Managing hydrographic data for multiple usage

Mark TERLIEN

Key words: Maritime Spatial Data Management, GeolinQ, Point clouds, WMS, WFS, REST

SUMMARY

Managing hydrographic data varying from point clouds to object data like buoys, contours and wrecks is getting more challenging, because of increasing volume of acquired data and multiple usage. The traditional use for hydrographic data for navigational purposes rapidly extending over the years to new fast growing usages like offshore windfarms, dredging and environmental issues. A new innovative approach of hydrographic data management is crucial to cope with these challenges.

The presentation starts with an assessment of the functionality needed for data management. Data management solutions are needed that offer a combination of fast and efficient storage, visualization and web distribution and flexibility with respect to point, raster and feature attributes, metadata and styling. Users need to be able to browse datasets based on metadata and location, visualize point cloud data according to their own needs and publish point cloud datasets based on customer requirements.

Second part of the presentation shows how thes new approach is implemented in GeolinQ. The data management approach based on the concept of flexible data modelling will be explained. Examples will demonstrate how the approach improves data quality by structuring and validation, allows fast data retrieval, customized visualization and multiple options to share or export hydrographic data. Experiences will be shared defeating the challenges managing hydrographic data for multiple usages.

1. INTRODUCTION

Managing hydrographic data varying from point clouds to object data like buoys, contours and wrecks is getting more challenging, because of increasing volume of acquired data and multiple usage. The traditional use of hydrographic data for navigational purposes is rapidly extending over the years to new fast growing usages like offshore windfarms, dredging and environmental issues.

These new usages require different hydrographic products where the primary objective is no longer safe navigation. No longer the shallowest depth is required for all purposes. This asks for a different way of managing and distributing hydrographic data. Different hydrographic products should be offered for different purposes. As the number of potential customers and the their requirements for hydrographic products are increasing all the time more and more hydrographic offices and research institutes decide to offer not only standard products such as charts for safety of navigation but also their data sources either as open data or as closed data. Open data is freely available to everyone to use and republish as they wish, without restrictions from copyright, patents or other mechanisms of control while closed data is made available for a charge.

2. NEW CHALLENGES

In recent years a number of distribution portals have been released to gives access to hydrographic products and their data sources for Europe (EMODNET), the Baltic sea (Baltic Sea Bathymetry Database), USA (NOAA). These portals offer a limited set of hydrographic products in different formats for download. If data products are needed with different specifications (e.g. different resolution, different contour intervals, different year, different color scheme), the portal gives access to the underlying datasets. These can be downloaded and further processed into tailor-made products. The next step is that these portals are also going to offer functionality to generate hydrographic products on-the-fly based on requirements entered by the user.

This new way of distribution hydrographic data and being able to offer functionality to generate hydrographic products on-the-fly requires a new innovative approach for maritime spatial data management. GeolinQ is a web-based spatial data management solution that can implement the backbone for such a new maritime spatial data infrastructure.

GeolinQ is a web-based solution to link and publish spatial data in an efficient and flexible way (see Figure 1). GeolinQ covers the complete value chain from import to the distribution of data. Data from various sources is imported and stored according to user-defined metadata models and data models and is immediately available for visualization and publication as web service.

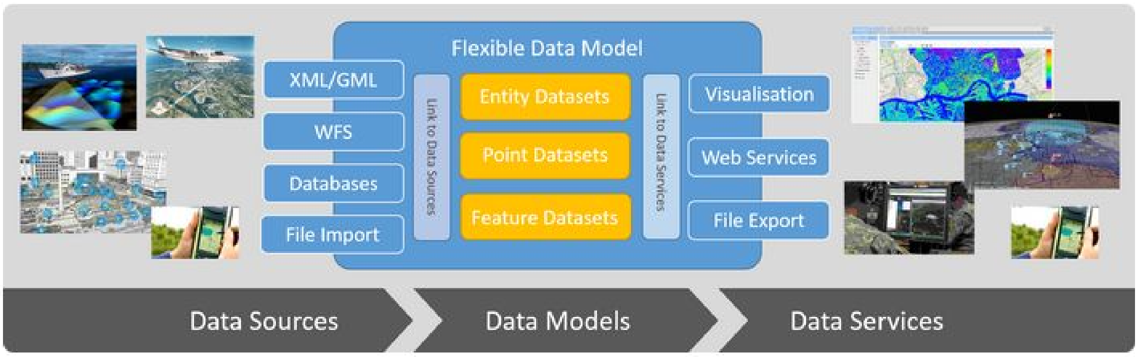

Figure 1: Business architecture GeolinQ

3. MARITIME SPATIAL DATA MANAGEMENT

Maritime spatial data management comprises the total hydrographic value chain from data collection through data processing, storage, product generation to distribution. During this whole process spatial data management must guarantee that data is discoverable, accessible and usable by a variety of users and applications.
Metadata is an important component of data management. Metadata is needed to discover and access data and to evaluate the usability of dataset. To evaluate usability aspects such as quality, resolution, post-processing steps and moment of collection must be known. Also traceability and liability should be an integral part of spatial data management.

