Feng *Open Geospatial Data, Software and Standards* (2016) 1:6
DOI 10.1186/s40965-016-0007-y

Open Geospatial Data, Software and Standards

ORIGINAL ARTICLE

Open Access

Geodata for everyone - model-driven development and an example of INSPIRE WFS service

Meixia Feng

Abstract

Background: On the first January 2013, everyone in Denmark received a wonderful new year present from the Danish government: "Free and Open Geo-data". This means that everyone is free to use, reuse and redistribute the geographic data published by the Danish Geodata Agency. As Denmark's central public source of geographic data, the Danish Geodata Agency is responsible for surveying, mapping and land registering of Denmark, Greenland, the Faroe Island and all waters associated with these. The overall goal of the organization is to supply and ensure that everyone in the Danish society has access to reliable and accurate maps and geographic information on all parts of the country.

Methods: This article presents a model-driven development process which has been established in the Danish Geodata Agency for unified distribution of geographic data. The process starts from the conceptual modeling of the geographic data to the end where data are distributed via WFS services and download services.

Results: By adopting this model-driven development process, we have automated the process from data modeling to data distribution. We have defined the national data models which comply to INSPIRE data models, thus we maximize the re-use of national spatial data for the INSPIRE services.

Conclusions: From our experience of using this model-driven development process, we conclude that it improves our INSPIRE compliant, national spatial data infrastructure and ensures free and open access to the high quality geographic data.

Keywords: Model-drive development, Geo-data, WFS, Basic data, eGovernment, UML, GML, ShapeChange, Database schema, DDL, GO Publisher, INSPIRE, QGIS

Background

Free and Open Geo-data is part of basic-data programme [1], which was launched in the autumn of 2012 as part of the country's eGovernment strategy for 2011–2015. The basic-data programme defines basic data as core information about individuals, businesses, real properties, buildings, addresses, geography, etc. As illustrated in Fig. 1, the basic data should be efficiently updated at one place, distributed via a common data distributor, and used by everyone. By releasing basic data, the public sector wants to remove the barriers of using public-sector basic data, improve efficiency, and create growth of new and innovative digital services.

Correspondence: mefen@sdfe.dk
Agency for Data Supply and Efficiency, Rentemestervej 8, 2400 Copenhagen, NV, Denmark

Making high-quality basic data open and easy-to-access involves five processes such as:

1. Make all public basic data free. Everyone can freely use public-sector basic data for commercial as well as for non-commercial purposes.
2. Warm wash of all public data. Enhance the quality of data and improve the coherence in the basic data registers.
3. Technical harmonization. Ensure that data conform to the same technical requirements.
4. Common public data distributor. Establish a common single distribution system to improve the distribution of data. From the beginning of 2016 the Data Distributor will distribute geographic data such as maps, cadastral maps and other geographic data.

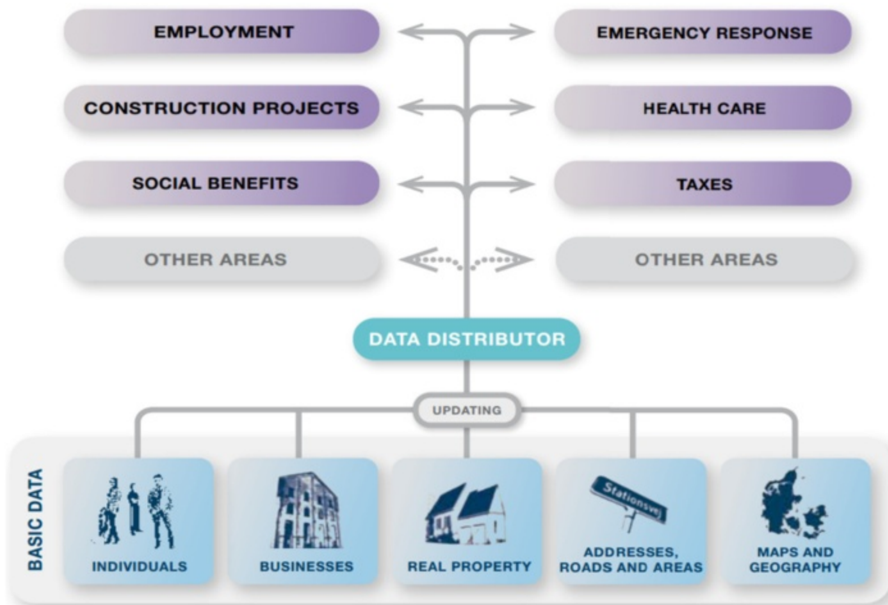


Fig. 1 basic data infrastructure

5. Governance. Establish a cross-institutional national board to ensure efficient, effective and coordinated development and use of basic data.

