Pitney Bowes
Business Insight

Meeting the Obligations of the EU INSPIRE Directive

White Paper #2

Metadata

The Key to Realising the Potential of Your Data

Contents

Meeting the Obligations of INSPIRE

Metadata – the key to realising the potential of your data

A guide to how Pitney Bowes Business Insight (PBBI) can assist with implementing the INSPIRE directives.

This white paper is one of a series that outlines how you can deliver an infrastructure to meet key requirements for the INSPIRE agenda. Pitney Bowes Business Insight (PBBI) is focused on building innovative, and easy to use solutions that can help you meet these obligations. This paper focuses on the obligations, processes and solutions for implementing a metadata catalogue that meets the requirements of INSPIRE. It also addresses the broader operational, communication and management processes that will be key to implementing a sustainable, scalable and robust business level integration. INSPIRE is not only about the technical architecture, it is has to deliver changes in every day procedure, policy and practise across relevant organisations and partnerships.

Introduction

> *All Government organisations face continual challenges with discovering data, and then enabling efficient and appropriate access to that data. The high level goal of INSPIRE is enabling member states of Europe to make better informed decisions, more insightful planning and more efficient resource deployments by a shared and collaborative approach to intelligence and key decision making. This goal will lead to improved operational excellence and lower overall costs.*

This paper will outline

1. The key obligations for maintaining a metadata catalogue and aspects to consider, in order to meet the deadlines for compliance.
2. How PBBI can deliver the requirements of a metadata catalogue and help you in your tasks of implementation. In addition this paper will briefly consider the wider benefits of these activities aside from simply complying with the INSPIRE mandate.
3. Provide a resource for future discussion, extension and update on how PBBI is helping organisations across Europe to deliver on the INSPIRE vision.

Audience and Assumptions

This paper assumes some familiarity with the requirements of INSPIRE, the directives and understanding of the key goals. The following roles can benefit from these topics:

- Data Custodians
- IT Managers
- GIS Users and Operators

Metadata

Simply put, metadata is data about data. It is descriptive information about a particular data set, object, or resource, including how it is formatted, and when and by whom it was collected. Metadata may be created automatically using software or entered by hand.

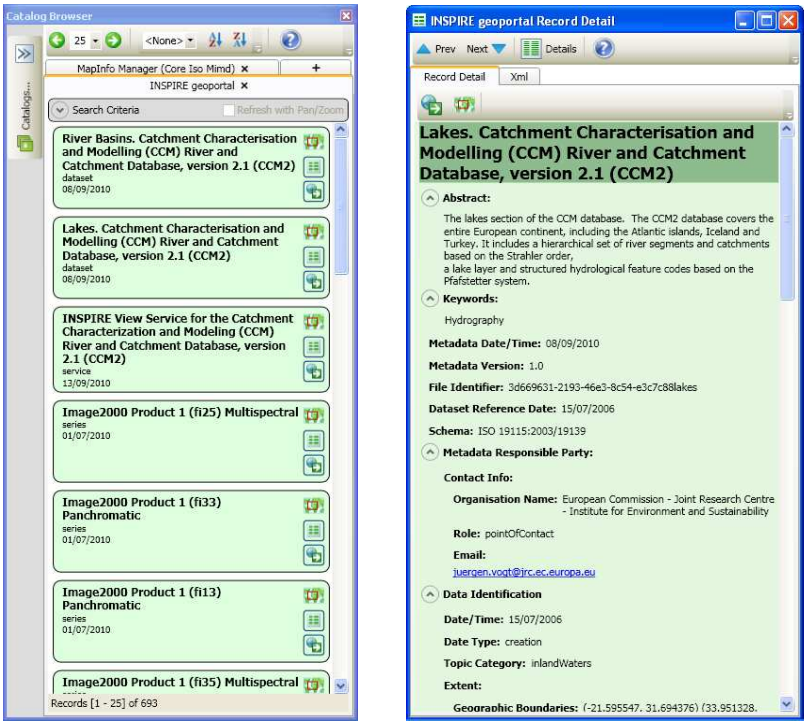
> *The publishing of metadata aids potential users of data in finding the data they need and understanding appropriate uses for it. Metadata also helps with understanding the ownership and authority of data. In addition a metadata repository helps reduce duplication of effort in collecting data and for organisations that purchase data can help them to avoid duplicate purchases taking place in different parts of the organisation.*

Public authorities in Europe and organisations who produce data for public authorities are required and legally mandated to produce INSPIRE compliant metadata. Metadata is a significant aspect of the INSPIRE Directive, and as such has it has its own set of supplementary directives.

