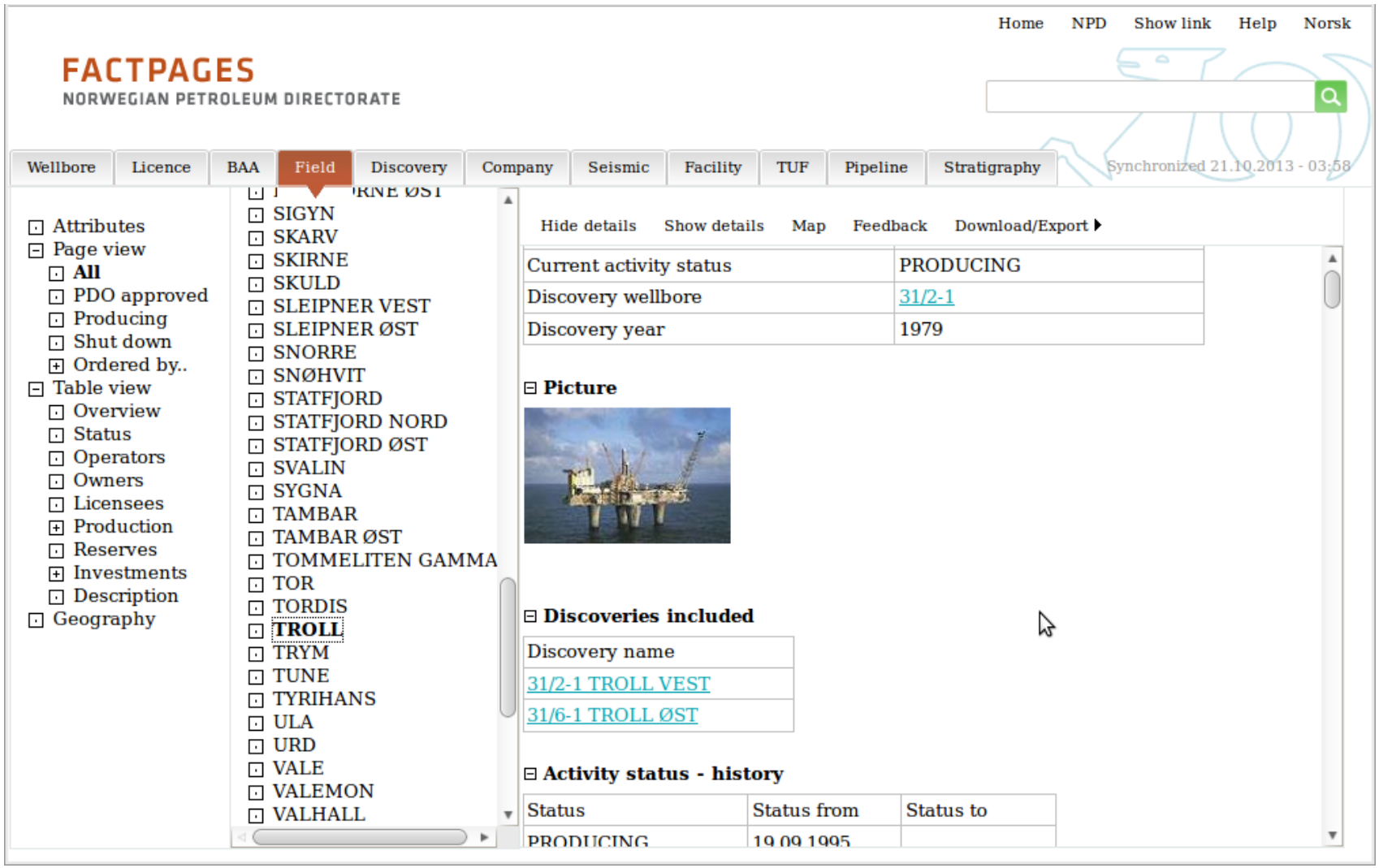
NPD FactPages

What is it?