The basic-data programme requires that all the authoritative basic data should be defined and standardized according to the same methods. As Denmark's central public source of geographic data, the Danish Geodata Agency has established a set of guidelines for future modeling and distribution of spatial data. Based on the guidelines, a model-driven development process has also been established to automate the process from business modeling to data distribution. This paper describes this model-driven development process and uses a practical implementation of INSPIRE Web Feature Service (WFS) as an example to show how this process improves our INSPIRE compliant, national spatial data infrastructure. Section 2 describes the architecture of our model-driven development process and the tools used of the process. Section 3 presents a WFS INSPIRE service as an example of our results and discuss the challenges. In Section 4 some conclusions, as well as future works, are presented.

Methods

Model-driven development of Geo-data

Geographic information has become a more and more important part of the application domains. To incorporate geographic features from the real world into digital data representations, modeling is a powerful abstraction mechanism. The relevance of conceptual modeling is widely applied in spatial data infrastructure (SDI) initiatives such as INSPIRE programme and European Location Framework (ELF) programme. Implementing an effective data distribution environment adhering to the conceptual data models is the primary idea behind establishing this model-driven development process in our organization. In this section the technical implementation and the tools used in the implementation are described.

The model-driven development process starts from the data modeling in Unified Modeling Language [2] (UML) to the end where data are distributed via WFS services and download services. ShapeChange [3] is used as a bridge between data modeling and data distribution. ShapeChange is a Java-based tool that takes an ISO 19109 compliant UML application schema and translates it into other output formats, for example Geography Markup Language (GML) application schema. The overview of the process and the components of the process are illustrated in Fig. 2 below.

From modeling to documentation

The first step in the process is to use UML to describe the business process or production system in Enterprise Architect. The basic-data committee has provided a model document, where common model rules and common basic metadata attributes for object classes are defined. It is a requirement that our geo-data must follow the model rules, have the required metadata attributes, and comply with the ISO standard and INSPIRE specification. Object Constraint Language [4] (OCL) is used for defining constraints, since ShapeChange supports parsing of OCL constrains. Data specification documentation for