Briefly, INSPIRE requires metadata to be collected as follows [3]

- Data that relates to an area where a European Union (EU) Member State has or exercises jurisdictional rights and the data is in electronic format
- Data that is held by or on behalf of a public authority, having been produced or received by a public authority or being managed or updated by that authority and falling within the scope of its tasks
- Data that relates to one or more of the themes listed in Annex I, II or III of the INSPIRE directive.

All metadata discovery ("discovery" refers to the ability to search metadata) will be free and in the public domain. This implies it should be easy to use, accessible and intuitive to access for all users. It should also be easy for those most closely involved with metadata within their organisation to deploy the necessary public metadata catalogs in order to meet the INSPIRE requirements.



An example of metadata records from the INSPIRE geoportal displayed in MapInfo Professional's Catalog Browser.

Federation

Don't assume that all content or metadata has to be put into one place, into one system, one database or a single Spatial Data Infrastructure (SDI). The use of standardised services is specifically designed to aid federation across several connected systems, such that the data remains close to, and in the custody of the originating organisation, while being discoverable and usable by anybody who has been granted access to it. Therefore data must be easy to discover across a distributed domain.

> *Consider the library metaphor, each library maintains a collection of books, and a search at any library catalogue will return books in other libraries. If your request returns a book in your local library, you can find it via its index. If the book is in a remote library, there are typically services to request that the book be delivered to your local library.*

An interconnected network of metadata discovery services, permits data searching (and with the appropriate additional capabilities in place, the viewing and downloading of data) from multiple catalogs at once. This concept is sometimes referred to as cascading.

Profiles

Metadata standards are a complex area and INSPIRE outlines key metadata fields, and defines mandatory elements that need to exist in your own metadata catalogue. This document, Technical Guidelines based on EN ISO 19115 and EN ISO 19119, regulates the form, structure and contents of the metadata catalogue. (ISO refers to the International Standards Organisation.) Different countries implement different metadata schemas (also known as profiles). For example in the UK, the applicable standard is GEMINI. Gemini is an ISO 19115 implementation and is compatible with INSPIRE. Similar standards for Germany and France are known as the GDE-Profile and ISO 19115 Fr respectively. Metadata standards and profiles rely heavily on techniques for storing, transforming and transferring metadata information as XML, such that it can be consumed by the widest number of systems. Although ultimately transparent to the end user, XML schemas are used to validate and form the documents, while style sheets are used to transform the document from one schema into another. Managing the different schemas and style sheets in addition to the core data will be the discipline of the metadata custodian.

Organisations should continue to capture and maintain their metadata in their current format, with the understanding that they may need to revisit the metadata they have collected in order to ensure it is compliant with the INSPIRE profile. Some metadata may need to be changed or some fields added. The profile will include both mandatory and optional metadata fields. Clearly the minimum requirement is to capture the mandatory metadata, but further business value is derived from establishing complete metadata records. In many cases metadata may be captured informally or by using tools such as Microsoft Office, spreadsheets or bespoke databases. If that is the case, you are again recommended to continue to maintain good discipline and be aware that metadata import/export tools are available to help you Extract, Transform and Load (ETL) metadata from these source systems into an INSPIRE standards compliant catalogue.

PBBI offers a range of toolkits for auditing, building and maintaining a metadata catalogue compliant with the appropriate ISO and Open Geospatial Consortium (OGC) standards, Pitney Bowes MapInfo® Manager™ from PBBI includes a range of modules for establishing a metadata framework for publishing, managing and integrating metadata throughout your organisation and integrating with the broader service oriented architecture.

Benefits of high quality metadata

Metadata is critical to a business. It is recommended that even if the INSPIRE directives do not require you to maintain metadata for some of your datasets, you should consider broadening your metadata capture to include as much of your organisation's data as possible. This could include spatial, non spatial and document based content. Value is delivered in the following ways, and this value increases with the volume, currency and quality of your metadata.