- 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

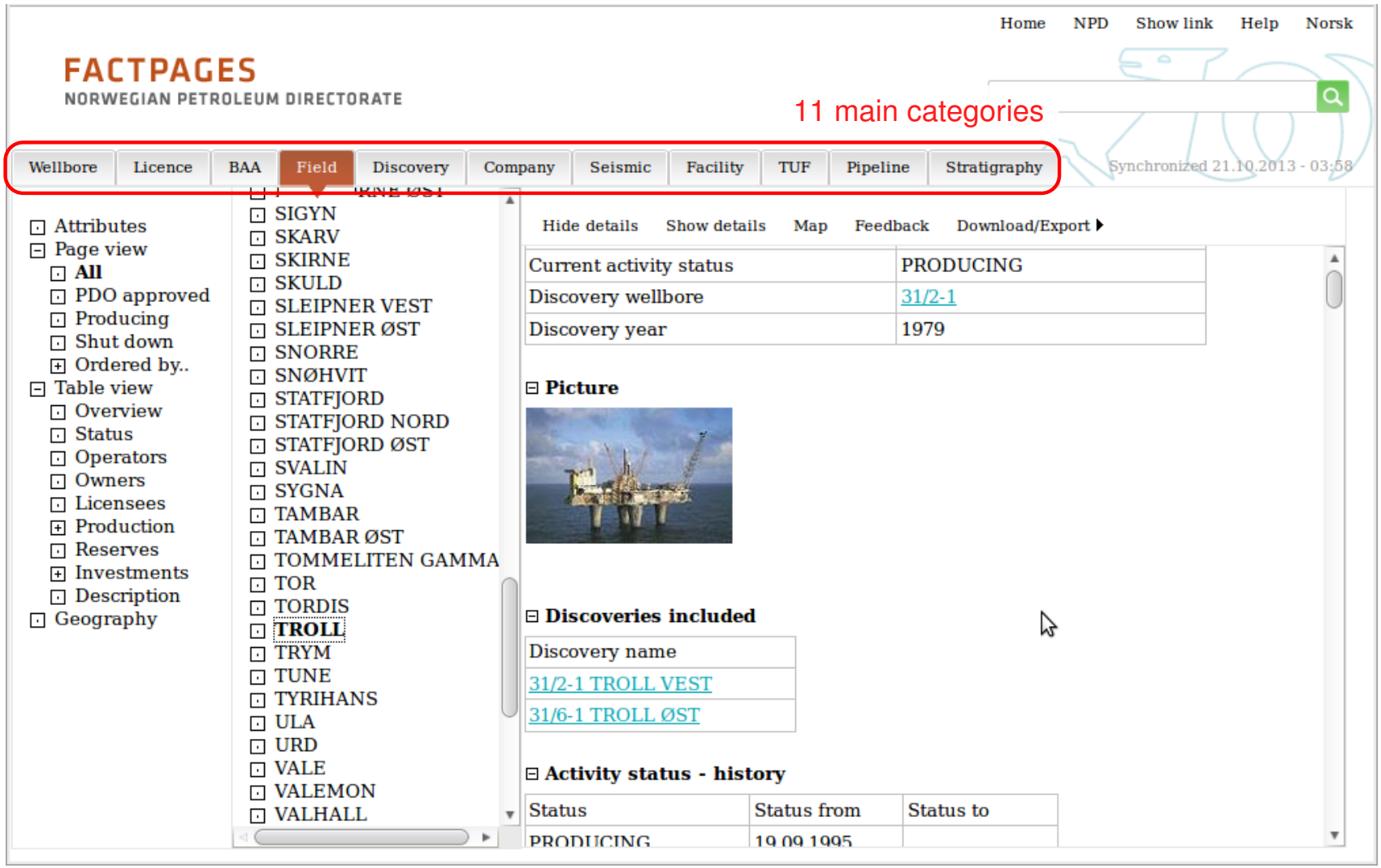
NPD FactPages

Demonstration



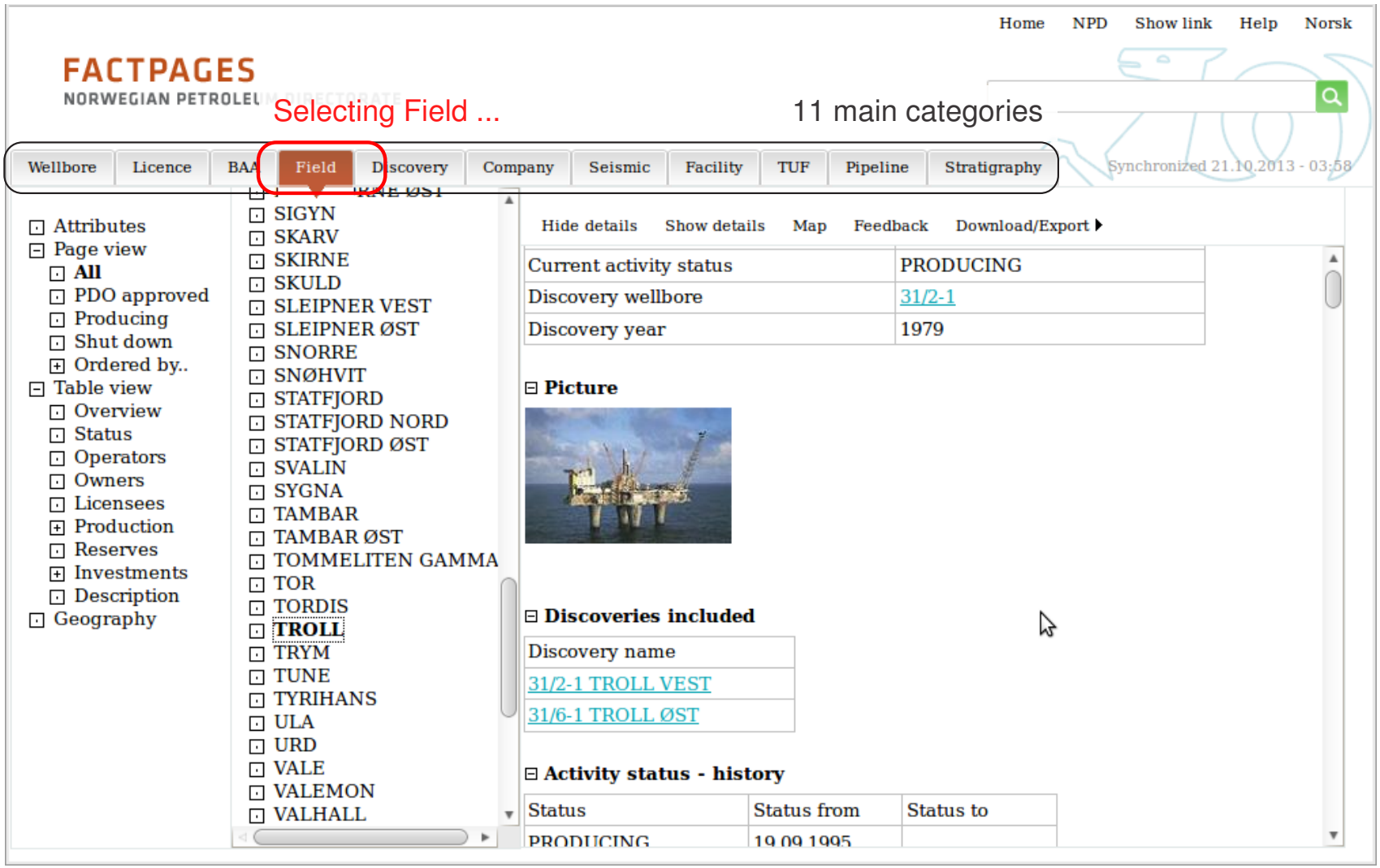
NPD FactPages

Demonstration



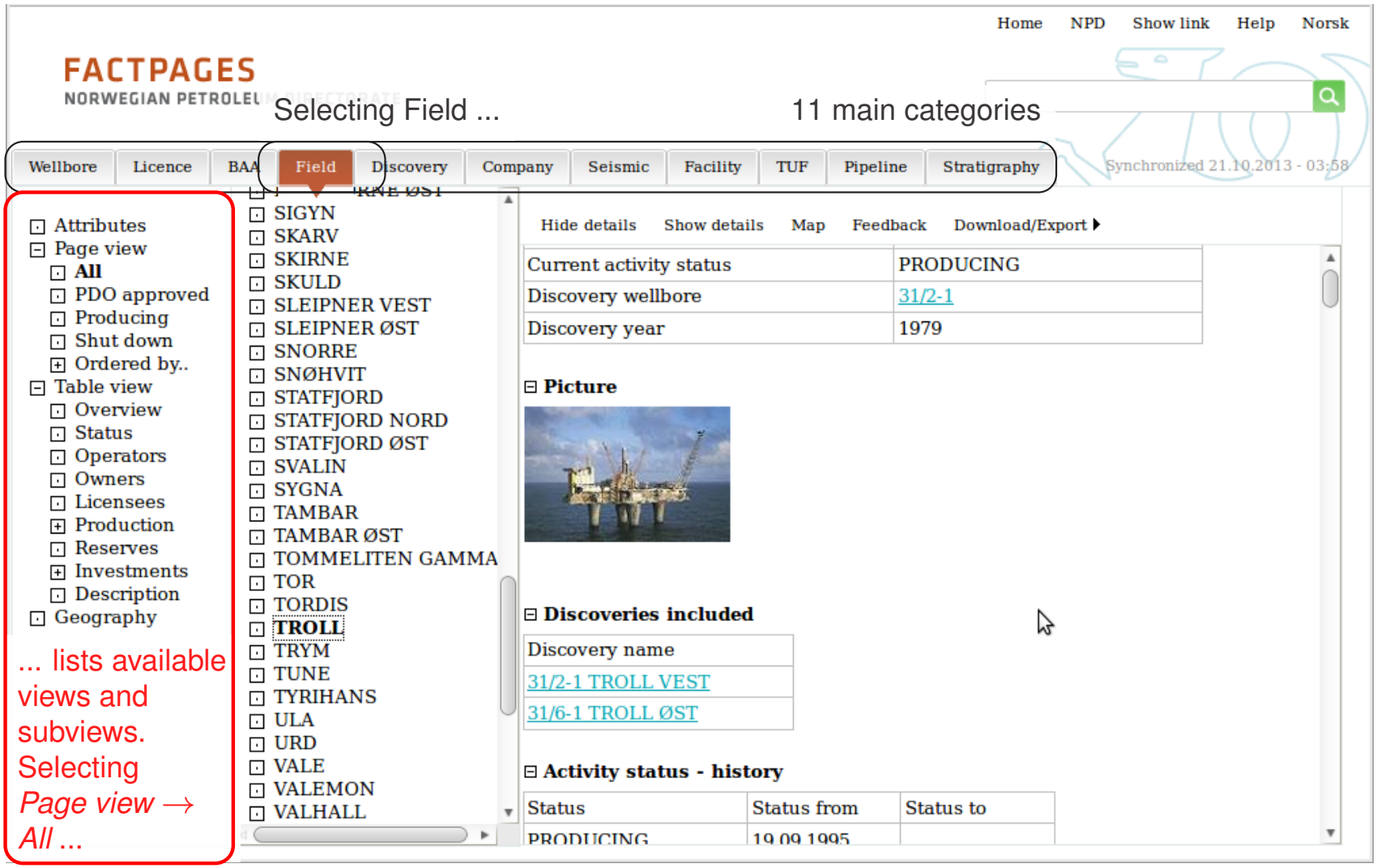
NPD FactPages

Demonstration



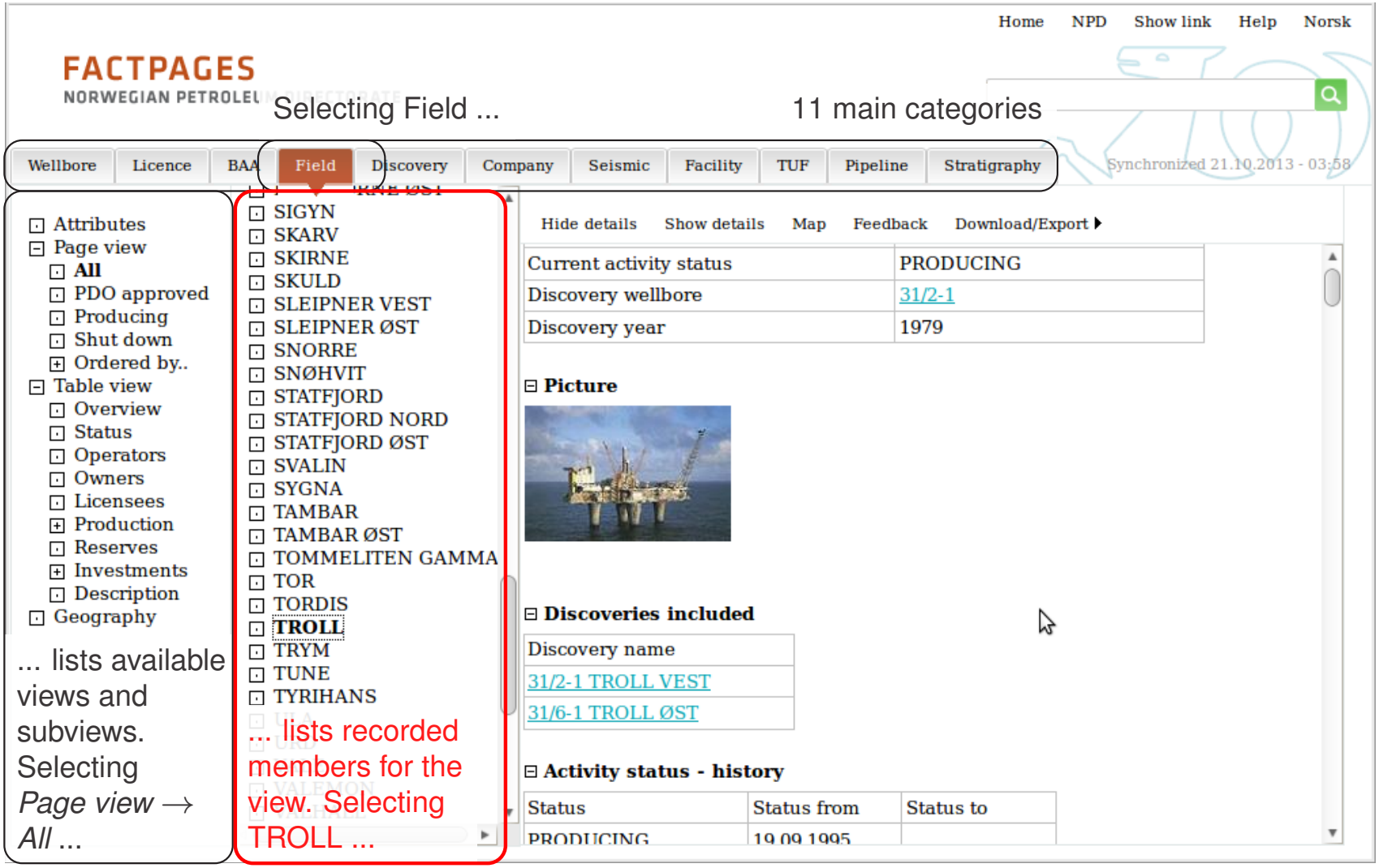
NPD FactPages

Demonstration



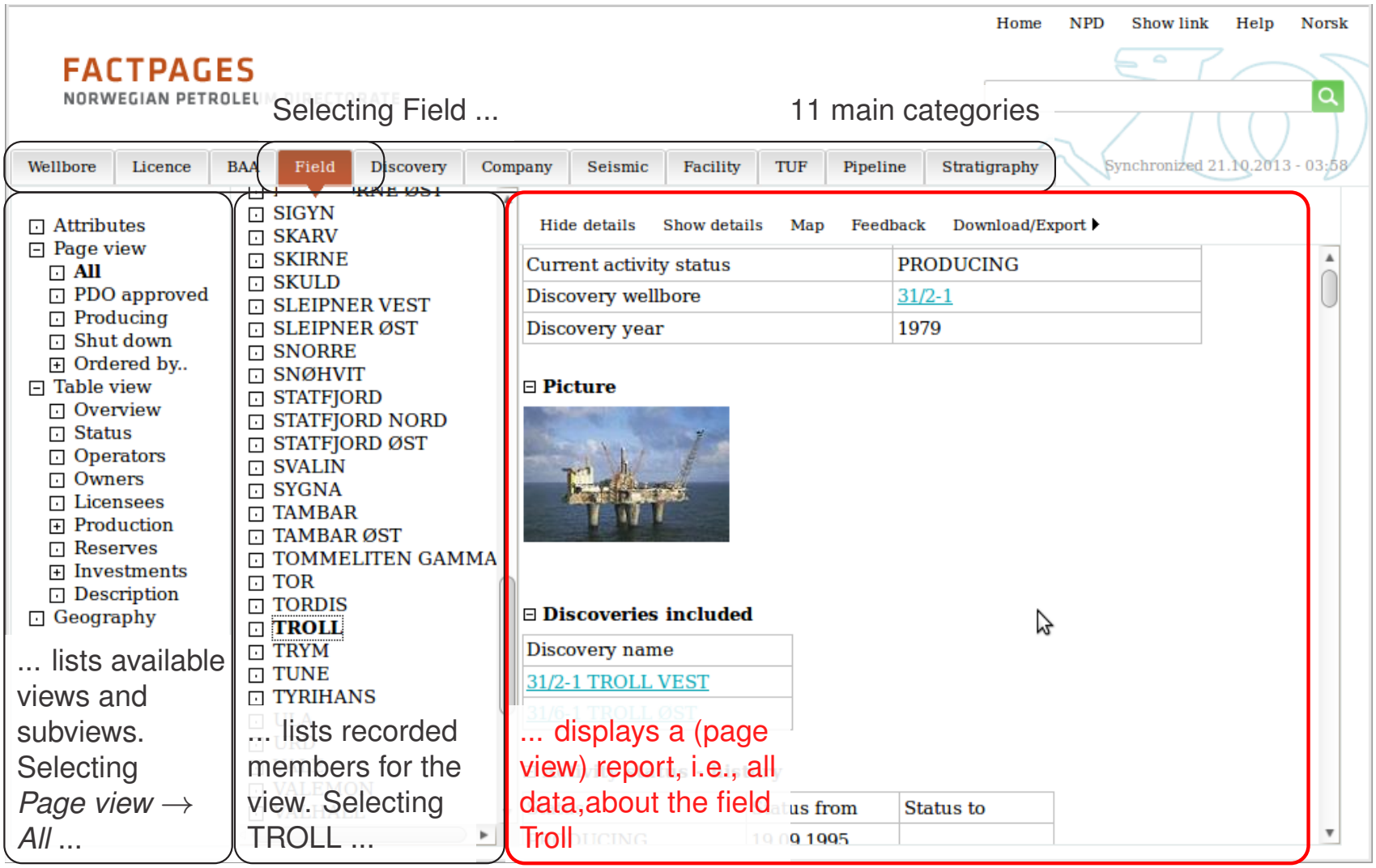
NPD FactPages

Demonstration



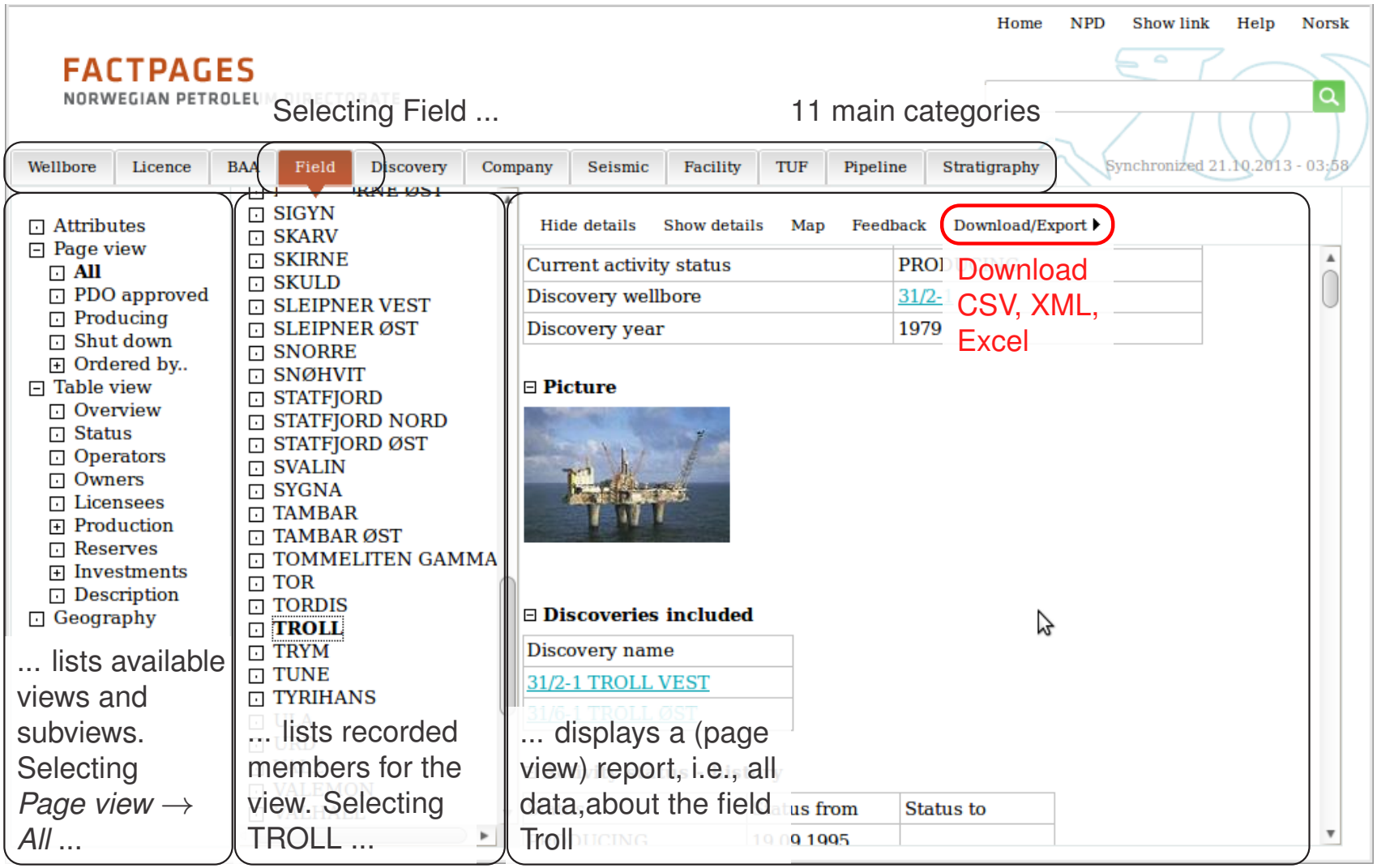
NPD FactPages

Demonstration



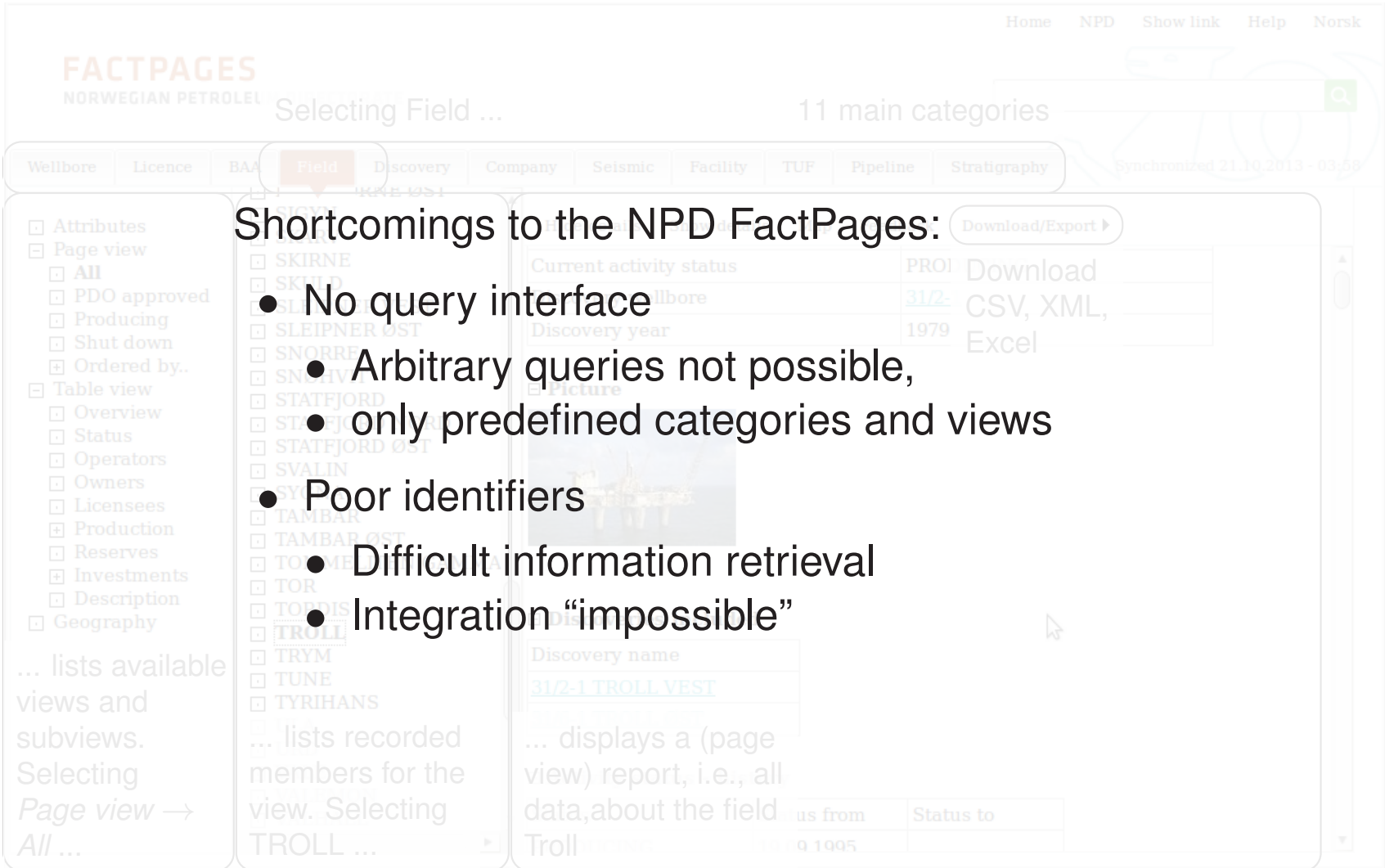
NPD FactPages

Demonstration



NPD FactPages

Demonstration



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

5-star Linked Open Data

Tim Berners-Lee's 5 Linked Open Data stars

5-star Linked Open Data

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

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

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

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

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

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)

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

- 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`

- 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

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 retrival 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 retrival 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/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 retrival 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`