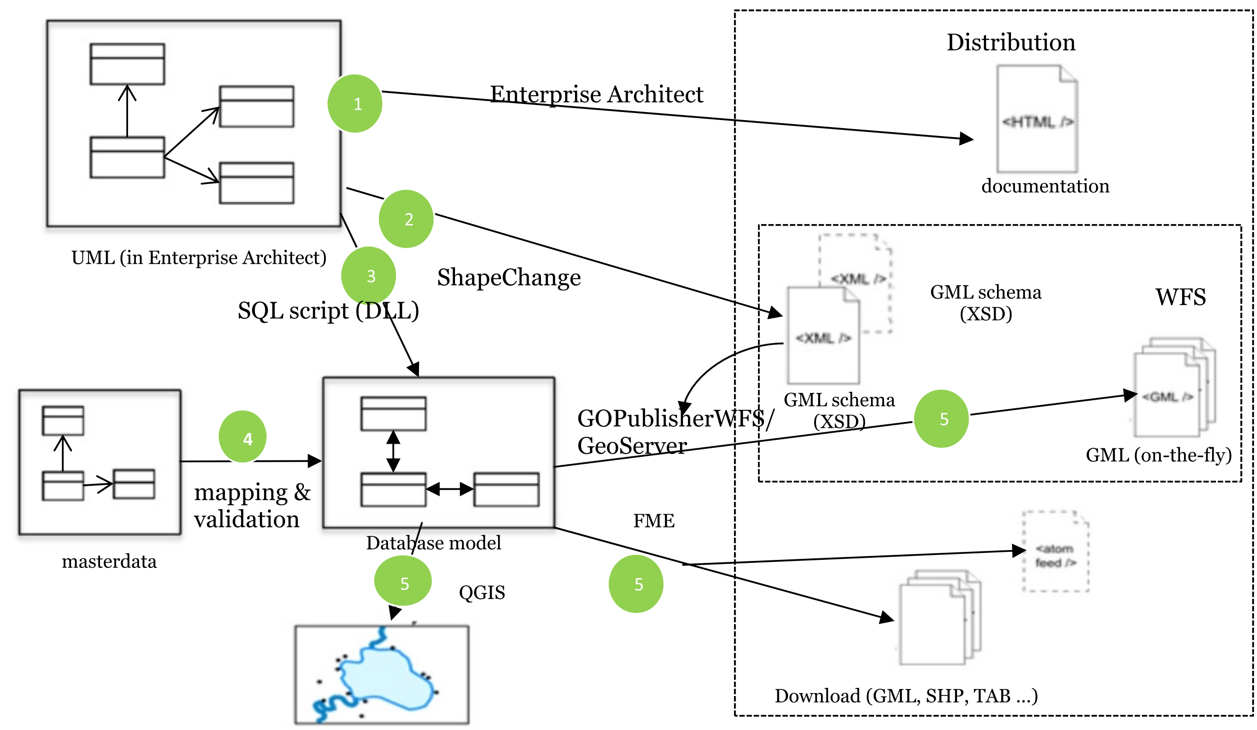


Fig. 2 Overview of Model-driven development process

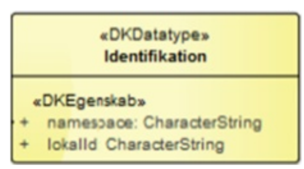
end-users is exported via dynamic HTML by the built-in functionality in Enterprise Architect for convenient browsing via the web.

From modeling to GML schema

After the UML model has been defined, we use ShapeChange to generate GML application schemas from the UML. ShapeChange uses a set of encoding rules to convert a UML model to GML schemas. Which combination of encoding rules will be used in ShapeChange will depend on which version of GML is desired in the output. By default, ShapeChange generates schemas that conform to the GML 3.2 standard, which is also used in the INSPIRE GML schemas. The encoding rules for the conversion are described in UML-to-GML application schema encoding rules of [OGC 07–036] Annex E [5].

Figure 3 shows a concrete example of our national class AdministrativInddeling (AdministrativeBoundaries in English) which is described in UML is converted to a complexType in XML schema.

The conversion includes among others the properties of the UML class are converted to the elements of the XML complexType, and the UML data types are converted to the proper XML data types. Properties of a class in UML model are not always defined as simple types. In our example above, the property "id" in UML model uses a complex feature type "Identifikation" which is defined as:

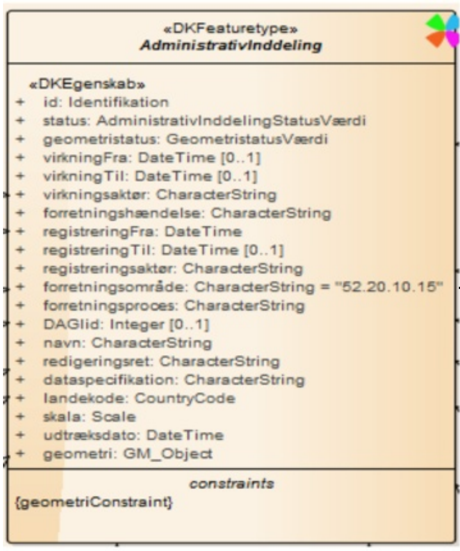


The resulting structures can become quite complex, which makes it difficult to convert data into formats that do not support such complexity, for example Shapefiles. To simplify the complexity, ShapeChange provides a flattening function. The Flattener supports a set of flattening rules. In our case, the property "id" is flattened by using rule "rule-trf-prop-flatten-types", and the "landekode" (country code in English) which is defined as "gmd:Country_PropertyType" is flattened by using rule "rule-trf-prop-flatten-codelists" to a string type.

The GML schemas generated from this process will be used later in the process by other tools to distribute or read the data.