Although there is a strong tendency towards standardization of metadata according to ISO19115 (e.g. S-102), many organizations still use their own metadata definitions. Therefore in GeolinQ users can define their own metadata schemas according to their own business needs. These schemas are created and modified on-the-fly. There is no need for new software releases or tailor-made add-ons. Metadata can be exported to an XML file. This XML file comes with a XSD and can be used for metadata distribution or further processing in other tooling. Using XML technologies such as XSLT or Xpath this XML can be translated into a standardized ISO19139 XML.

Not only the metadata model but also the data model to store the hydrographic datasets must be customizable by the user. Survey data is no longer delivered and distributed as only x,y,z. Current acquisition equipment is capable of collection a large number of addition point attributes that may interest data users as well. Therefore these attributes must be kept during the whole workflow and that must also be available for distribution. Data users are interested in more than only the depth attribute.

Since S-100 allows plug and play updating of data models without new versions of product, product feature catalogs must be more flexible and capable of expansion. In GeolinQ the data model to store feature catalogs can be created manually or can be generated from XSD files at runtime allowing a great flexibility is storage and usage of feature data.

GeolinQ stores all data as well as the metadata in a relational database management system (RDBMS). One of the main advantages of an RDBMS is the possibility to query and search all data (administrative and spatial) and metadata using a standard query language (SQL). Furthermore a RDBMS supports transaction management and improves scalability, manageability, backup and restore and possibilities for performance tuning. RDBMS indexing functionality on administrative and spatial data guarantees fast access with growing data volumes.

4. MANAGING POINT CLOUDS

Survey data is usually delivered as point clouds (ASCII, LAS) or grids (e.g. GeoTIFF, BAG). One survey can contain up to one billion points making data processing, storing, visualization and distribution as web service a challenging task. GeolinQ is shipped with an algorithm for fast storing, visualization and retrieval of point clouds and grids (see Figure 2). The algorithm has no limitations on number of points, point cloud size, physical memory of hard disk size and does not require any prior knowledge about number of points or MBR of the point clouds. Also appending new files to already stored point cloud is allowed. The entire point cloud is stored in the database using data chunks; No proprietary data types are used (database independent) which guarantees an open solution. During import a visualization pyramid and a spatial index are generated for fast retrieval and visualization on all scale levels. At the most detailed level, the individual points are shown. The footprint (hull) of point cloud is automatically generated based on a TIN.