- Using existing "web machinery"

- 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`

- 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

- 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

- RDF query language, similar to SQL

Not possible to answer with existing NPD FactPages

- 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

- 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

- 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

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

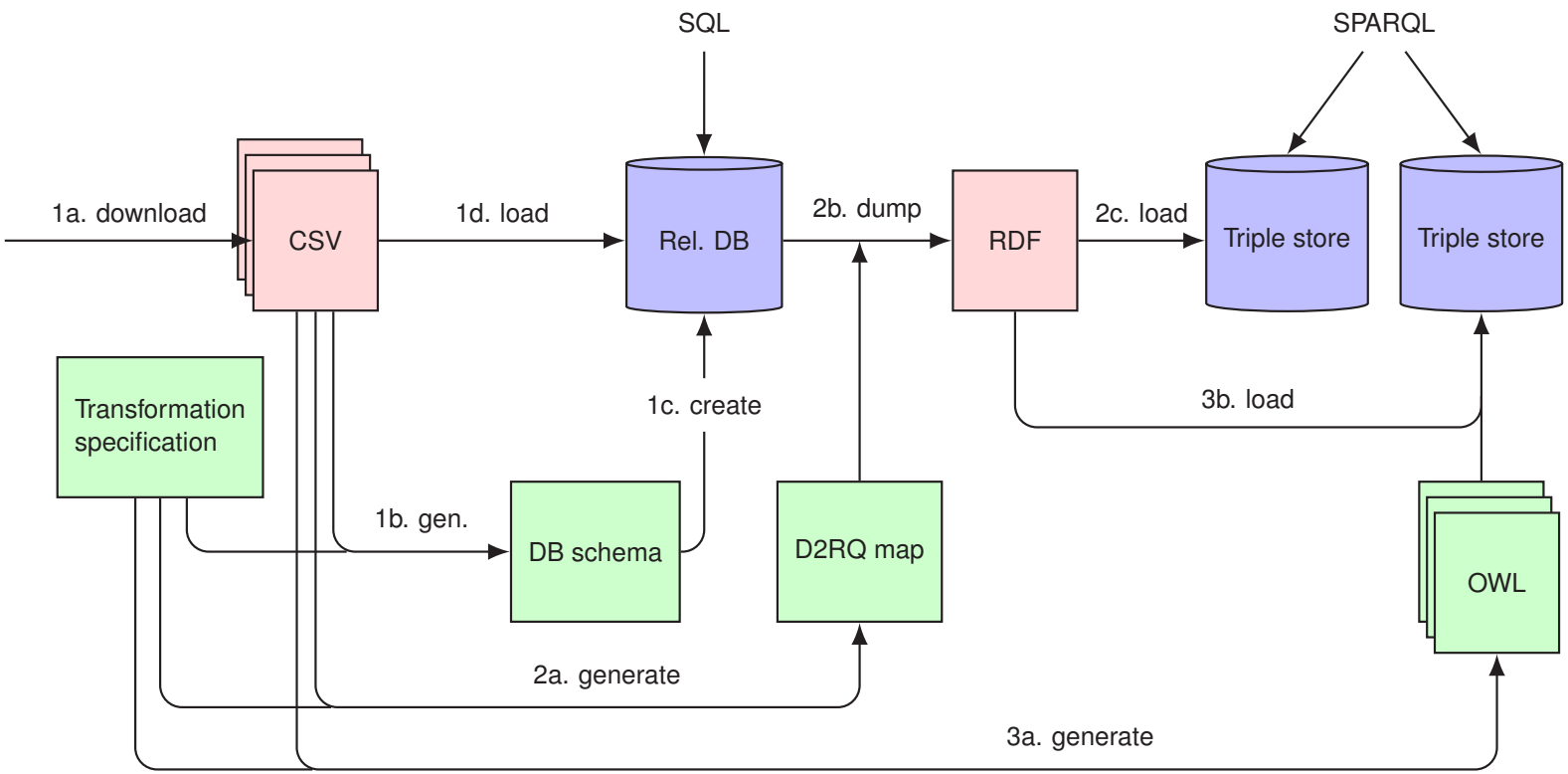
- Language for formally defining vocabulary semantics
  - what the data *means*, not how to store it
- Domain-centric vocabulary
  - Data is accessible to the “right people”
  - Easier to formulate queries

- 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 intregration using relations between ontologies

- 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 intregration using relations between ontologies
- Formal semantics allows reasoning
  - infer new facts
  - check for inconsistencies
  - explain consequences and query answers

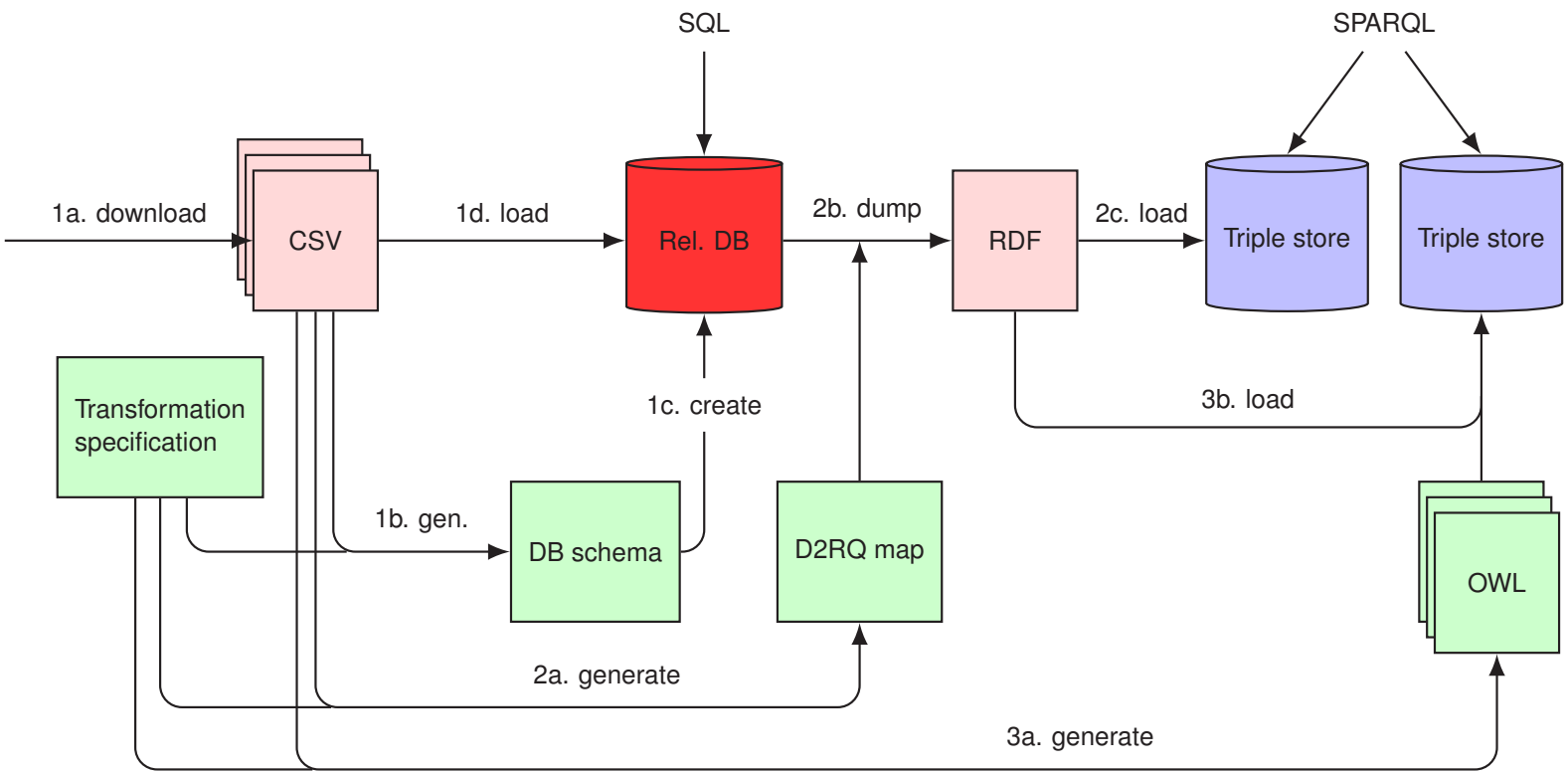
Outline

Conversion Process



Conversion Process

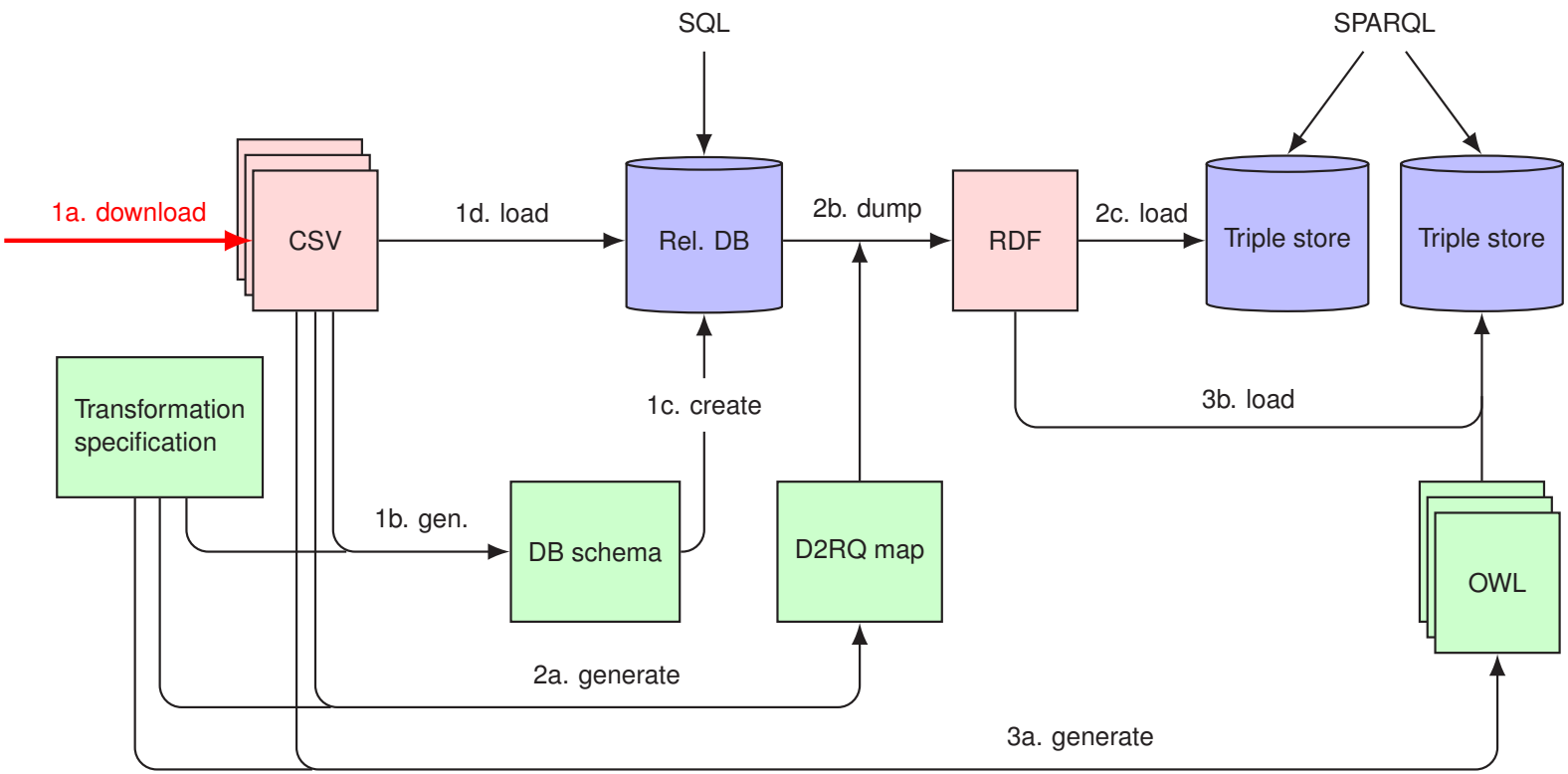
1. Create relational database representation