From modeling to database schema

As depicted in step 3 of Fig. 2, ShapeChange is also used to generate database schema from UML model. The database schema, in our system called geodatabank, will




```xml
<!--XML Schema document created by ShapeChange - http://shapechange.net/-->
<element abstract="true" name="AdministrativInddeling"
substitutionGroup="gml:AbstractFeature" type="dagi:AdministrativInddelingType"/>
<complexType abstract="true" name="AdministrativInddelingType">
  <complexContent>
    <extension base="gml:AbstractFeatureType">
      <sequence>
        <element name="id" type="gd:IdentifikationType"/>
        <element name="status" type="dagi:AdministrativInddelingStatusVaerdiType"/>
        <element name="geometristatus" type="geo:GeometristatusVaerdiType"/>
        <element minOccurs="0" name="virkningFra" type="dateTime"/>
        <element minOccurs="0" name="virkningTil" type="dateTime"/>
        <element name="virkningsaktoer" type="string"/>
        <element name="forretningshaendelse" type="string"/>
        <element name="registreringFra" type="dateTime"/>
        <element minOccurs="0" name="registreringTil" type="dateTime"/>
        <element name="registreringsaktoer" type="string"/>
        <element name="forretningsomraade" type="string"/>
        <element name="forretningsproces" type="string"/>
        <element minOccurs="0" name="DAGIid" type="integer"/>
        <element name="navn" type="string"/>
        <element name="redigeringsret" type="string"/>
        <element name="dataspecifikation" type="string"/>
        <element name="landekode" type="gmd:Country_PropertyType"/>
        <element name="skala" type="string"/>
        <element name="udtraeksdato" type="dateTime"/>
        <element name="geometri" type="gml:GeometryPropertyType"/>
      </sequence>
    </extension>
  </complexContent>
</complexType>
```

Fig. 3 (left) UML to (right) GML Conversion

serve as the container of the data. It is generated by a set of SQL Data Definition Language (DDL) instructions which are derived from UML models. For this purpose, the output targert "sqlddl" of ShapeChange is used. In order to achieve a simple structure of tables, as well as to cope with the potential complexity of the application schema, the Flattener of ShapeChange is used again to execute a number of flattening rules, such as rule-trf-cls-flatten-inheritance, rule-trf-prop-flatten-multiplicity, rule-trf-prop-flatten-types, rule-trf-all-flatten-name. The general principals used in mapping an object-oriented UML model to a relational database are:

- Each concrete class is mapped to a table.
- Object relationships in UML models are implemented as primary key and foreign key in database.
- Each property is converted to a column in the database.

After the database schema has been generated, it's also important to make the necessary performance tuning such as building effective indexes.

With the distribution database schema ready, the data from production can be loaded into geodatabank by using a set of mapping and validation rules.

From modeling to distribution

With data in the database and the application schema generated by ShapeChange, we can now distribute data via different channels.

- We use GOPublisherWFS or GeoServer to set up WFS services and distribute data via WFS services. These services are easy to access and are provided free of charge. For details, see http://www.kortforsyningen.dk/.
- The internal users can use GIS tools to connect to the database directly.
- FME is used to generate predefined dataset in different formats for download. They are available for download via our web shop free of charge. For details, see http://download.kortforsyningen.dk/.

Results and discussion

Using model-driven development to implement INSPIRE WFS service

INSPIRE is a proposed European Directive that aims to create the legal framework for making relevant, standardized and quality assured spatial data available for the Member States of the European Union (EU) [6]. INSPIRE is based on the infrastructures for spatial information that are created by the Member States and that are made compatible with common implementing rules. For each Member State it is necessary to establish a search and distributions system based on network services and portals to make access to spatial data easier and to improve compatibility and interoperability, as well as it is necessary to develop agreements on access to and use of spatial data. The Joint Research Centre of the European Commission has developed a prototype EU Geoportal that is an entry point to the discovery and viewing of spatial data sets and services [7]. The implementation of a prototype EU geoportal helps to better understand user requirements and identify potential problems, and to