Metadata helps to:

- Clearly define ownership (including copyrights) and responsibility for data. Specifying copyright information in your metadata informs others of the copyright terms that are assigned to your data.
- Increase the usage of your data assets. A metadata catalog assists all who might need to use data to be able to learn of its existence as well as to access and use the data.
- Discover publicly accessible data that can be used to enrich or enhance existing data, or be incorporated into your business processes.
- Identify data redundancy and opportunities to reduce maintenance costs, including avoiding errors such as purchasing the same commercial dataset more than once in an organisation.
- Ensure quality assurance on data and assets, the metadata helps the user to determine the data's appropriate uses and context.
- Improve the understanding of overlaps in business areas. Metadata is instrumental in helping users to understand complex datasets which may form interrelationships amongst different departments or users within an organisation.
- Improve the accuracy of end-user analysis of access and usage of the data. It becomes easier to develop business cases for valuable dataset maintenance.
- Enable managers to make quicker and more insightful decisions in data use.
- Share information much more easily with partners, agencies and the public. This promotes better decision making, shared services and increases citizen satisfaction.

You should consider deploying an internal use metadata catalogue for the purposes of improving the overall finding and using of information within your organisation. It is worth noting that for some organisations, the bigger benefit for having to go through the exercise to comply with the INSPIRE directive may come from the improved understanding and optimised use of all of their datasets internally.

In considering a metadata strategy, it may be worth the incremental effort to go beyond the INSPIRE metadata requirements in order to collect additional information which may be of specific benefit to your organisation. This can include extending your metadata collection efforts to all of your datasets (not just those that are specified by the INSPIRE directive) and to perhaps collect some additional useful metadata. For example, information about data purchase prices, brands and providers (for commercially acquired datasets), digital signatures, licensing terms, change history/revision dates and more can prove valuable.

ROI / Benefits of Metadata

The value of metadata potentially outweighs the initial upfront investment for capture of the metadata, and the ongoing maintenance costs. Data is becoming an organisation's most critical asset; in many cases the value of data can easily exceed several million Pounds or Euros in ongoing value. Data is also at the heart of most business critical systems. It is essential that the data is utilised in the most efficient manner by all who need it, so that its benefits can be maximised, and gains realised both initially and continuously.

PBBI can help you develop the business case for metadata, and execute on implementation quickly and efficiently.

Customer Requirements

PBBI is dedicated to providing a full suite of integrated capabilities to help our customers not only meet their INSPIRE requirements but also to get the best possible return on their effort and investment. These include:

Data auditing – it is necessary to know, in the first place, what datasets and data assets the organisation has.

Automated metadata capture

- "Harvesting" is the term used for the bulk collection of metadata. Many organisations will find that they will need to engage in a large initial data collection effort. Some metadata items, for example the bounding coordinates of a dataset, the date the data was created, the date of the last revision of the data, the name of the organisation, the location of the data, a contact e-mail address and more might be collected in automated ways.

  In addition to harvesting, metadata template structures are required in order to easily apply repeated values of metadata across a number of data files. For example, the name of the organisation or the contact person may be the same for some or all of the data sets.

- Metadata import and export. An organisation that has engaged in some form of collection of metadata may wish to reuse this in any new metadata catalogs that they deploy. As such it is important to be able to import existing sets of metadata and in some cases map the fields from the original metadata set to different fields in the store of metadata being built. Also, the ability to export metadata makes it easier to share with other systems.

Manual and ad-hoc metadata capture and editing

- Individual users who create a new dataset or make changes to an existing dataset will have the requirement to update the metadata catalogue. This requirement may stem from the GIS Users in the organisation (such as those working with Pitney Bowes® MapInfo Professional®) but it can also extend in some cases to non-expert users of web based systems. As such an organisation may find that a variety of their client software requires these capabilities.

Deploy the metadata catalogs internally and externally

- INSPIRE requires the deployment of metadata catalogs using an open standard defined by the OGC called Catalog Services for the Web (CSW). It is necessary for an organisation to deploy a publicly accessible catalogue built to the standard. It is understood that many organisations will not be able to hire specialist personnel to do this and as a result the capability to do this must be a simple and easy configuration task.

  The deadline for the collection of the metadata detailed in Annexes I and II is in December 2010. Looking beyond this there are additional requirements for making the data available to users. These will include deploying open standard services such as Web Map Services (WMS), Web Feature Services (WFS), direct provision of "flat" or compressed archives of data files and by other means. We will be releasing another white paper on this topic shortly.

Integration with existing systems

- Some organisations may benefit from programmatic capabilities in order to be able to integrate metadata processes and workflows into their existing systems.

What about the quality of the data?

- INSPIRE requires that metadata (and the data itself) be collected and shared regardless of the quality of the data. This has implications for the reputation of the organisation that produced the data or for whoever is responsible for sharing it. As such, some organisations may find it necessary to engage in a "clean up" and enrichment effort up front, as well as on an on-going basis.

The Pitney Bowes Business Insight Portfolio

PBBI provides an integrated set of capabilities across our MapInfo product line to help our customers meet their INSPIRE requirements.