Conversion Process

1. Create relational database representation

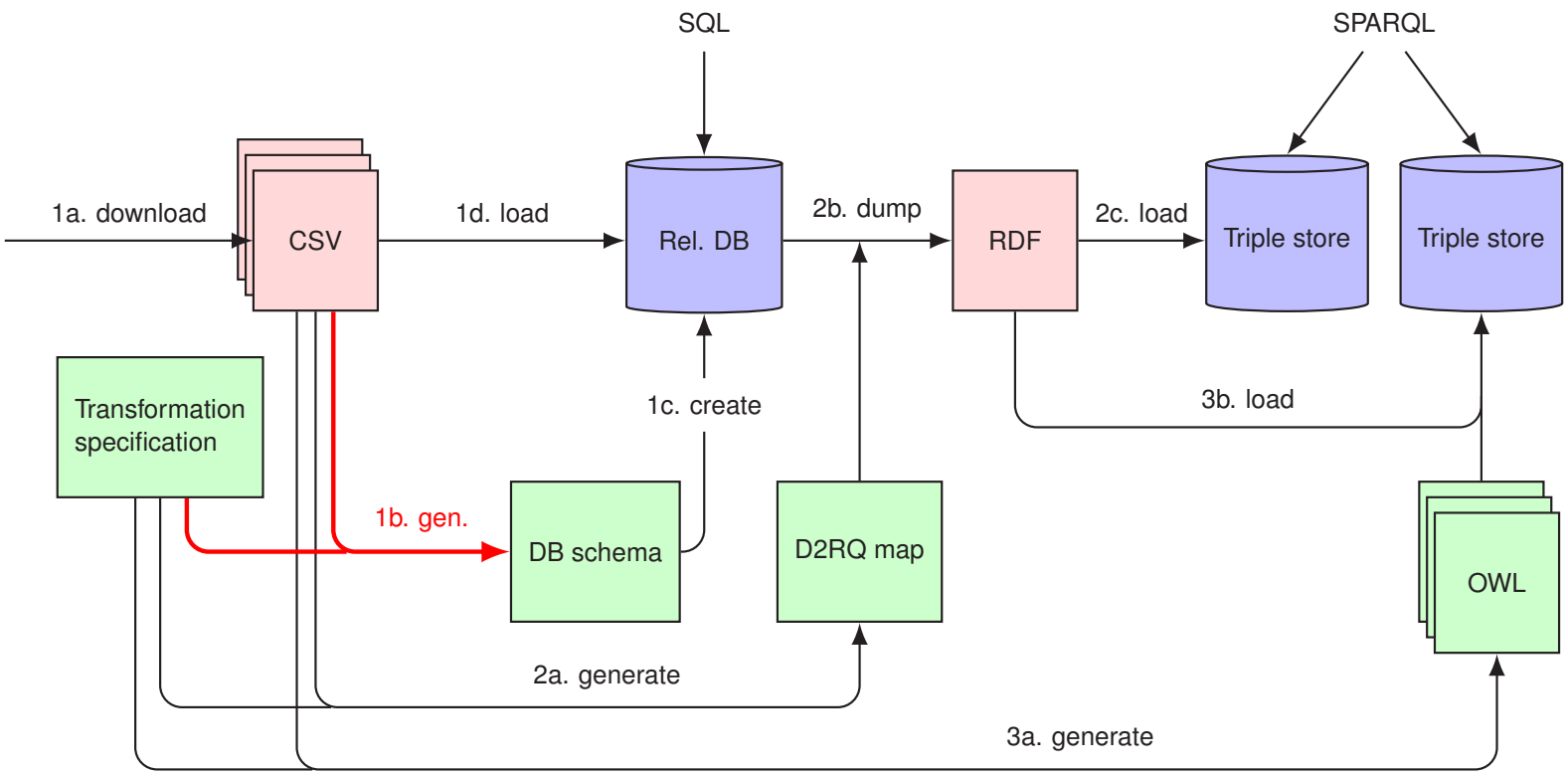
   a) Download CSV files from NPD FactPages website



Conversion Process

1. Create relational database representation

   b) Generate database schema—based on CSV files and transformation spec.