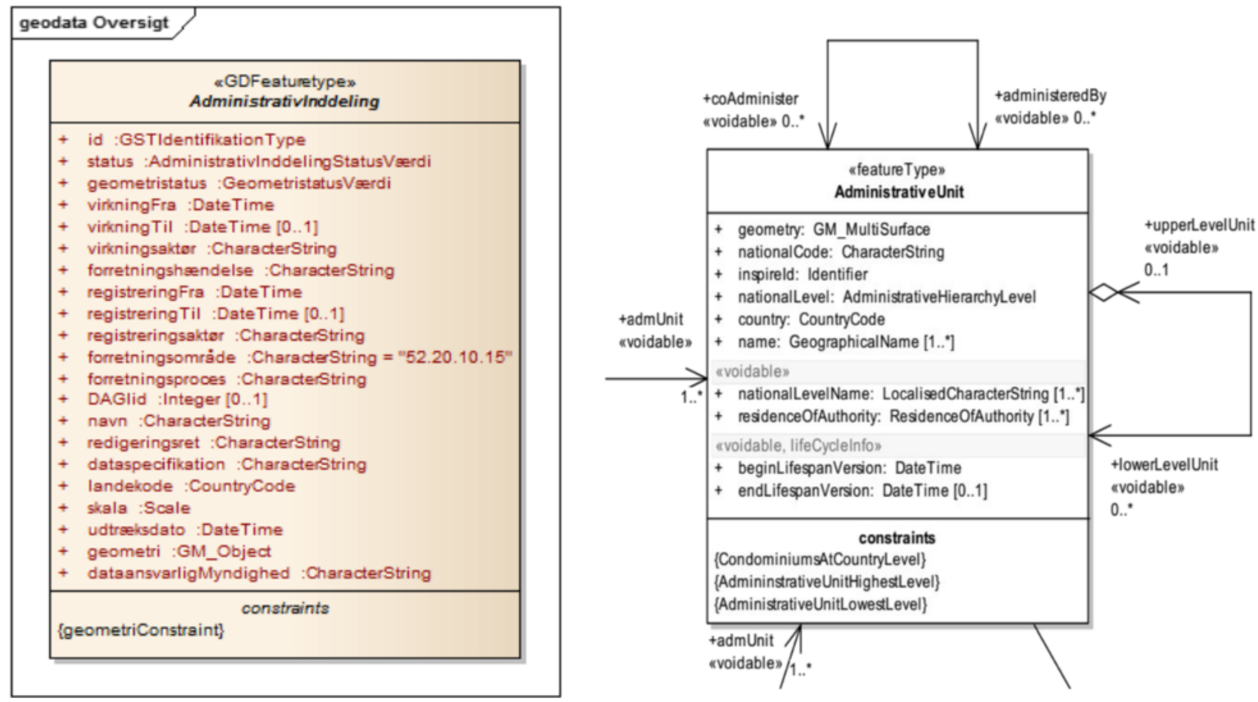


Fig. 4 (left) National Object to (right) INSPIRE Theme

define the standards-based technical specifications of the actual implementation of INSPIRE.

The INSPIRE Directive is transposed into Danish law concerning the setup of an Infrastructure for Geographic Information. This law came into force in May 2009 and hereby geographic information becomes a key component in eGovernment. The Danish implementation of INSPIRE is as far as possible be linked to actual developments of the Danish SDI and the eGovernment initiative. With free and open Geo-data as part of the country's eGovernment strategy for 2011–2015, it is expected to overcome the data sharing challenge which is one of the three strategic challenges for SDI development addressed by Masser et al. [8]. In order to overcome the challenge on enabling platforms, the Danish SDI includes also platforms such as www.kortforsyningen.dk and download.kortforsyningen.dk.