The MapInfo Enterprise Location Intelligence Suite

- MapInfo Professional – PBBI's flagship Desktop Mapping System offers comprehensive Location Intelligence and GIS capabilities and is renowned for its ease of use. MapInfo Professional includes the ability to search CSW catalogs for data as well as to create and edit metadata. MapInfo Professional also supports directly accessing the data from the CSW server, when the data is made available. Data can be downloaded as flat files, accessed through open standard services such as a Web Map Service, or Web Feature Service and through Tile Servers.

- MapInfo Manager – PBBI's spatial data management solution designed to allow organisations to harvest and catalogue their metadata and to make it available through Web Map Services, Web Feature Services and directly as files. MapInfo Manager is specifically designed to be easy to use such that the skill sets required are available within the GIS department or IT department of the typical organisation. In addition, MapInfo Manager includes API level access to spatial functions for those who may wish to integrate MapInfo Manager with internal systems is desired.

- MapInfo® Stratus™ – PBBI's latest web mapping offering. MapInfo Stratus should be considered by organisations that wish to deploy capabilities to meet their INSPIRE requirements without the need for installing and maintaining software on their own IT resources. MapInfo Stratus offers a rich Web 2.0 interface and is deployed and customised through an intuitive configuration interface. No programming is required to deploy and start meeting INSPIRE obligations. In addition, a set of Rich Internet Application (RIA) controls are also available for quickly and easily embedding Stratus capabilities anywhere both on an organisation's web site or internal applications.

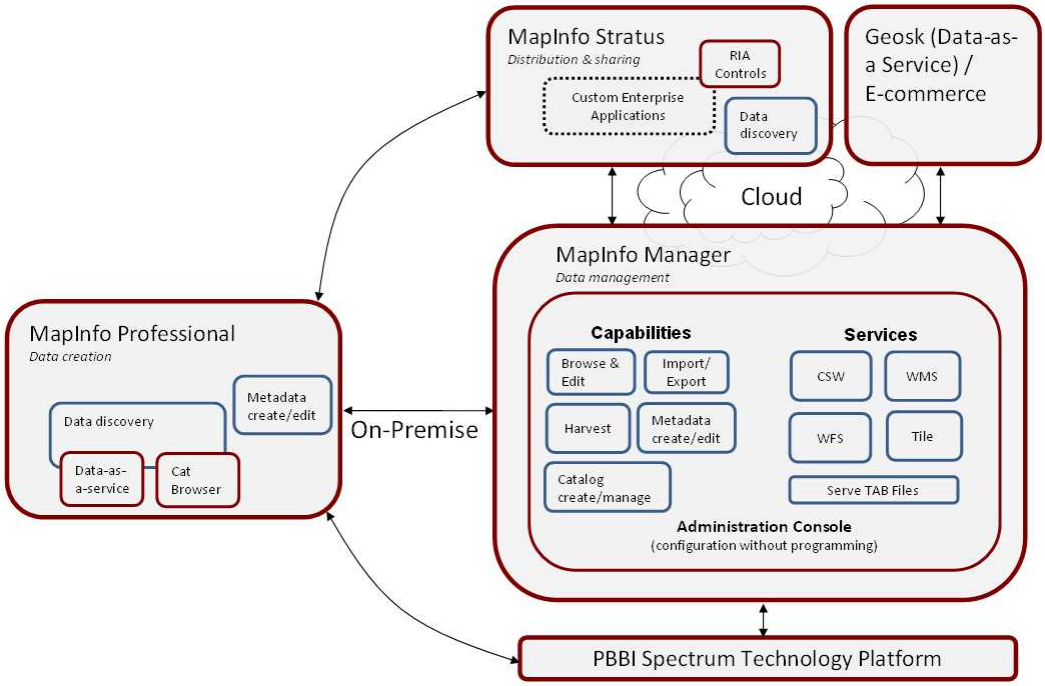
Data Quality and Data Integration

Many organisations are faced with challenges around the overall quality of their data holdings. Data Quality problems can include misspellings, duplicates, transposed fields in tables, inconsistent reference values (primary keys) in different sources and more. Also, what compounds the problem is that data holdings in an organisation may be dispersed in a large number of different formats, network locations and systems.

- The Pitney Bowes Spectrum™ Technology Platform provides data quality and data integration capabilities. These can be used to ensure the data is of sufficient quality and is suitable for its intended purpose and that an organisation's myriad of source data systems can be directly accessed and/or aggregated to provide automated capabilities for keeping your data and metadata up-to-date and to the highest possible standard.