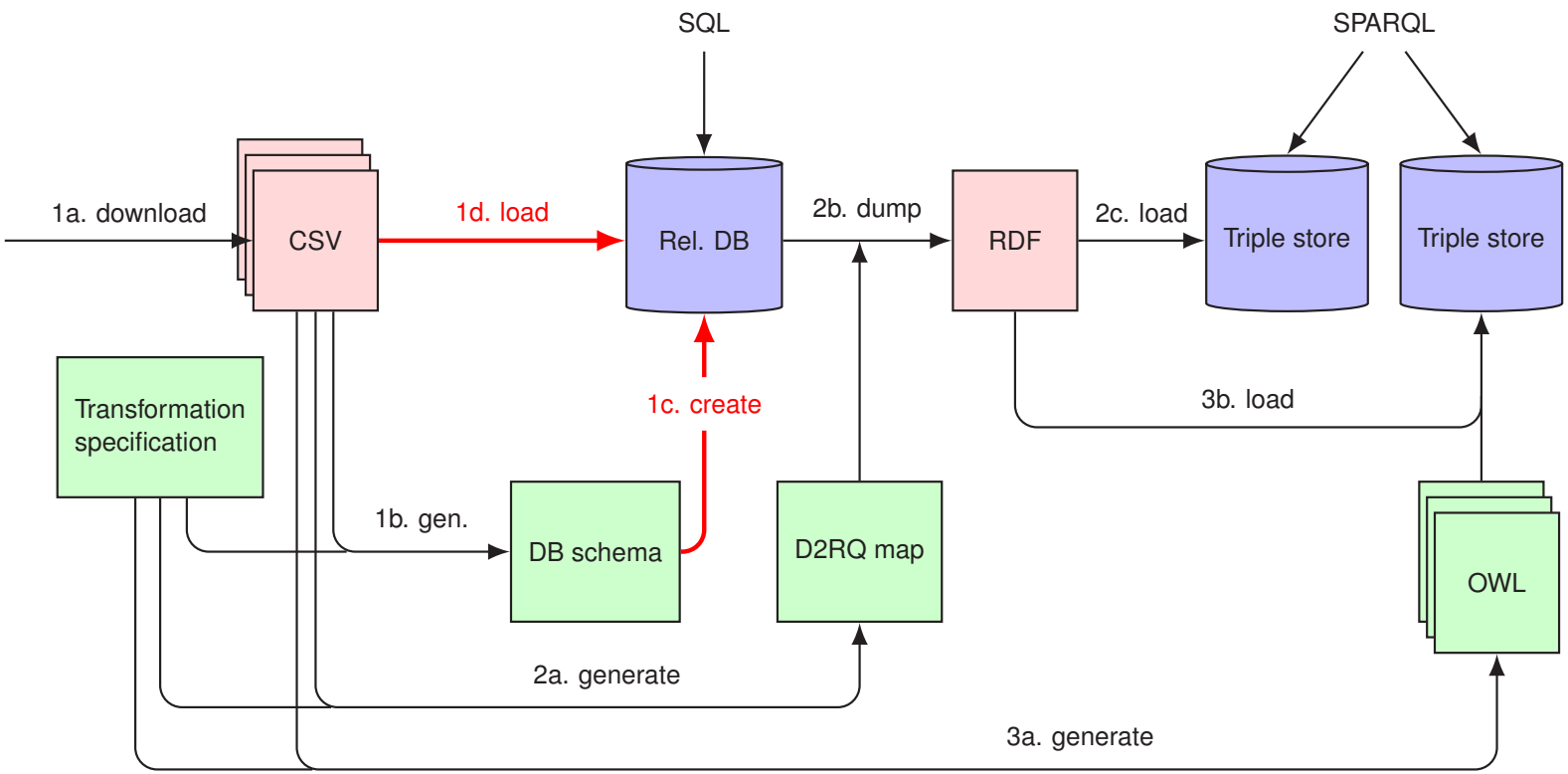


Conversion Process

1. Create relational database representation

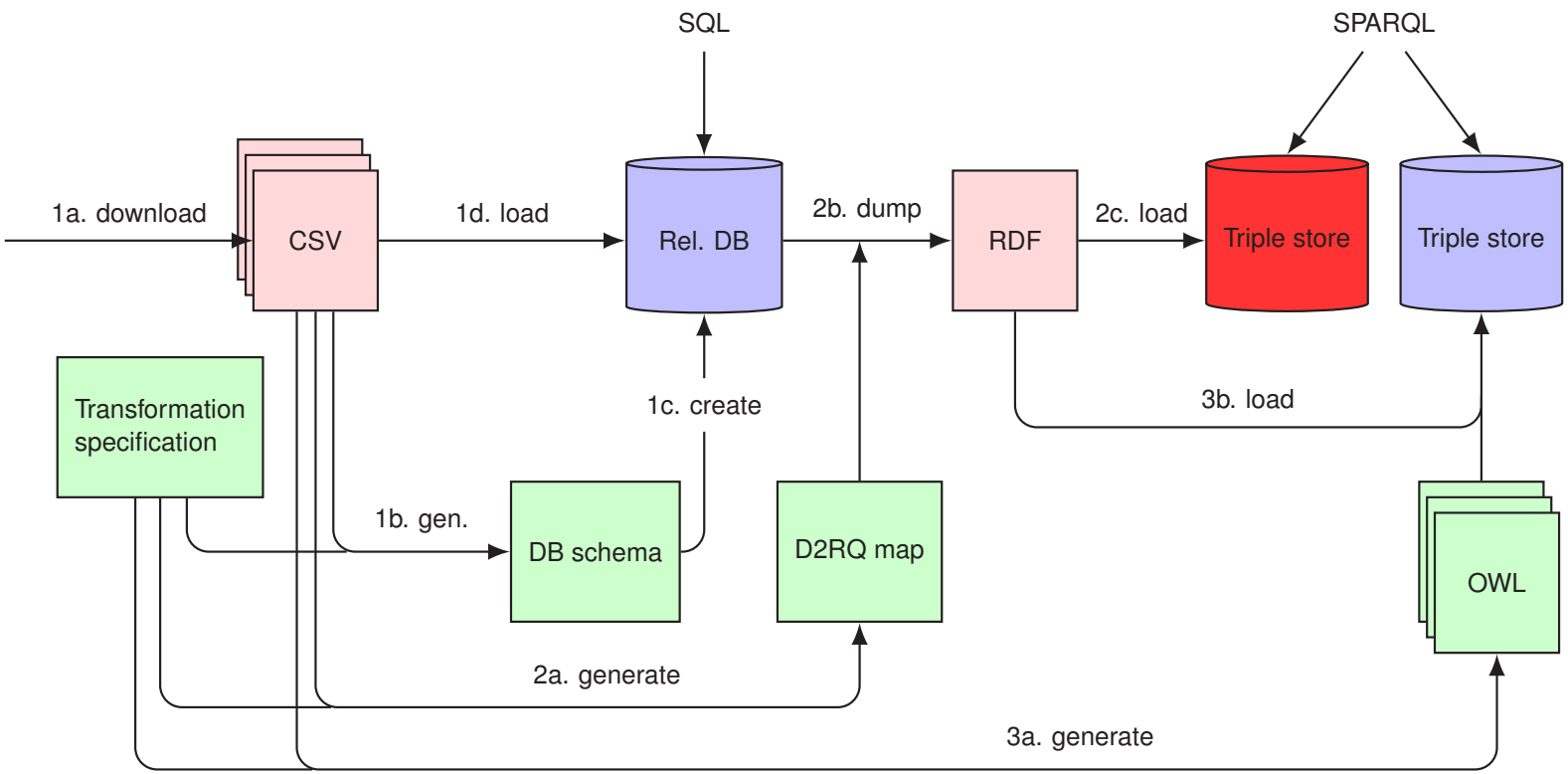
   c) Create and

   d) load the database



Conversion Process

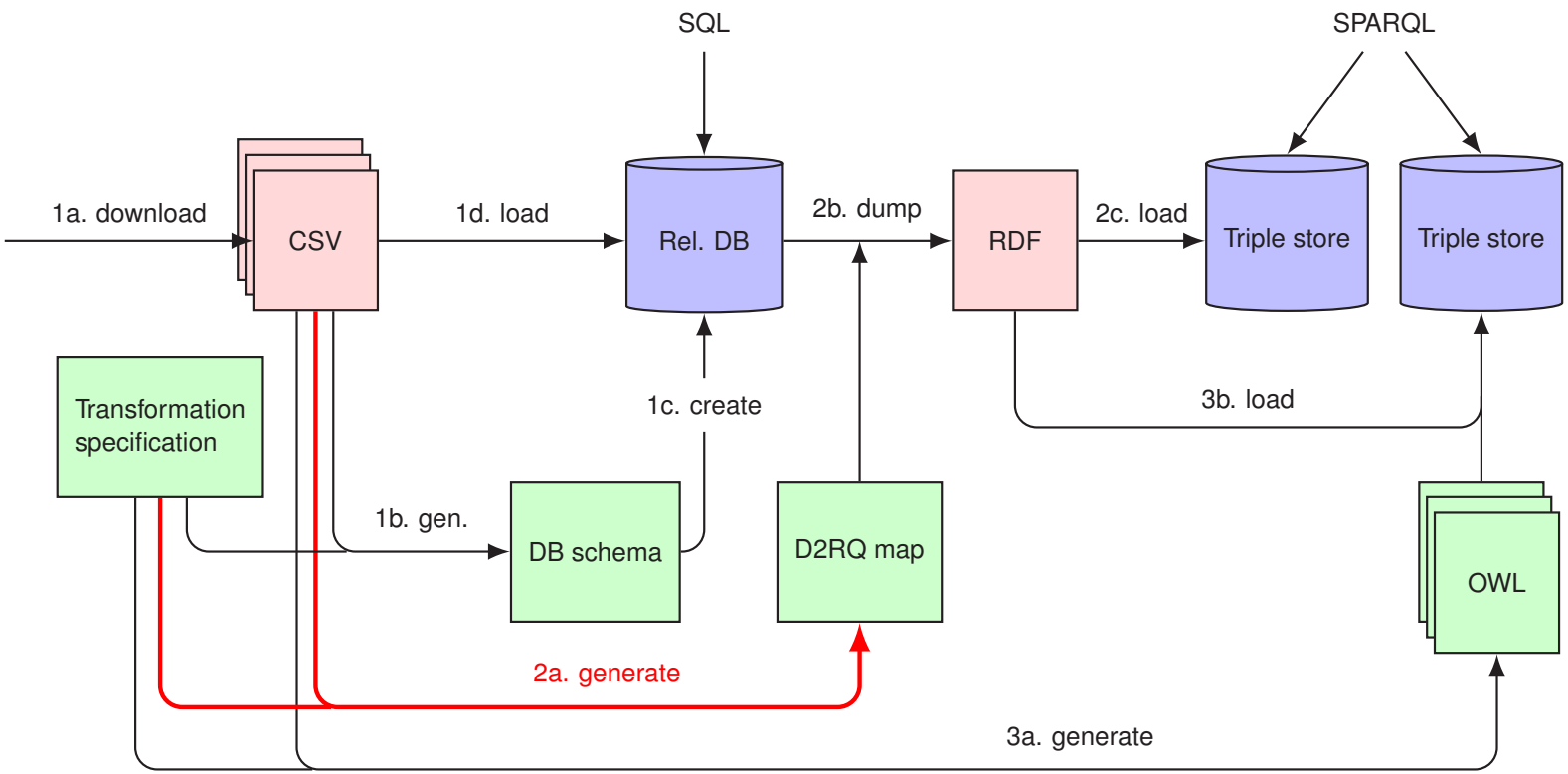
2. Create RDF/LOD representation



Conversion Process

2. Create RDF/LOD representation

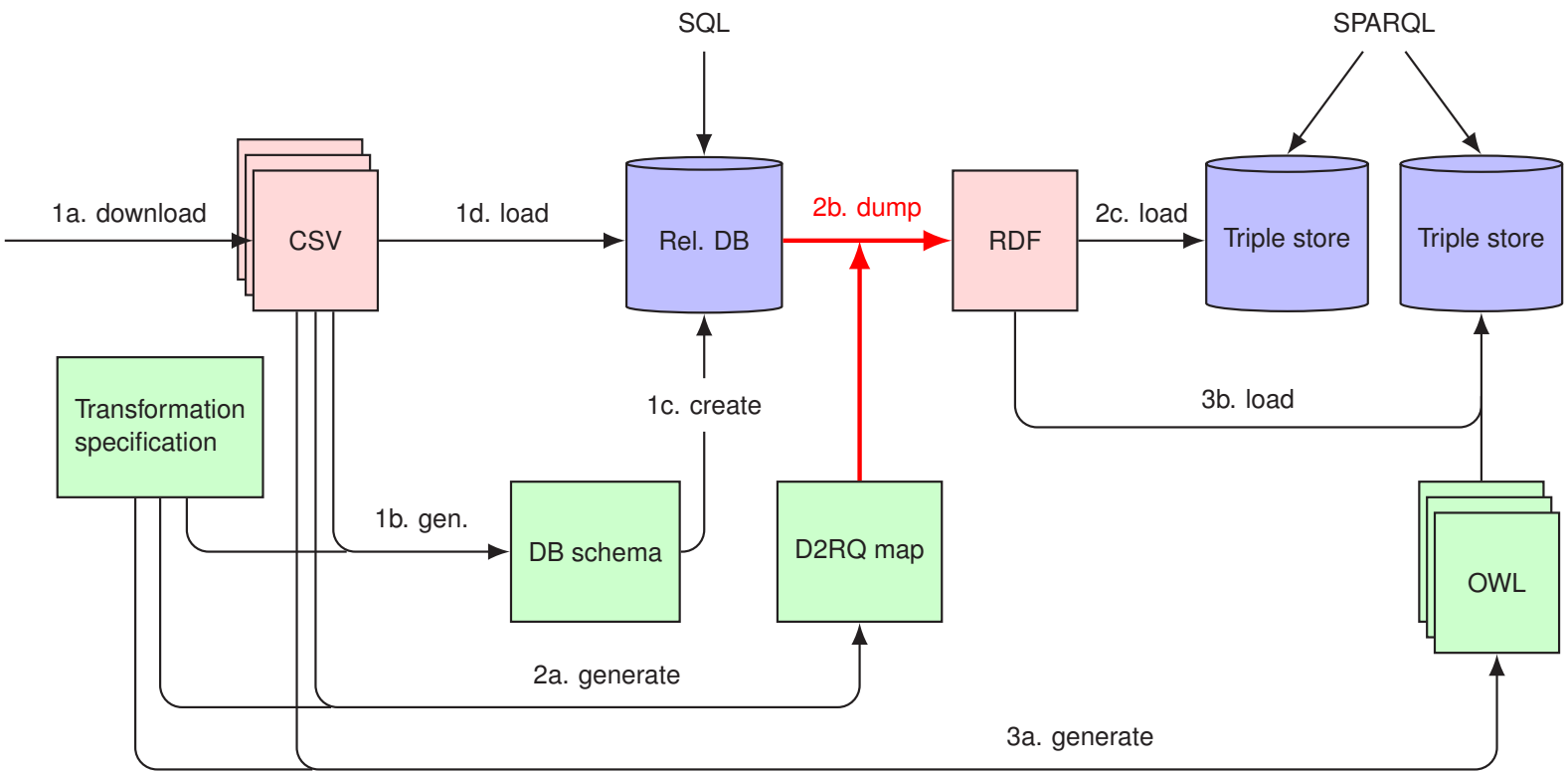
   a) Generate D2RQ map—based on transformation spec. and data