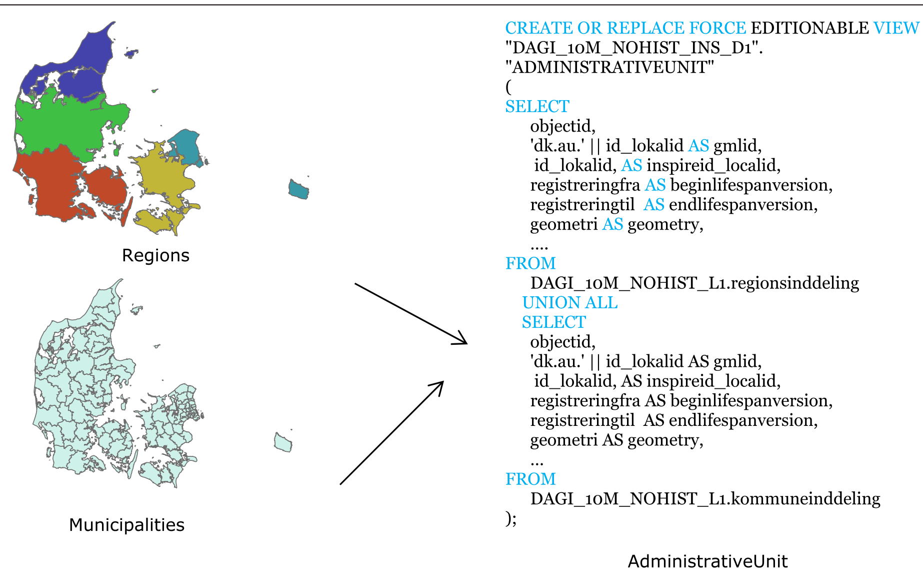


Fig. 5 National Dataset to INSPIRE Dataset

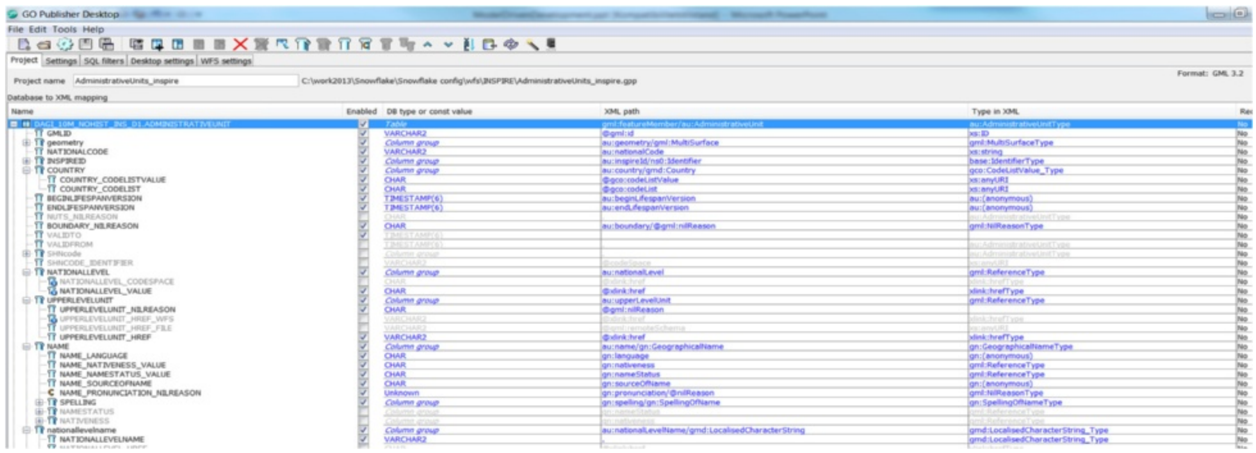


Fig. 6 GO Publisher Desktop

These platforms provide the access to the geographic information and services, and link the data providers and data users.

From national object to INSPIRE theme

Our basic data are "Inspired by" INSPIRE principles [9]. This means that data are collected only once, distributed according to different specifications and shared with many users and applications. In this chapter, we are going to show how we re-use the national data produced by this model-driven development process to implement an INSPIRE WFS service.

We continue to take our national model AdministrativInddeling (AdministrativeBoundaries in English) as an example. It is intended to fulfil the INSPIRE specification on Administrative Units theme [10] in Fig. 4.

From national dataset to INSPIRE dataset

To make the data ready in the database for implementing the INSPIRE service, we created a view which derives and reconstructs data about regions and municipalities from the National AdministrativInddeling dataset, see Fig. 5. With this database view, we tried to create the source data which match the INSPIRE output as closely as possible. For example, for use as value of gml:id attribute in INSPIRE feature, we created a "gmlid" column in the view from the "id_lokalid" column in the table.