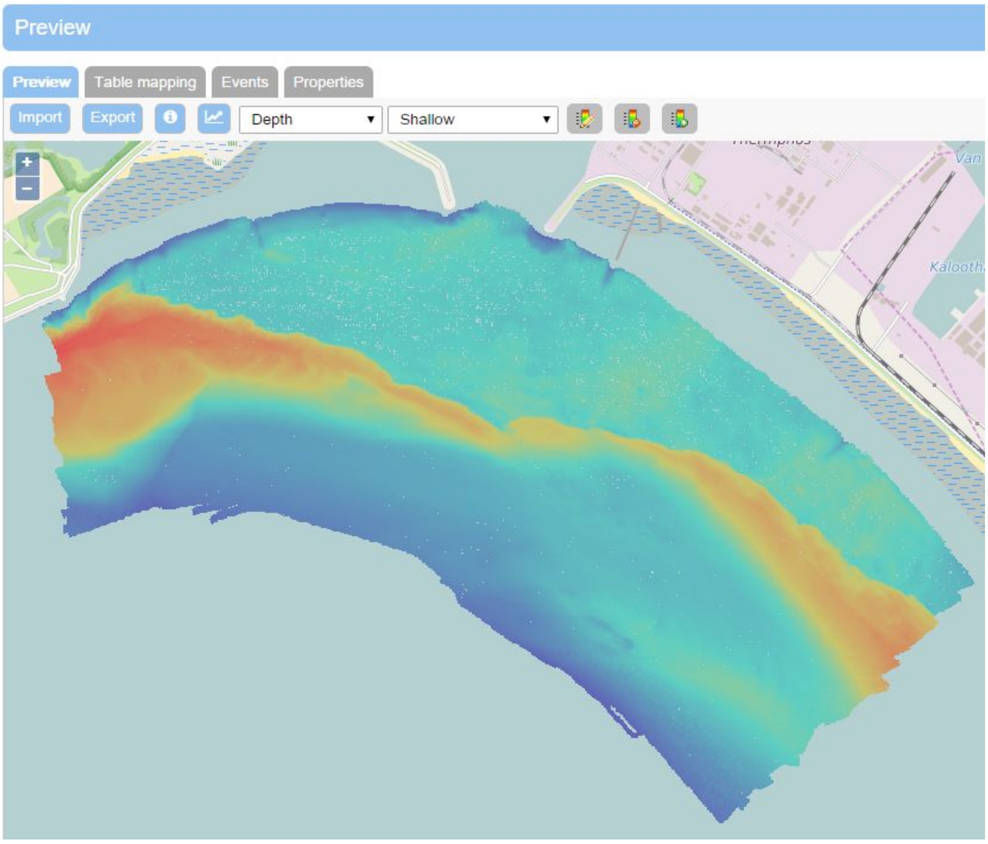

Figure 2: Example of point cloud visualization

To validate data during import, domains can be defined for each attribute. A domain defines the range of attribute values allowed. Points with values outside the specified domain will be skipped. All coordinate systems from the EPSG database are supported as input coordinate system. In the database, the points are stored in WGS84 or ETRS89.

5. MANAGING FEATURE DATA

Feature data can be delivered as files (Shape, XML/GML) or can be downloaded from web services (e.g. WFS services). To store features in GeolinQ first a data model must be generated. The data model is either generated directly from the file (e.g. Shape) or is generated from the XSD (e.g. XML/GML). In the case of a WFS the feature definition is requested with a DescribeFeature request to generate the data model. All coordinate systems from the EPSG database are supported as input coordinate system. In the database, the features are stored in WGS84 or ETRS89.

6. DATA VISUALIZATION

Data stored in GeolinQ is accessible in various ways. GeolinQ offers a web portal that provides internet access to the data and metadata. The metadata of all system-defined and user-defined metadata attributes can be queried to search for datasets. Metadata queries can be extended with a spatial constraint to find only datasets within a specific area. Footprints of all datasets or of single datasets are plotted on a chart for data validation, to check if the correct coordinate transformation is applied, the correct vertical datum is defined and if attributes have the same units.

As GeolinQ gives direct access to the data and not to a pre-defined product, users can interact with the data and change the data visualization according to their business needs. Color scales for point clouds, styling for features can be created, modified and applied interactively (see Figure 3). Styled layers can be published a WMS.

