Using INSPIRE data on the web

Submission to the Linking Geospatial Data workshop

Clemens Portele, interactive instruments GmbH

14.01.2014

1 Introduction

This paper builds upon the talk "How to use INSPIRE data" given at the INSPIRE Conference 2013[1,2] and describes experiences from the ongoing work in the European Location Framework (ELF) project[3]. It discusses patterns how the usability of INSPIRE data on the web can be improved and identifies open challenges that we are facing. This paper represents views of the author and does not necessarily represent the view of the ELF consortium.

This paper focuses on two aspects. The first aspect relates to requirements and patterns for reference geography on the web. The second aspect is about bridging the different platforms and technologies. There are other important issues that deserve a discussion, too, e.g. data quality, license conditions, access control, etc., but these are outside of the scope of this paper.

In the talk, if accepted, I plan to discuss the aspects discussed in this paper using live services and data.

2 Reference geographies on the web

A principle of INSPIRE is to support linking data to geospatial reference data in order to

- geo-reference "business data";
- improve data integrity and reliability.

This is discussed in detail in the INSPIRE Generic Conceptual Model[4]. Topographic, address and cadastral datasets are typical examples of sources for geospatial reference data.

Support for distributed systems with links across system boundaries is important so that always the latest, up-to-date content is used. The pattern that an application provider downloads and integrates the data and ships it with the app is often insufficient.

To make this work in practice and on the web, there are a number of prerequisites. Most of these are already addressed in the current plans for ELF, but some will need more work.

---

[1] Paper: http://inspire.jrc.ec.europa.eu/events/conferences/inspire_2013/schedule/submissions/307.doc

[2] Slides: http://inspire.jrc.ec.europa.eu/events/conferences/inspire_2013/pdfs/25-06-2013_ROOM-3_14.00 - 15.30_45-Clemens Portele_Clemens-Portele.pdf

[3] The European Location Framework (ELF) is a technical infrastructure which delivers up-to-date geospatial reference data from public administrations in Europe, in particular mapping and cadastral agencies. To kick-start the European Location Framework, a three-year project is co-funded under the ICT PSP programme of the European Commission (Project reference 325140). The project runs from March 2013 to February 2016. More information about the project is available at http://www.elfproject.eu/.

[4] http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/D2.5_v3.4rc3.pdf, see Chapters 13 and Annex D

- The infrastructure needs to support access via http to individual features as well as support queries on datasets based on spatial, temporal or thematic characteristics.
    - In INSPIRE, this is not a requirement. It is only a requirement to provide access on the dataset level – and not necessarily via http (ftp, for example, would be allowed, too).
    - In ELF it is planned to provide access to individual features as well as to support queries via web services (using OGC Web Feature Services and GeoServices REST API Feature Services).
- Persistent identifiers are assigned to all features that are part of datasets that constitute the reference geography, e.g. cadastral parcels, transport network elements, administrative units, statistical units, etc.
    - This is a requirement for some of the INSPIRE feature types, but not for all of them.
    - In ELF it is planned to assign persistent identifiers to all features.
- On the web, persistent identifiers should be http URIs.
    - This is not a requirement in INSPIRE, but a recommendation. See also Annex H of the Generic Conceptual Model.
    - In ELF all features will be accessible via http URIs, but these are document URIs using OGC Web Feature Service queries or GeoServices REST API URIs, and not necessarily persistent. As these are document URIs and the same feature is available via different services, the same resource has multiple URIs, there is not yet a single persistent URI for every feature.
    - If ELF would assign persistent URIs for its features, it would have to support redirecting to appropriate document URIs for the requested media type.
- Quite often linking will also occur via thematic identifiers (identifiers of the real-world phenomenon, not the feature); for example, road names, cadastral parcel codes, NUTS/LAU codes, TMC codes, or postal codes.
    - In INSPIRE, if included in the scope, these are represented as properties of the feature, but there is no guarantee that datasets provide the information.
    - In ELF the existence of thematic identifiers will depend on the existence of this information in the source datasets, but the importance of such identifiers is well understood. To qualify as reference geography, it will be important in the future to add additional thematic identifiers based on identified needs.
    - ELF will also provide the GeoLocator, a gazetteer and geocoding service supporting geographical names, addresses and administrative units.
- In order to be attractive to application developers and users it is essential that tools and APIs are available that make it really simple to use the data.
    - This is not part of the legal framework of INSPIRE. However, several projects including Are3na, ELF and many others are working on providing tools. At the moment most tools focus on supporting the legal obligations related to providing geospatial data, less on using the data.
    - In ELF the Finnish mapping agency is enhancing its OSKARI platform to support integrating ELF/INSPIRE data from distributed sources and provide access via a showcase mapping application.
    - In addition, ELF/INSPIRE data is made available via ArcGIS Online (more on this below in the next section). The rich sets of web APIs for the development of map-based applications for all major devices simplify the use of ELF/INSPIRE data.
- In practice, links will often be implicit, for example, when LAU codes are used as attribute values. The infrastructure needs to provide a mechanism, typically via APIs, that makes the relation explicit and establishes the join between the business data and the geospatial reference data. The infrastructure also needs to provide a mechanism to show features in maps that get their geo-reference from reference features.
    - INSPIRE does not provide any support for this at the moment. In principle, spatial data services could provide such a capability.
    - For ELF reference data published in ArcGIS Online, automated geo-processing is typically used to geo-reference or otherwise enrich data. In order to provide good performance when using such data in maps etc., the enriched data is stored as a new dataset. Using ArcGIS Online, this can be