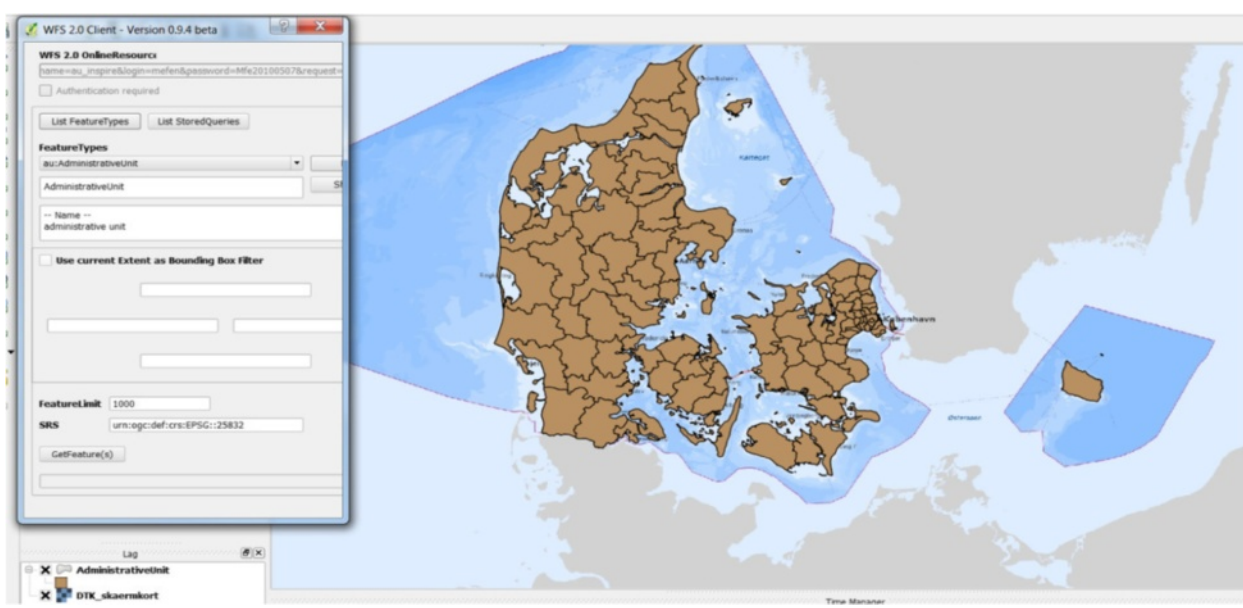
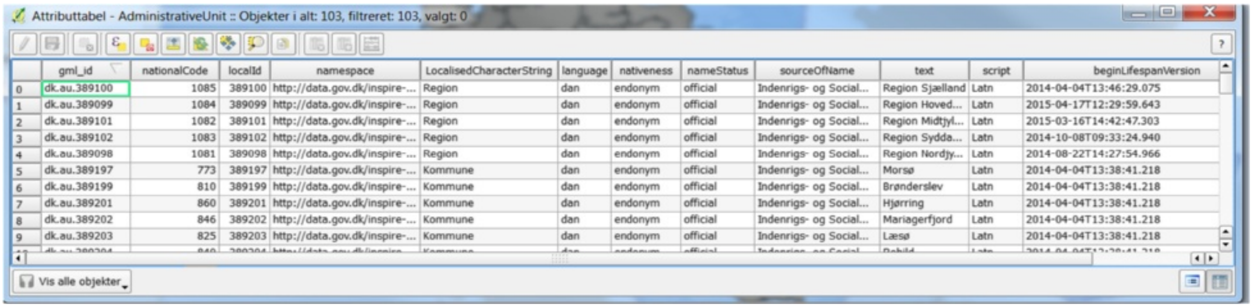


Fig. 7 WFS AdministrativeUnit in QGIS



	gml_id	nationalCode	localId	namespace	LocalisedCharacterString	language	nativeness	nameStatus	sourceOfName	text	script	beginLifespanVersion
0	dk.au.389100	1085	389100	http://data.gov.dk/inspire-...	Region	dan	endonym	official	Indenrigs- og Social...	Region Sjælland	Latn	2014-04-04T13:46:29.075
1	dk.au.389099	1084	389099	http://data.gov.dk/inspire-...	Region	dan	endonym	official	Indenrigs- og Social...	Region Hoved...	Latn	2015-04-17T12:29:59.643
2	dk.au.389101	1082	389101	http://data.gov.dk/inspire-...	Region	dan	endonym	official	Indenrigs- og Social...	Region Midtjyl...	Latn	2015-03-16T14:42:47.303
3	dk.au.389102	1083	389102	http://data.gov.dk/inspire-...	Region	dan	endonym	official	Indenrigs- og Social...	Region Sydda...	Latn	2014-10-08T09:33:24.940
4	dk.au.389098	1081	389098	http://data.gov.dk/inspire-...	Region	dan	endonym	official	Indenrigs- og Social...	Region Nordjy...	Latn	2014-08-22T14:27:54.966
5	dk.au.389197	773	389197	http://data.gov.dk/inspire-...	Kommune	dan	endonym	official	Indenrigs- og Social...	Morsø	Latn	2014-04-04T13:38:41.218
6	dk.au.389199	810	389199	http://data.gov.dk/inspire-...	Kommune	dan	endonym	official	Indenrigs- og Social...	Brønderslev	Latn	2014-04-04T13:38:41.218
7	dk.au.389201	860	389201	http://data.gov.dk/inspire-...	Kommune	dan	endonym	official	Indenrigs- og Social...	Hjørring	Latn	2014-04-04T13:38:41.218
8	dk.au.389202	846	389202	http://data.gov.dk/inspire-...	Kommune	dan	endonym	official	Indenrigs- og Social...	Mariagerfjord	Latn	2014-04-04T13:38:41.218
9	dk.au.389203	825	389203	http://data.gov.dk/inspire-...	Kommune	dan	endonym	official	Indenrigs- og Social...	Læsø	Latn	2014-04-04T13:38:41.218

Fig. 8 QGIS attribute table

Mapping dataset to INSPIRE schema

We use GO Publisher Desktop to map the columns of database to the attributes and geometry of the inspire application schema. As Fig. 6 shows, this tool provides a comprehensive graphical user interface with 'point and click' functionality for mapping. After the mapping is done, this tool generates a web archive (war) file, which can be deployed as web service and publishes data via Web Feature Service. This Web Feature Service supports WFS standards 1.0.0, 1.1.0 and 2.0.0. It copes with providing valid GML 3.2-encodings of rich geo objects such as the INSPIRE features. It supports Reference to other features or GML objects, properties that contain GML core data types which are not geometries such as code types or units of measure.