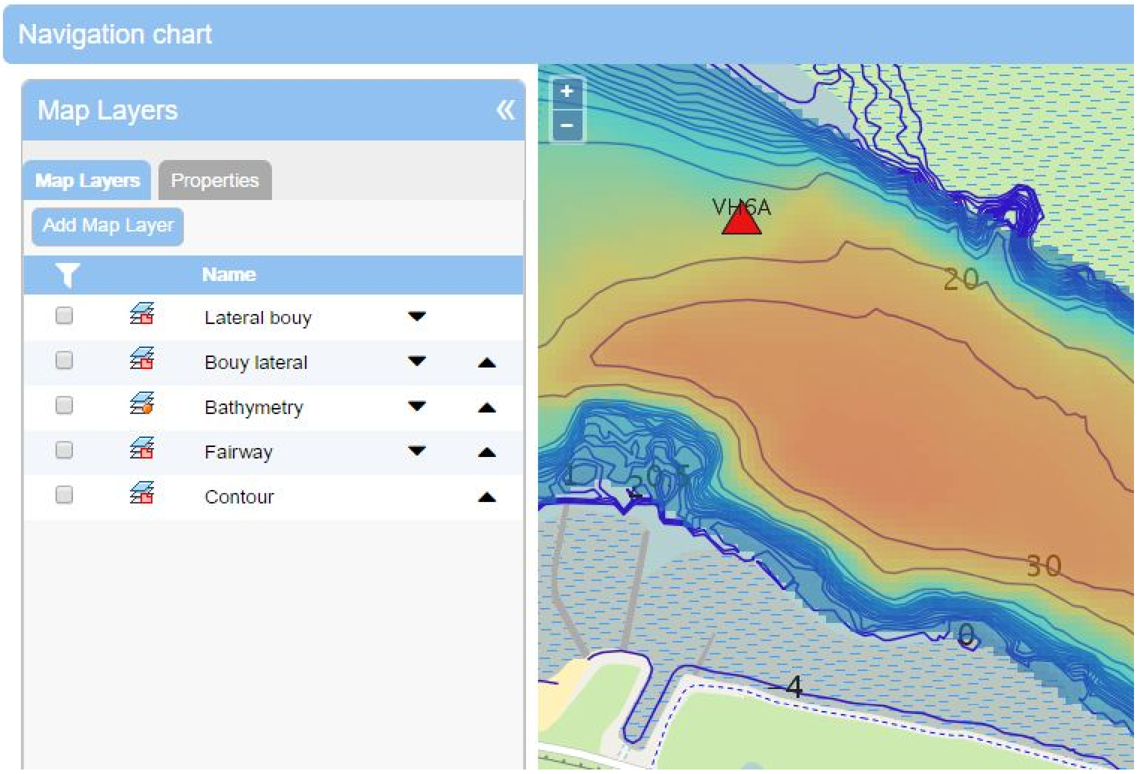

Figure 3: Example of feature visualization

7. DATA DISTRIBUTION

Data and metadata access and distribution are the cornerstones of a spatial data infrastructure (SDI). A SDI provides a basis for spatial data discovery, evaluation, and application for users and providers within all levels of government, the commercial sector, the non-profit sector, academia and by citizens in general.
A Marine Spatial Data Infrastructure (MSDI) encompasses marine and coastal geographic and business information in its widest sense. An MSDI would typically include information on seabed bathymetry (elevation), geology, infrastructure (e.g. wrecks, offshore installations, pipelines, cables); administrative and legal boundaries, areas of conservation and marine habitats and oceanography.

INSPIRE has a defined a number of core marine and coastal themes and a set of themes of regular interest to different stakeholders working in the coastal/marine environments. INSPIRE also defines the services that have to be deployed for data discovery, data viewing and downloading. The INSPIRE services are based on the OGC standards for data discovery (CSW), data viewing (WMS) and data downloading (WFS).

Point clouds stored in GeolinQ can be published as WMS layers. The point clouds that belong to a WMS layer are dynamically selected by a query on the point cloud metadata attributes. Using a query to select the point clouds has the advantage that new point clouds are automatically added to the WMS layer when they meet the search criteria of the metadata query. Therefore a WMS layer may consist of one single point cloud or of a collection of point clouds. The color scales that are created for visualization in the web portal can also be used as styles of a point cloud WMS layer.

An export to download the points in a point cloud to a file with the possibility to select the area, attributes and coordinate system to export depending on the business usage is also supported. Together with the point cloud the point cloud metadata can be exported to an XML file. The structure of this XML file is defined by the metadata model that is created. A XSD is generated as well for validation purposes.

Features can be published as WMS layers. The styles that are created in the web portal for visualization and thematic mapping of features can be used as styles of features of a feature WMS layer. A WMS returns a rendered image but does not give access to the individual features. To download the individual features, GeolinQ offers a WFS service. The feature definition of the features served by the WFS are dictated by the user-defined or generated feature definitions. Features can also be accessed using the GeolinQ REST interface.
For features also file export is available. Features can be exported to Shape, XML or GML file. The related metadata can be exported to a XML file. For XML/GML files with features as well as for the metadata XML a XSD is generated as well for validation purposes.

No only data and metadata stored in GeolinQ can be distributed and shared. Also the data and metadata definitions, including color scales and feature styling can be shared between different GeolinQ instances. GeolinQ has a number of pre-defined data and metadata definitions, for example for ISO19115.

8. REST INTERFACE

REST stands for Representational State Transfer. It relies on a stateless, client-server, cacheable communications HTTP protocol. REST is a lightweight alternative to mechanisms like RPC (Remote Procedure Calls) and Web Services (SOAP, WSDL, WFS). REST is becoming a popular building style for cloud-based APIs also in the geospatial world.

Data and metadata stored in GeolinQ can be queried using a REST interface. The response is returned as XML or as (Geo)JSON. The structure of the response is dictated by the user-defined or generated data models in GeolinQ. For every new data model and metadata model that is created or generated, automatically a REST interface is available without any need for programming or configuration.

9. CONCLUSION

Because the traditional use of hydrographic data for navigational purposes is rapidly extending over the years to new fast growing usages like offshore windfarms, dredging and environmental issues, clients of hydrographic offices and research institutes are not only interested in the pre-defined hydrographic products with limited usage (e.g. ENC's) but also in the hydrographic data that was used to create these products. Therefore hydrographic offices need a new architecture to manage and distribute their hydrographic data.

Hydrographic data has to be made available over the internet. This requires accurate and up-to-date metadata for data discovery, flexibility in modelling of data and metadata, new technologies to access point clouds directly from the internet and their not derived products such are resampled grids or contours and a wide variety of services (WMS, WFS, REST) to access and download point clouds and feature data and their metadata.

GeolinQ is a platform that provides the functionality to implement such a new architecture to manage and distribute hydrographic data for multiple usage.

REFERENCES

Links:
http://www.geolinq.com, last accessed on September 12, 2016

CONTACTS

Mark Terlien
IntellinQ
Vasteland 78
Rotterdam
Netherlands
Tel. + 31 6 24624844
Email: m.terlien@intellinq.nl
Web site: www.intellinq.com