achieved on the web without programming or local GIS software. At the moment there are restrictions regarding the data that can be used as input for such geo-processing tasks in ArcGIS Online without programming.
  - Alternatively, the Esri Maps for Office extension could be used to create dynamic maps of data in Excel using the reference geography provided by ArcGIS Online. Other platforms provide related capabilities, too, for example Fusion Tables from Google.
  - OGC has specified the Table Joining Service standard that provides the capability to establish joins between datasets, but the uptake of this standard is minimal[5]. Nevertheless, the Table Join Service may be useful for this and in ELF an application plans to make use of the specification to link health statistics data to reference geography provided by ELF.
  - The OGC Web Feature Service supports joins, too, but requires that all data is available in the same service, which typically is not the case when linking business data to reference geography.

To summarize:

- Implementing INSPIRE just by the letter of the law will not result in geospatial data that is ready for use on the web.
- Easy-to-use solutions become increasingly available through capabilities of commercial platforms that support web-based mapping. Nevertheless, this is not yet fully sufficient for specifying reference geographies and using them to geo-reference business data as it is envisioned by INSPIRE.

3 Bridging between platforms and technologies

It is a reality that there are and will continue to be multiple platforms where geospatial data is used and value is created through applications. These include Spatial Data Infrastructures (based on OGC Web Services), ArcGIS, Google Maps, MapBox, Nokia Here, the Linked Data Cloud, and others.

To some extent the platforms support the same specifications, so there is some permeability between the platforms. However, in most cases this is limited to data that has been processed for immediate presentation to the user, e.g. map tiles or KML documents.

Developers that use a particular platform for their application of course depend on the availability of the geospatial data that they need in their platform of choice. It thus seems important to make INSPIRE feature and map data available in platforms with a large potential for data reuse in applications – using the "native" representations and APIs of the particular platform.

The ELF project therefore explores mechanisms how cached copies of ELF/INSPIRE datasets can be made available – and kept up-to-date – in other platforms. The project uses ArcGIS as a sample platform for this purpose, as ArcGIS is a major platform for developing map- and location-based applications.

This involves several actions (ignoring other relevant issues, like licensing of the data for now):

- Transform the ELF/INSPIRE data: ArcGIS uses a simpler feature model than the ISO 19100 series and INSPIRE. To meet the expectations of application developers on the ArcGIS platform, the data needs to be simplified. For most application developers this simplification is a real benefit as the resulting data is much easier to use than complete ELF/INSPIRE data. Nevertheless, it is essential in this process to still include all important information as feature attributes – including identifiers and other information relevant for symbology and known use cases.
- Develop symbology: ELF will develop portrayal rules encoded using the SLD and SE standards of OGC. The portrayal rules are expected to support the range from small to large scales. The rules and symbolizers will

[5] No product is currently registered to implement the standard, see http://www.opengeospatial.org/resource/products, and it does not seem to play any role in current presentations about OGC.

have to be processed to convert them to GeoServices renderers based on the simplified data model. It is also an option to develop specific portrayal rules for maps for use in ArcGIS or to skip rendering maps in the ArcGIS platform altogether and access map tiles from the OGC WMTS provided by ELF.

- Support for geodetic coordinates and the Spherical Mercator projection: By now, the standard coordinate reference system for map displays in most platforms is the Spherical Mercator projection (EPSG code 3857). Feature data will often be requested in geodetic WGS84 coordinates (which can be considered identical with ETRS89 in most scales). The coordinate reference systems should thus be supported. Map tiles from an external WMTS can only really be used (see above), if the WMTS uses the standard tiling scheme and the Spherical Mercator projection.
- Publish in ArcGIS Online: The feature data and map tiles are published in ArcGIS Online as hosted services supporting the GeoServices REST API. Using the elasticity and scalability that comes with cloud computing these services will accommodate very demanding applications and a large number of concurrent users. The feature services also support geometry generalization depending on the current map scale as well as http caching to minimize network traffic and make the use of feature data responsive enough for their use in applications.
- Update data: INSPIRE provides no mandatory mechanism how updates to datasets are exchanged. One of the tasks in the ELF project is to explore with the data providers whether information about changes can be provided continuously or periodically, so that the feature data and map tiles in ArcGIS Online can be kept up-to-date. The use of Atom feeds or a geosynchronization service are options that will be discussed.

As a result, if successful, the reference geography established by ELF data will be available to all users and application developers of the ArcGIS platform.

Independent of the ELF project, interactive instruments is also exploring other mechanisms to provide bridges between platforms. The upcoming XtraProxy for WFS product[6] accesses OGC Web Feature Services and publishes them as GeoServices REST API Feature Services without a noticeable performance overhead. This allows clients in the ArcGIS platform to access OGC Web Feature Services directly, too, but usability depends on the performance of the OGC Web Feature Service.

Similar mechanisms like in the ArcGIS case would apply to other platforms, too. Of course, each platform differs, so somewhat different measures will be required to establish reference geography from ELF data depending on the platform.

[6] See http://www.interactive-instruments.de/en/xtraproxy/.