To ensure the mapping from our national data to inspire data is efficient and lossless, we conducted the following attribute mapping. First, we identified the attributes in the INSPIRE schema that we have national data to fulfil. Secondly, we identified the attributes in the INSPIRE schema that can be mapped to a constant value. Thirdly, we identified the attributes in the INSPIRE schema that should refer to the data sets in INSPIRE, for example "CountryCode". Lastly, there are also quite some "nillable" attributes in the INSPIRE schema that we have no national data to fulfil. They were defined as xsi:nil = "true" with proper nilReason.

Access WFS with GIS client

One of the free GIS clients, which have support for showing and querying WFS 2.0, is the QGIS with WFS2.0 [11] plugin. It is used to access the WFS AdministrativeUnit as Fig. 7 shows.

Fig. 9 WFS object in GML

Comparing the data in the QGIS attribute table in Fig. 8 to the data in GML from the same WFS AdministrativeUnit in Fig. 9, we can see that there are some challenges for GIS clients to cope with GML 3.2-encodings of complex objects, such as:

- no Access to codelists
- structure is not reflected, they are flattened in the attribute table
- attributes such as units, nill-reasons are not accessible (e.g. units, nill-resons)

Conclusion

In this paper, we have presented the background of establishing a model-driven development process and described how we use it to develop and distribute our geographical data. By adopting this model-driven development process, we have automated the process from business modeling to data distribution and reduced the time taken to implement a new product. This process ensures consistency between the various implementation schemas (GML, DDL, ArcGIS, etc.) as they are no longer manually created by different developers. The defined models in our process comply to INSPIRE data models, which maximize the re-use of national spatial data for the INSPIRE services. As an important part of the national SDI, it has resulted in a much broader access to the national geo-data collection and greater use of these data in the public and private sectors.

Currently, we continue producing INSPIRE and ELF services by using this model-driven process. We're sharing experience and results with other European countries.

As future work, we will extend this process to other public geospatial data producers. We expect that the common data modeling principles and standards will improve the effectiveness and interoperability of national and regional SDIs. It is also expected that this model-driven development process will ensure the INSPIRE and ELF implementation in Denmark on a more authoritative basis.

Competing interests

The author declares that she has no competing interests.

Received: 15 February 2016 Accepted: 2 June 2016
Published online: 16 June 2016

References

1. Good Basic Data for everyone" - The Danish Government/Local Government Denmark, October 2012. http://uk.fm.dk/publications/2012/good-basic-data-for-everyone/~/media/Publikationer/Imported/2012/Gode%20grunddata%20til%20alle/BasicData_UK_web_2012.10.08.ashx.
2. Unified Modeling Language. https://en.wikipedia.org/wiki/Unified_Modeling_Language. [Accessed 10 June 2016]
3. ShapeChange. Processing application schemas for geographic information. http://shapechange.net.
4. Object Constraint Language. http://modeling-languages.com/ocl-tutorial/. [Accessed 10 June 2016]
5. Inc OGC. OpenGIS Geography Markup Language (GML) encoding standard. 2007.
6. Craglia M, Annoni A. INSPIRE: an innovative approach to the development of spatial data infrastructures in Europe. Research and theory in advancing spatial data infrastructure concepts. 2007. p. 93–105.
7. Bernard L, Kanellopoulos I, Annoni A, Smits P. The European geoportal–one step towards the establishment of a European spatial data infrastructure. Comp Environ Urban Syst. 2005;29(1):15–31.
8. Masser I, Rajabifard A, Williamson I. Spatially enabling governments through SDI implementation. Int J Geogr Inf Sci. 2008;22(1):5–20.
9. About INSPIRE. http://inspire.ec.europa.eu/index.cfm/pageid/48. [Accessed 10 June 2016]
10. INSPIRE. Infrastructure for spatial information in Europe. 2010. http://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_AU_v3.0.1.pdf. [Accessed 10 June 2016]
11. Open Geospatial Consortium Inc. OpenGIS web feature service 2.0 interface standard. 2010.