Conversion Process

2. Create RDF/LOD representation

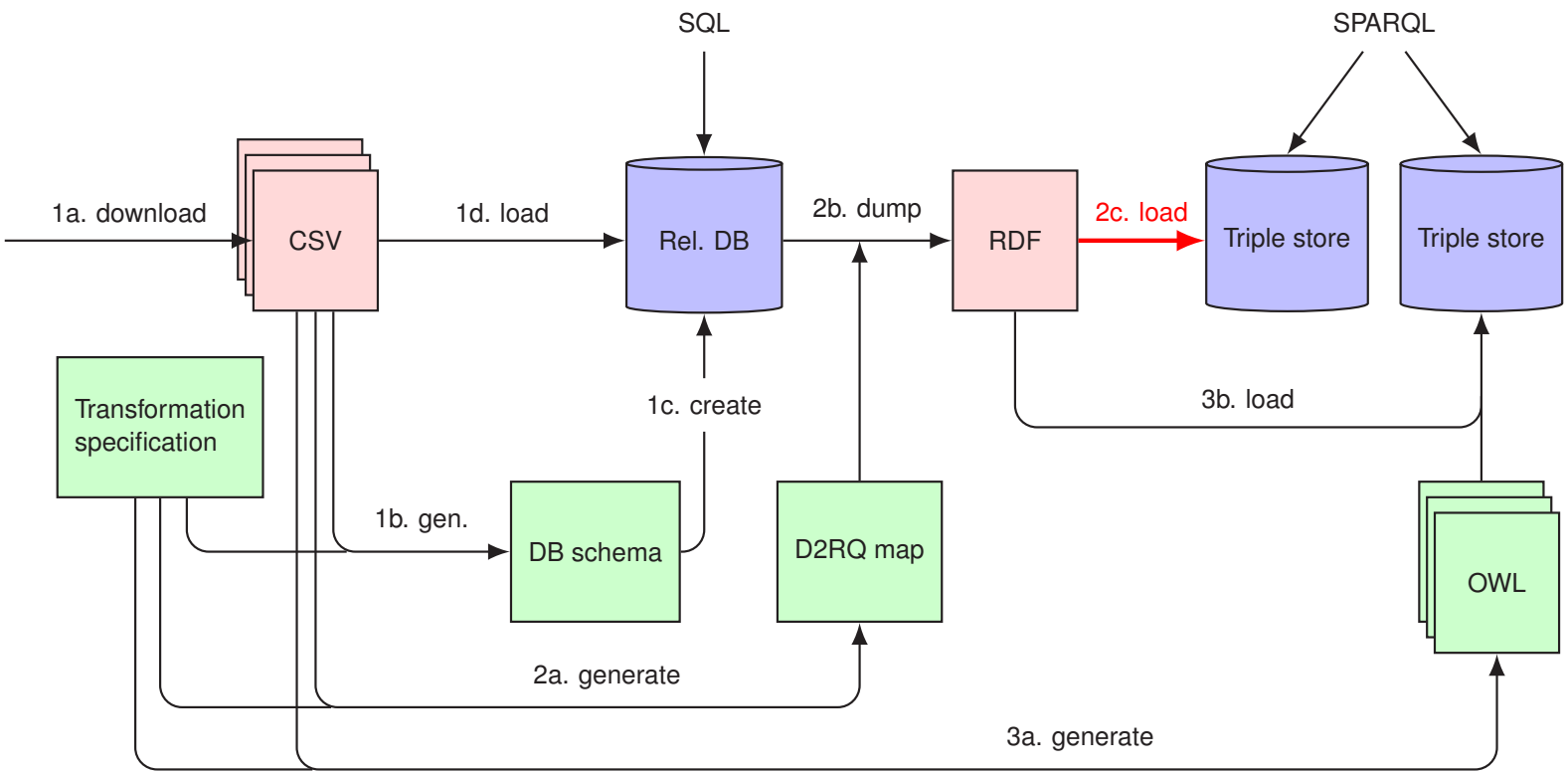
   b) Dump database to RDF using map



Conversion Process

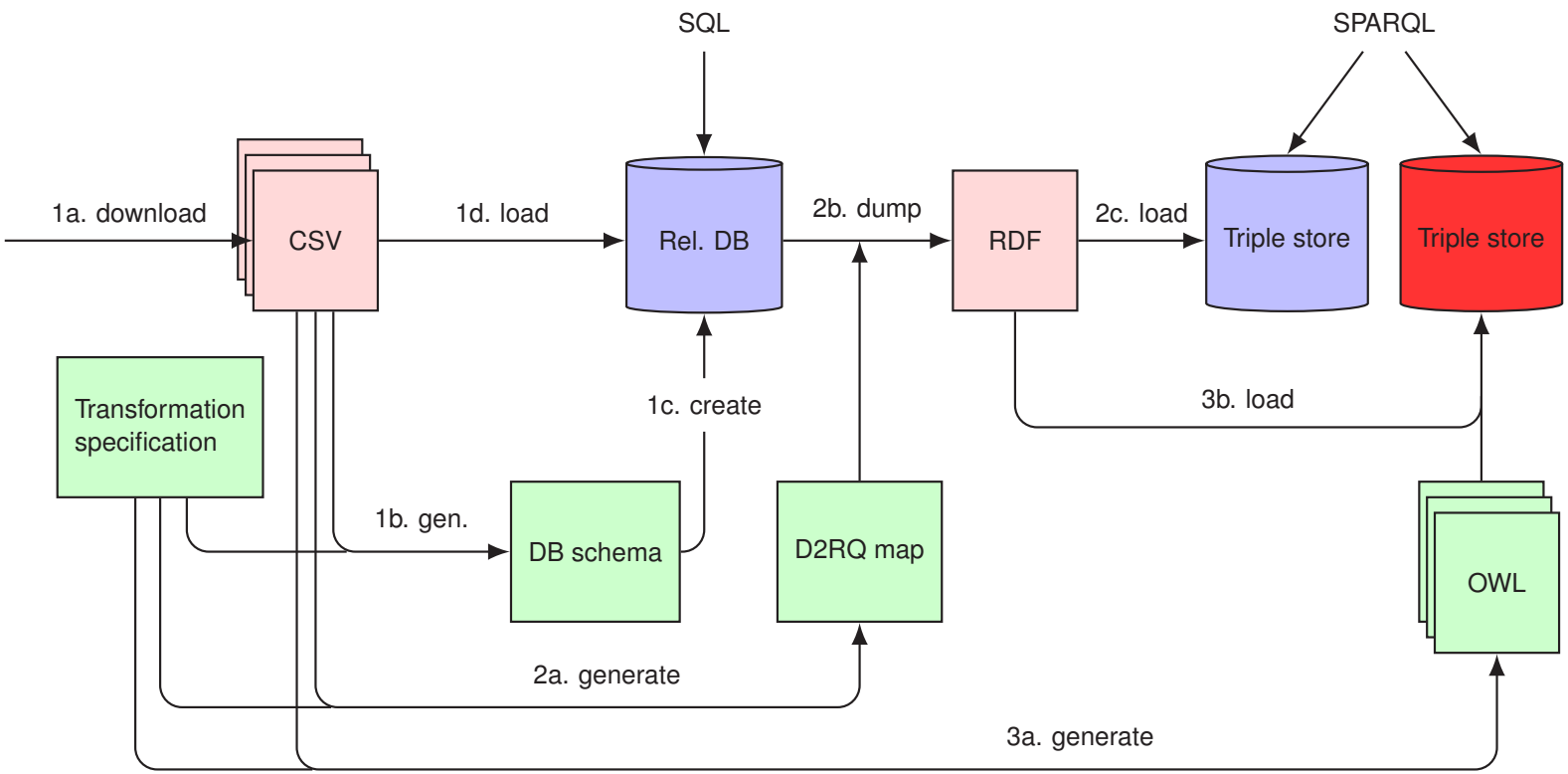
2. Create RDF/LOD representation

   c) and load into triple store



Conversion Process

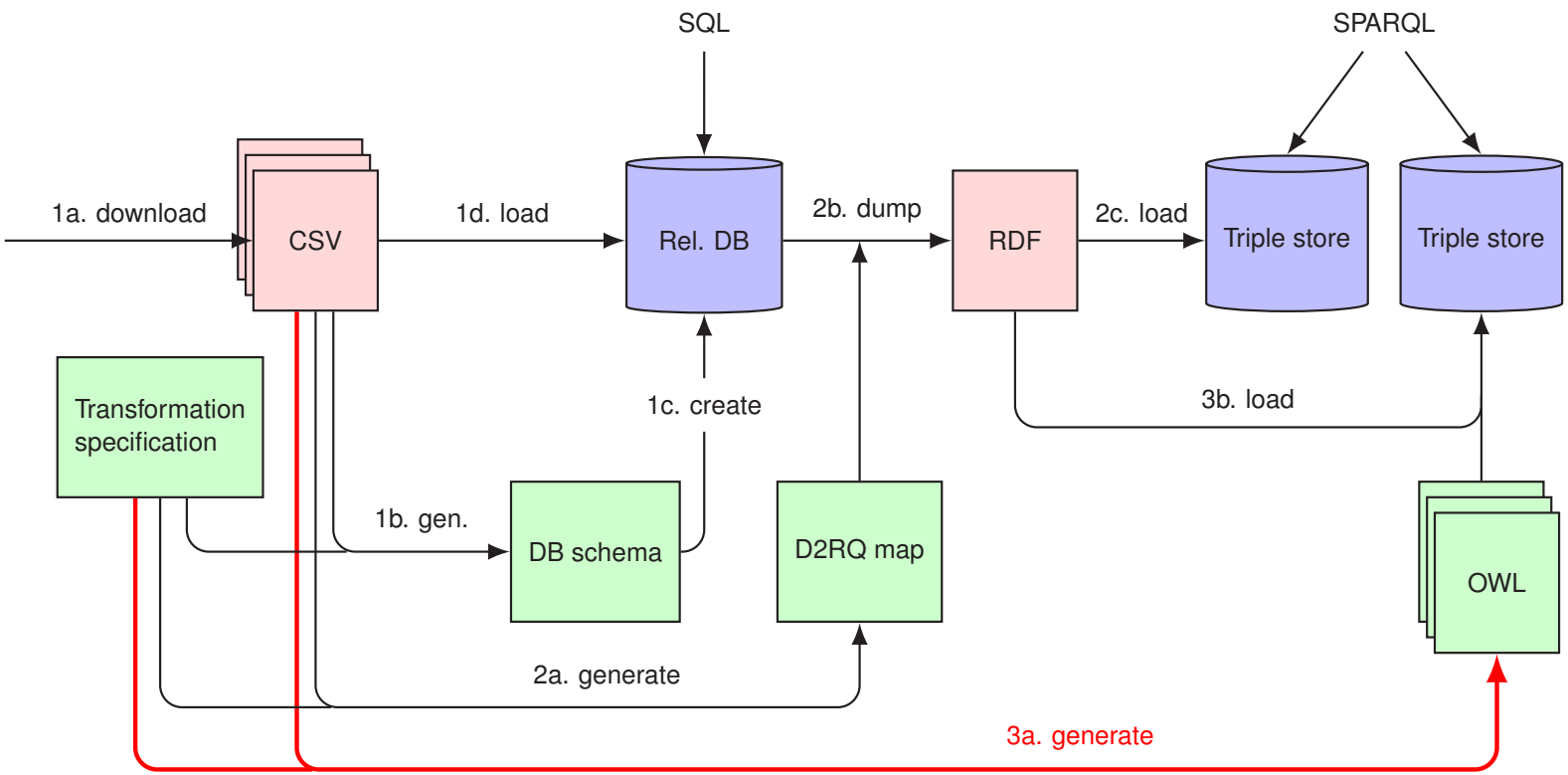
3. Create ontology representation



Conversion Process

3. Create ontology representation

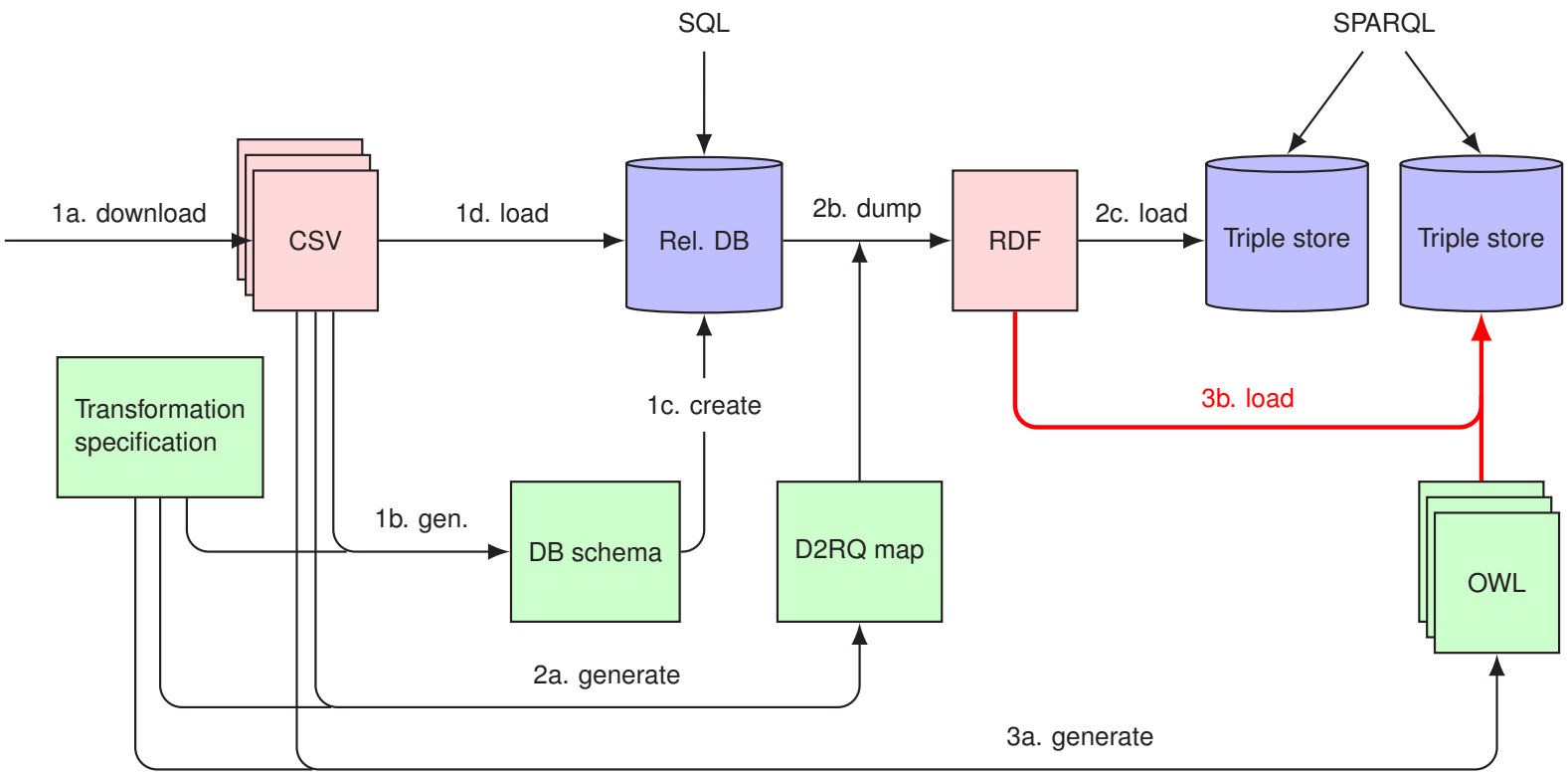
   a) Generate OWL ontology—based on transformation spec. and data—and extend manually