Data Vending and E-commerce

The Pitney Bowes Geosk™ Platform offers data vending and E-commerce capabilities allowing organisations to acquire and manage data content. One aspect of this is the Geosk Marketplace which allows our customers to obtain data in the right format, at the right level of aggregation, of the right quality and within the right timeframe. In the near future, additional capabilities for Data Use and Content Management will be delivered through the Geosk Library, through software integration and a Content Management as a Service offering. With respect to INSPIRE, organisations can consider selling their data for commercial use through Geosk via metadata catalogs meeting INSPIRE requirements. A separate white paper will be dedicated to the topic of INSPIRE and e-Commerce.



The PBBI Product Portfolio

Summary

Compliance with the INSPIRE directives creates additional responsibilities for those organisations directly impacted by the mandate (and for many organisations indirectly impacted, as well). PBBI recommends an approach whereby you strategise around gaining the most benefit from your investment in an effort to meet the INSPIRE mandate.

The benefits of a metadata repository include ensuring that the data required by those in your organisation (and those outside of it) is known and utilised in an appropriate manner. Time and effort are saved through avoiding in appropriate analysis or unnecessary searching around for the required data. Also, time effort and money can be saved as data collection efforts (and data purchases) are not duplicated and decisions around data ownership, maintenance and upkeep are optimised to real world usage and requirements.

The PBBI approach is to consider INSPIRE as a platform to significantly improve your organisations overall performance around a Spatial Data strategy. We offer our customers more than just a “tick the box” solution to meet INSPIRE requirements. Our portfolio includes not only the metadata catalogue and data sharing services required by INSPIRE but also additional capabilities to help you ensure your data is of the highest quality and that background processes in your organisation support not just your INSPIRE commitments but optimise many business operations generally.

References

[1] Directive 2007/2/EC, 25/4/2007

http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:108:0001:0014:EN:PDF

[2] eSDI-Net+, Community Network

http://www.esdinetplus.eu/

[3] Metadata Regulation No 1205/2008, 3/12/2008

http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:326:0012:0030:EN:PDF

[4] UK Ordnance Survey TOID

http://www.ordnancesurvey.co.uk/oswebsite/freefun/geofacts/geo1201.html

[5] UK National Land and Property Gazetteer

http://www.nlpg.org.uk/nlpg/

Resources

INSPIRE Main Web site

http://inspire.jrc.ec.europa.eu/

INSPIRE Conference 2010

http://inspire.jrc.ec.europa.eu/events/conferences/inspire_2010/

Pitney Bowes Business Insight

http://www.pbinsight.com

United Kingdom GIGateway – Metadata Standard Page

http://www.gigateway.org.uk/metadata/standards.html

MapInfo Stratus online Blog

http://stratus.pbbiblogs.com/

About Pitney Bowes Business Insight

Pitney Bowes Business Insight, a division of Pitney Bowes Software Inc., a wholly owned subsidiary of Pitney Bowes Inc., helps organisations to acquire, serve and grow customer relationships. Our leading-edge solutions in the areas of Customer Intelligence, Customer Communications and Customer Care enhance our customers' operational systems and workflows, enabling them to manage their customer relationships more effectively. We offer unique and compelling capabilities, including location intelligence; GIS; predictive analytics; data quality, management and integration; and customer communications management. In combination, our solutions and capabilities deliver customer insights that create competitive advantages. Leading companies, government agencies and systems integrators rely on our global expertise and decades of leadership to improve their operational effectiveness and business results. Visit http://www.pbinsight.co.uk and http://www.pb.com for more information.

EUROPE/UNITED KINGDOM

Minton Place
Victoria Street
Windsor, Berkshire SL4 1EG
main: 44.800.840.0001
fax: 44.1753.621140
pbbi.europe@pb.com
www.pbinsight.co.uk

CANADA

26 Wellington Street East
Suite 500
Toronto, Ontario
M5E 1S2
main: 416.594.5200
fax: 416.594.5201
pbbi.canada.sales@pb.com

UNITED STATES

One Global View
Troy, NY 12180-8399
main: 518.285.6000
1.800.327.8627
fax: 518.285.6070
pbbi.sales@pb.com
www.pbinsight.com

ASIA-PACIFIC/AUSTRALIA

Level 7
1 Elizabeth Plaza
North Sydney
NSW 2060
main: 61.2.9437.6255
fax: 61.2.9439.1773
pbbi.australia@pb.com
www.pbinsight.com.au