Conversion Process

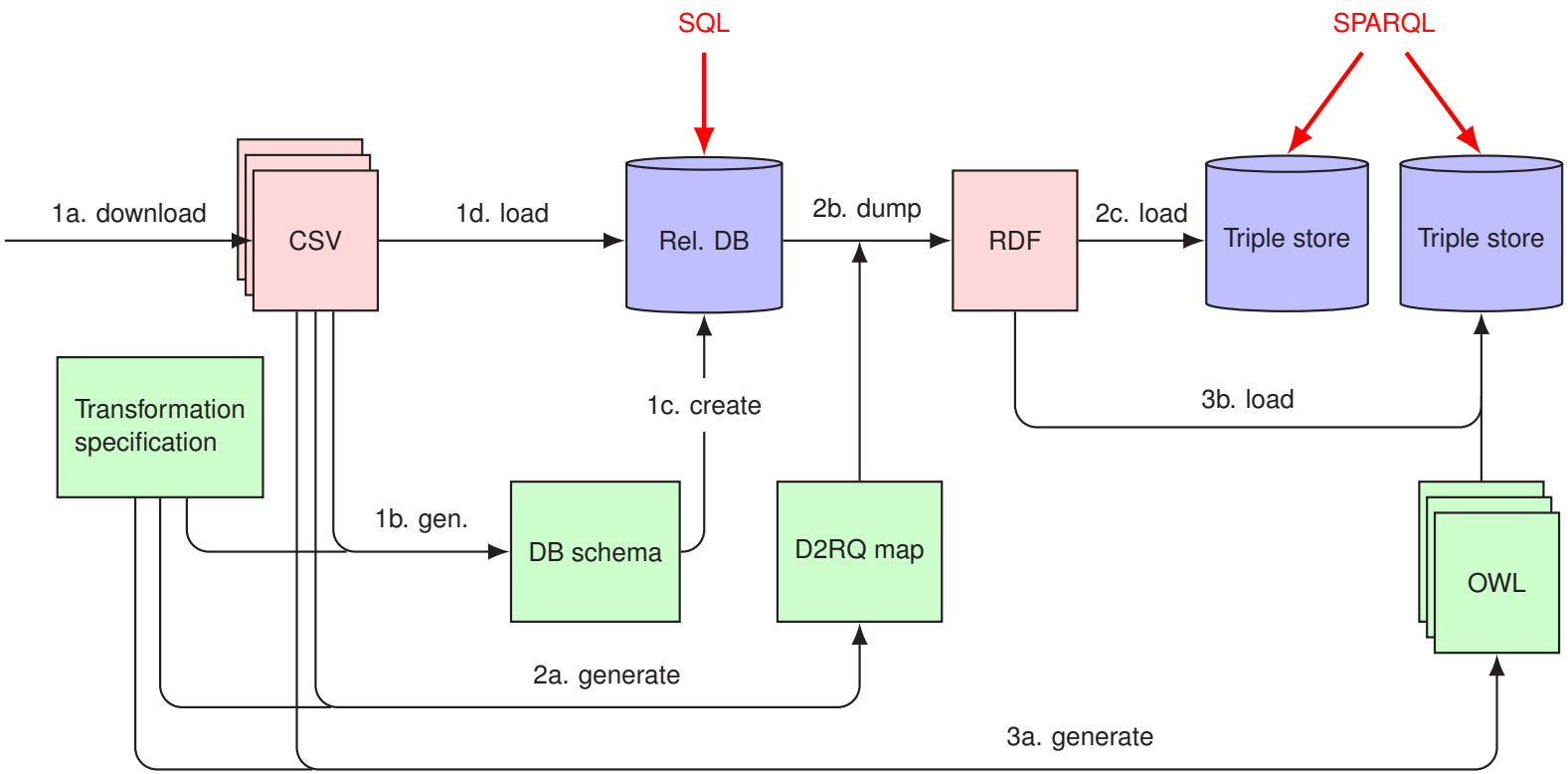
3. Create ontology representation

   b) Load RDF + OWL into triple store



Conversion Process

We test the different representations in our query evaluation experiment



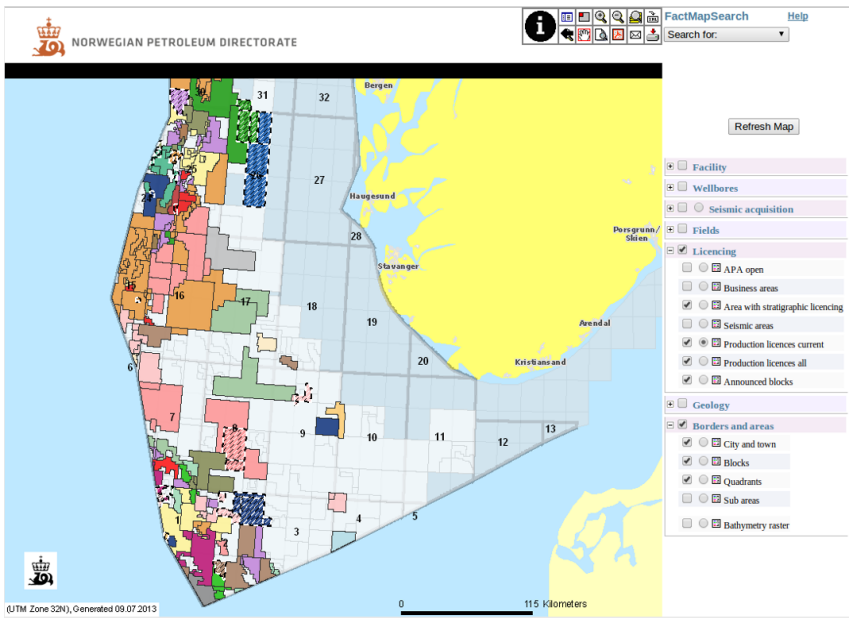
Results

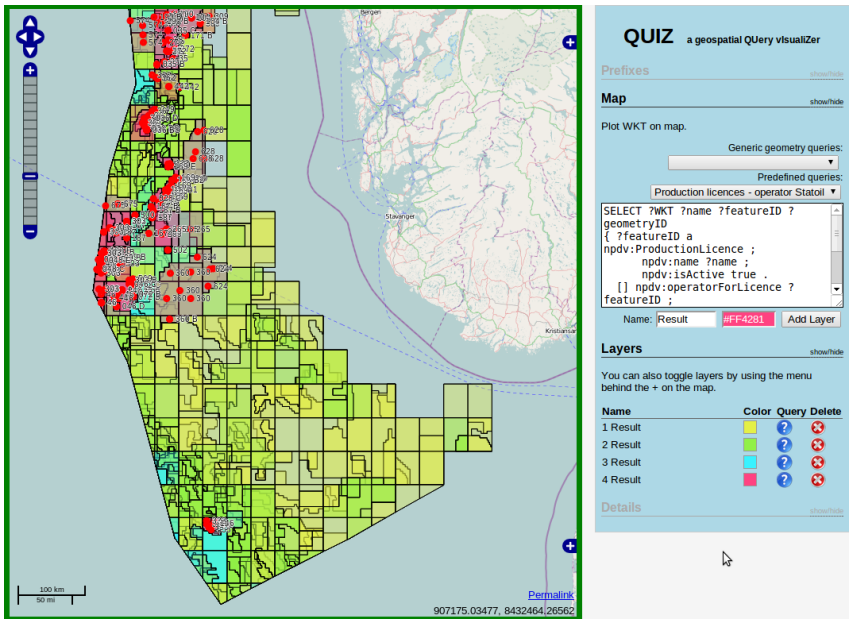
- Availability of data is better
  - Each thing has an identifier
  - Find information using identifer
- Easier to exhange and integrate data
- More data is available
  - Arbitrary SPARQL queries
  - New facts added by reasoning with semantics
- Semantics are more explicit
- All using standardised languages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

Outline

Example: Map Visualization

NPD FactMap vs. “our” SPARQL map visualiser.

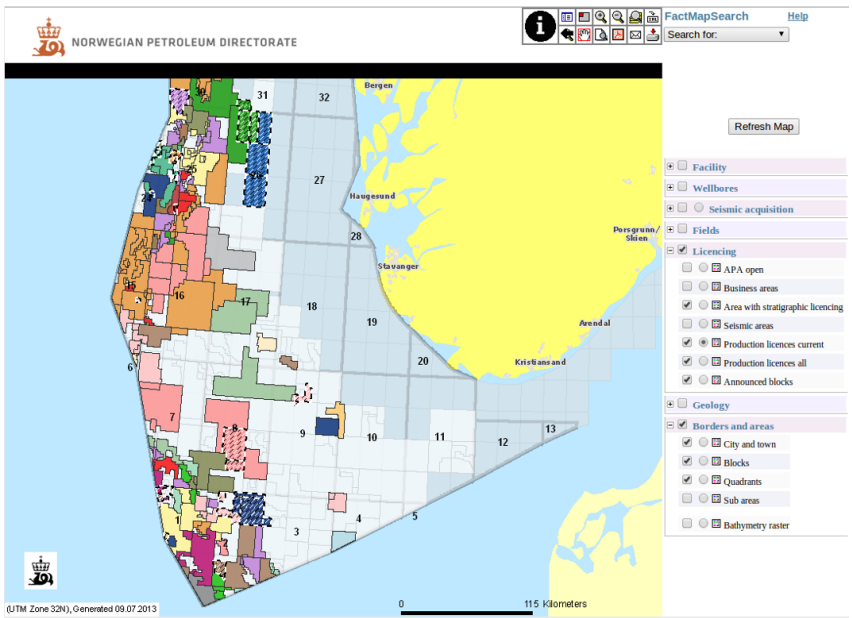


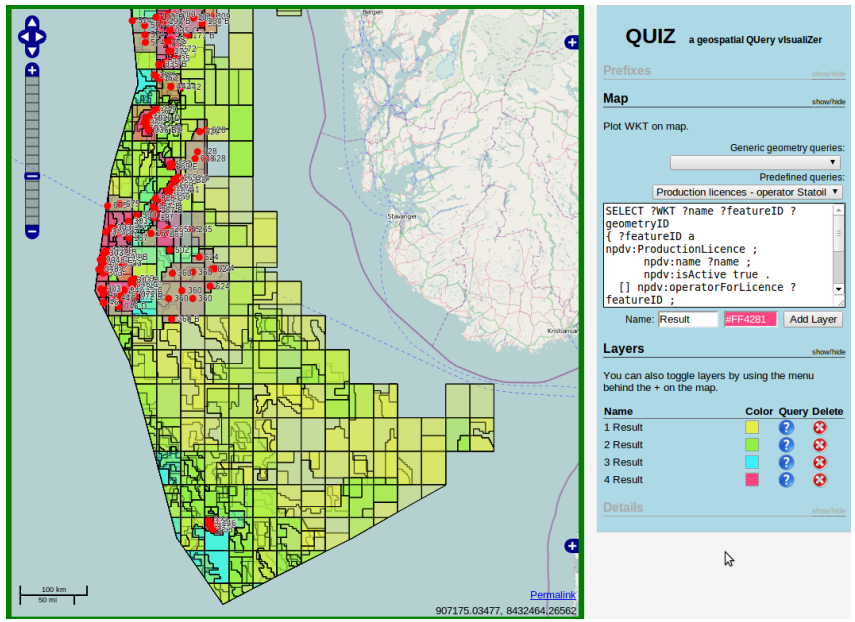


- Official map app.

Example: Map Visualization

NPD FactMap vs. “our” SPARQL map visualiser.

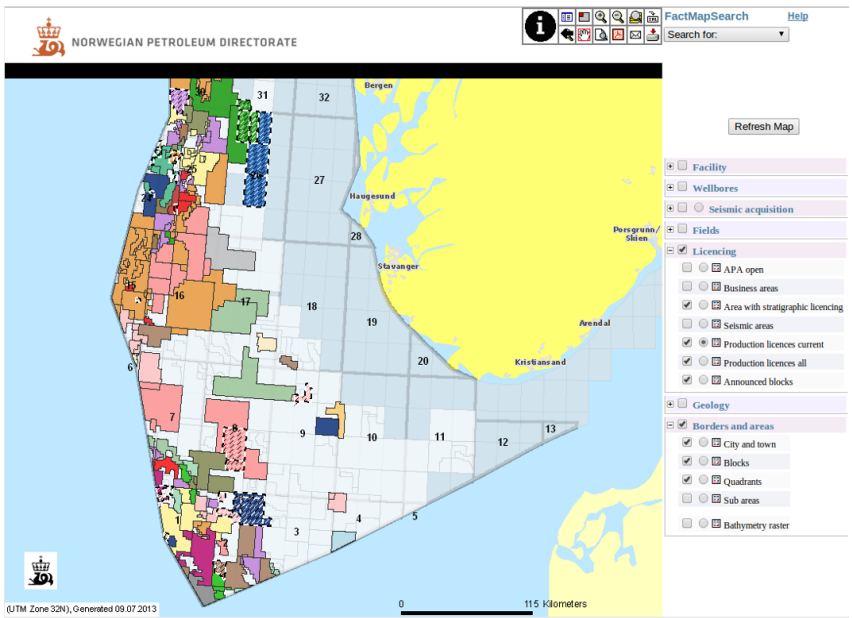


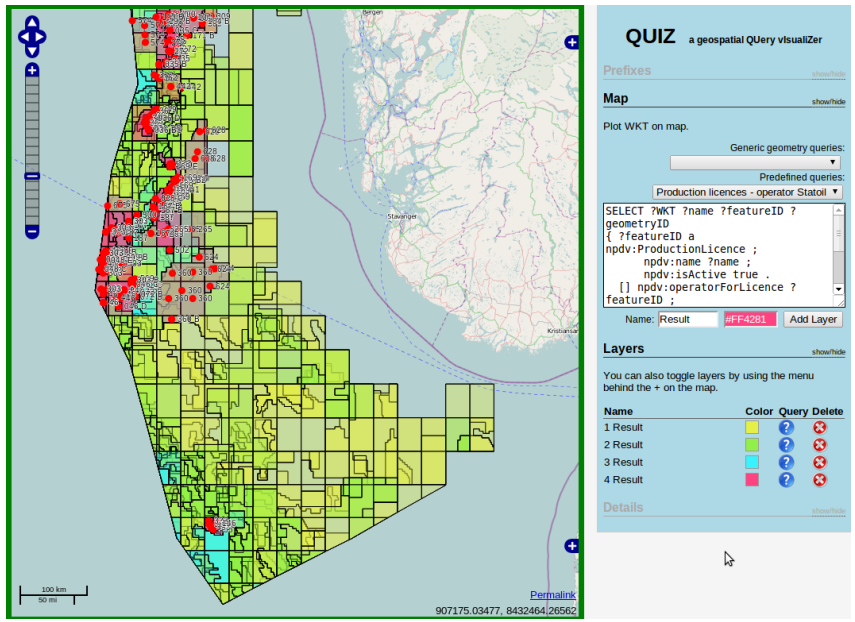


- Official map app.
- Server side

Example: Map Visualization

NPD FactMap vs. “our” SPARQL map visualiser.

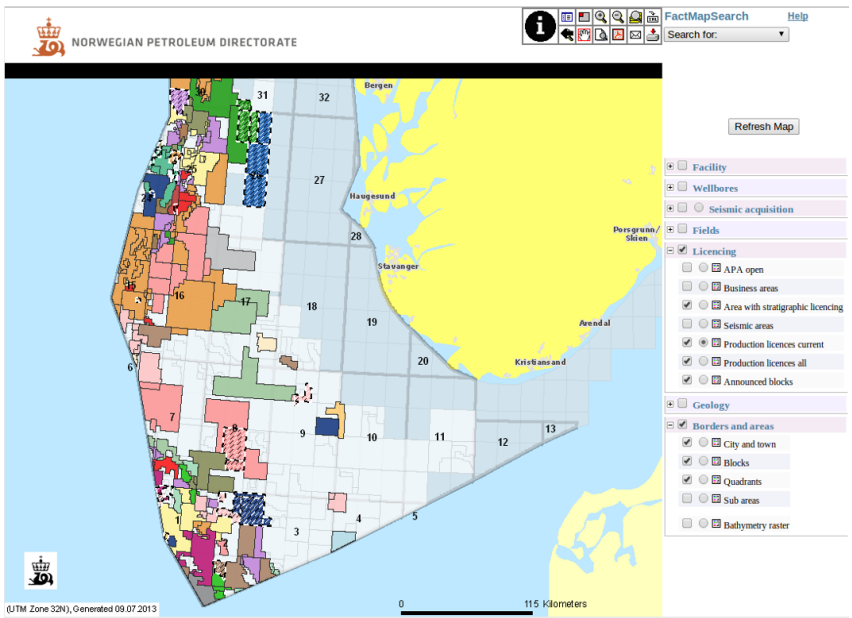


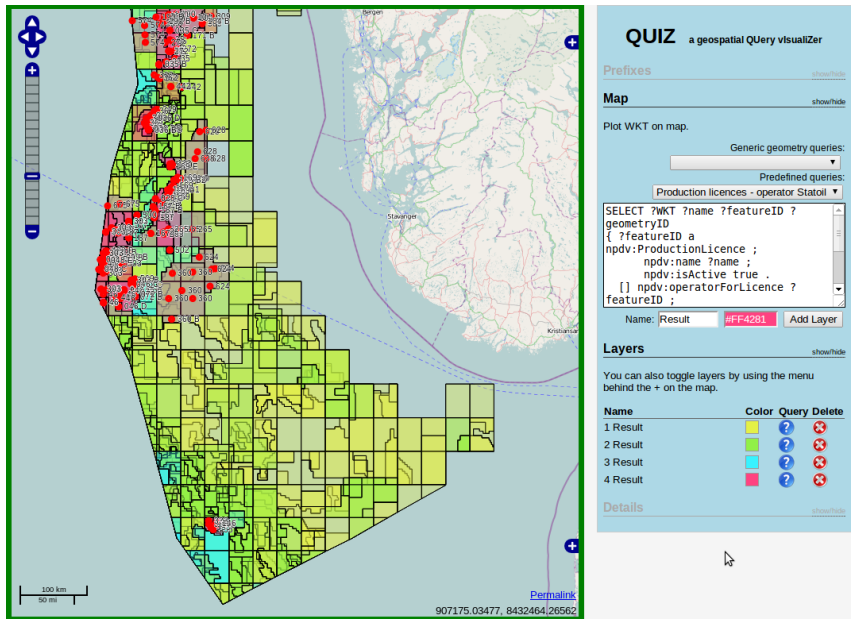


- Official map app.
- Server side
- Predefined queries only

Example: Map Visualization

NPD FactMap vs. “our” SPARQL map visualiser.

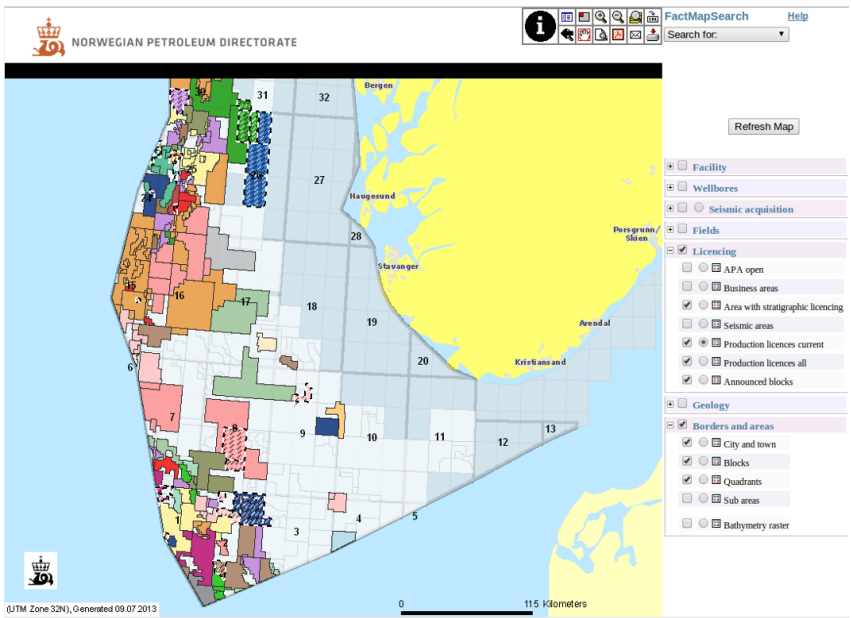




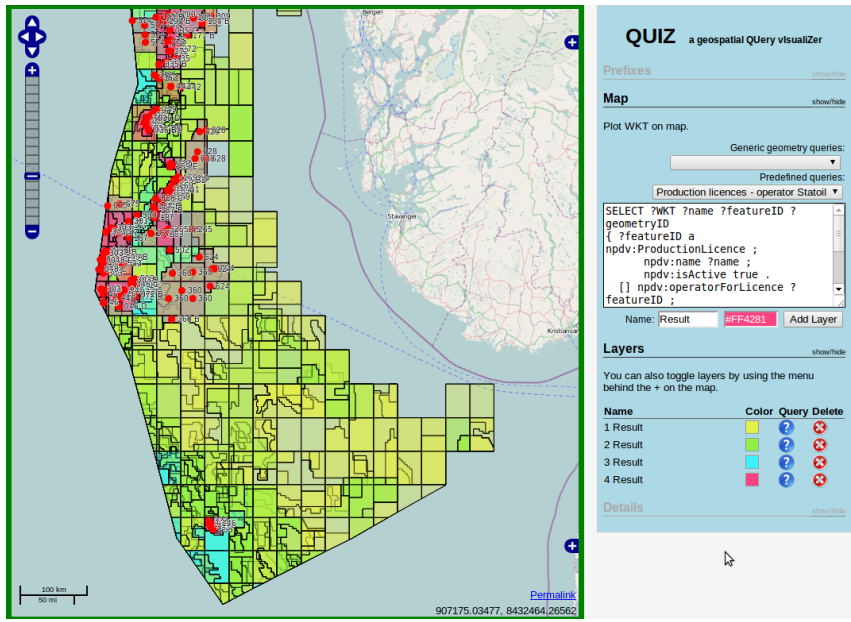
- Official map app.
- Server side
- Predefined queries only
- Map results only

Example: Map Visualization

NPD FactMap vs. “our” SPARQL map visualiser.



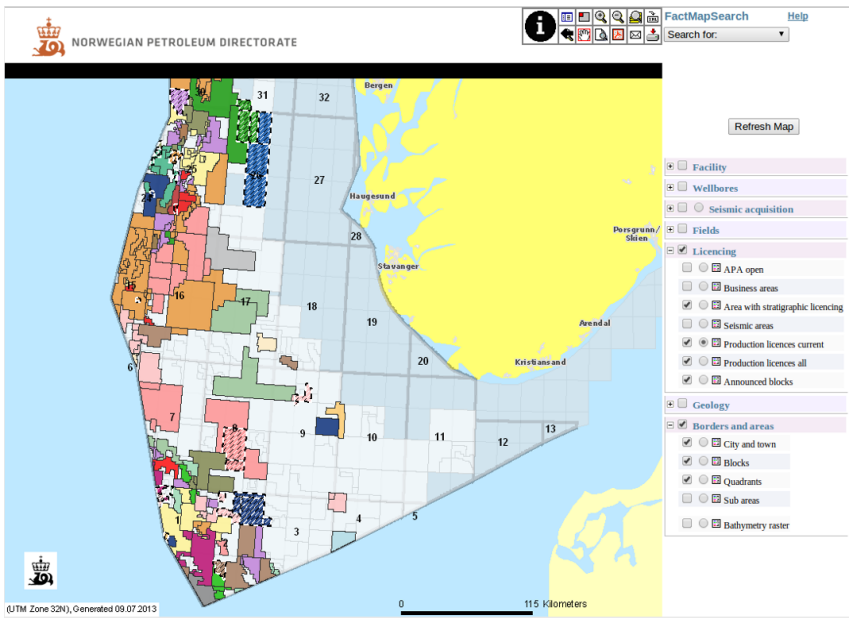
- Official map app.
- Server side
- Predefined queries only
- Map results only



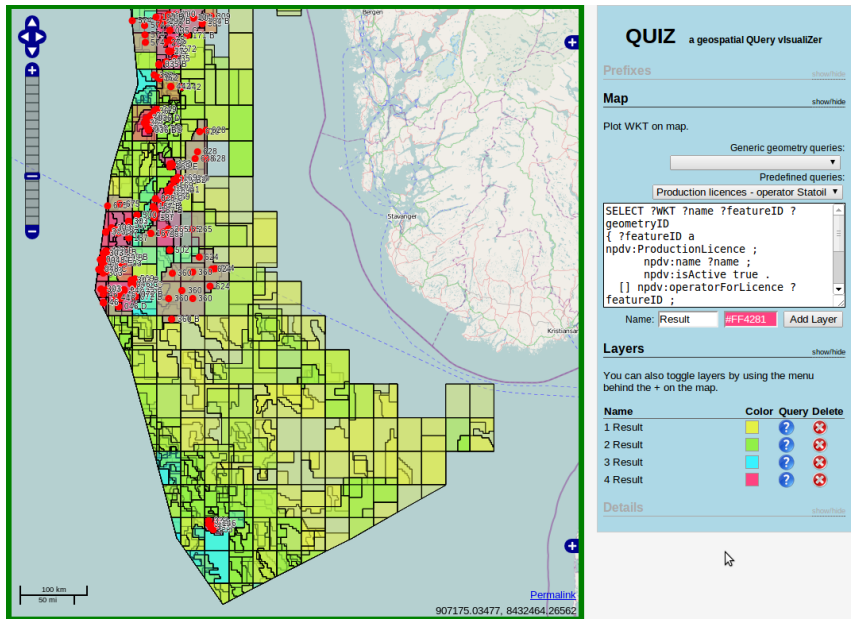
- Generic map app

Example: Map Visualization

NPD FactMap vs. “our” SPARQL map visualiser.



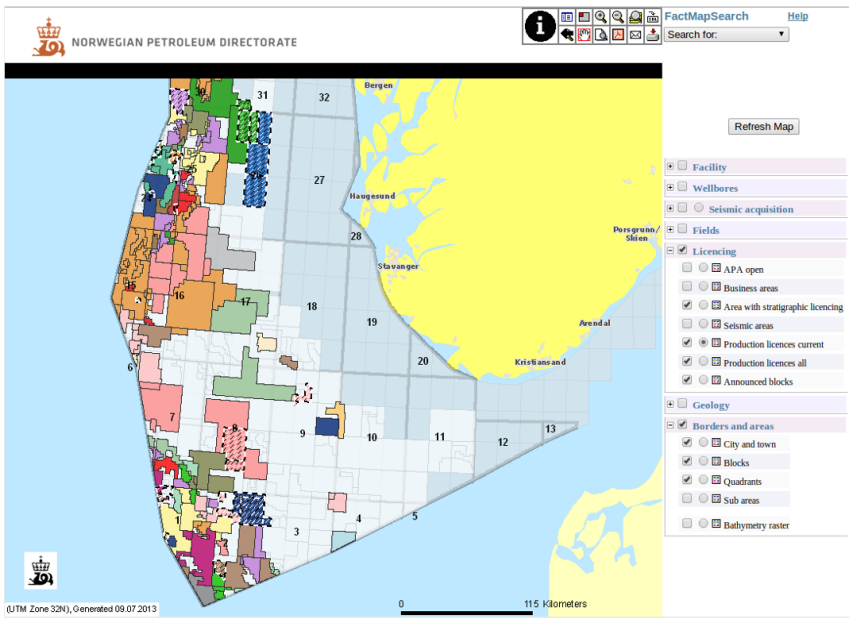
- Official map app.
- Server side
- Predefined queries only
- Map results only

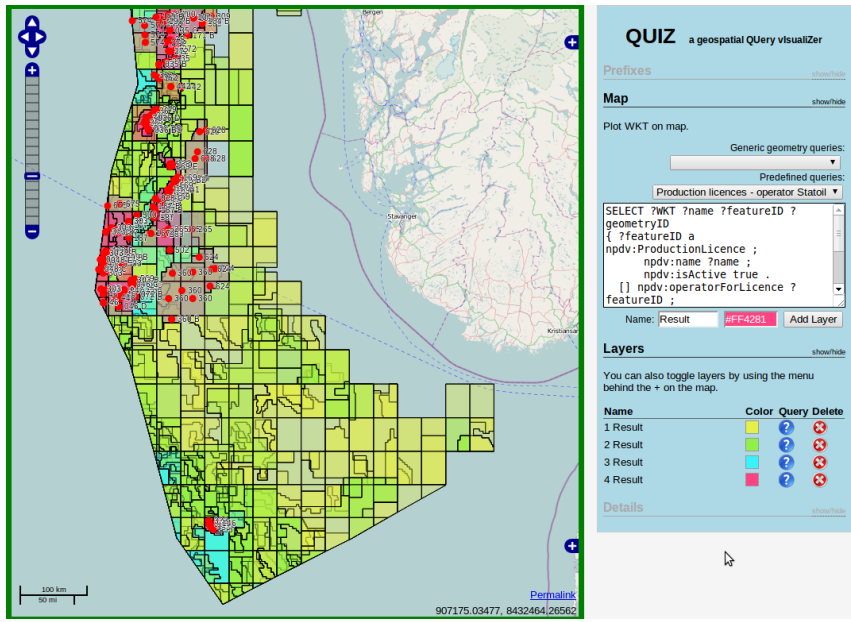


- Generic map app
- Client side, SPARQL based, open source

Example: Map Visualization

NPD FactMap vs. “our” SPARQL map visualiser.



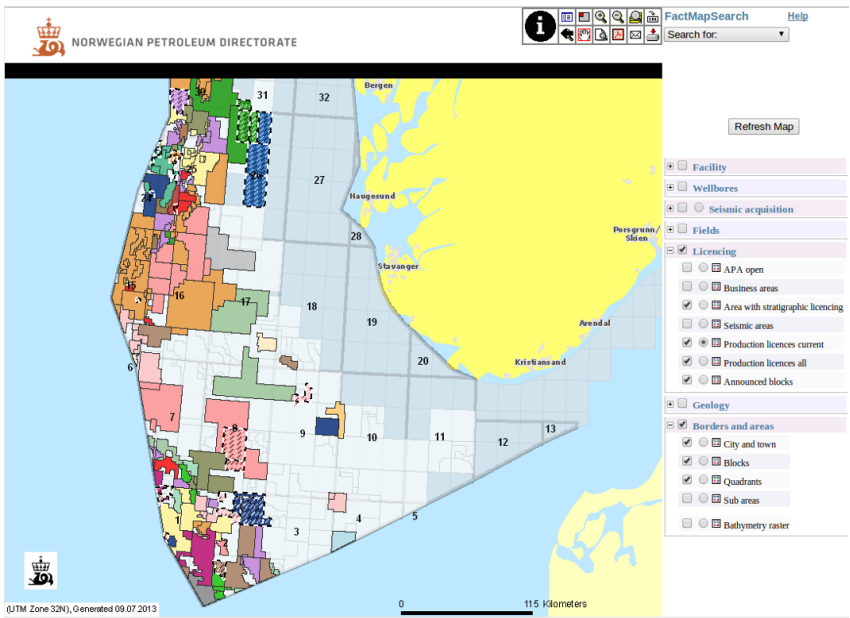


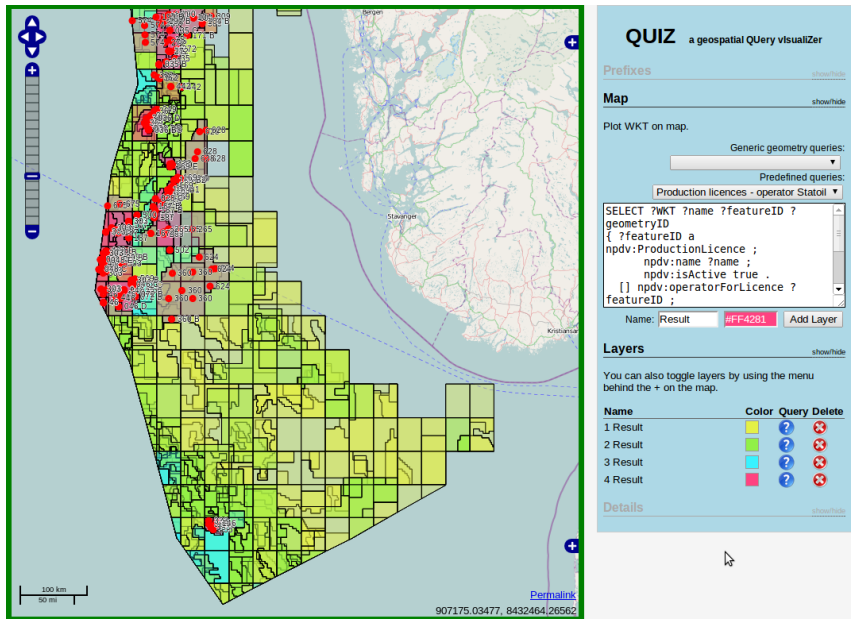
- Official map app.
- Server side
- Predefined queries only
- Map results only

- 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/`
- Field
- Company
  - Name: Statoil Petroleum ASA
- Facility
  - facility function: GAS PRODUCER
- Show SPARQL
- Run query
- Select SNØHVIT

NPD FactPages

Project website

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