Implementation of the Metadata Elements of the Inspire Directive

Fiona Lawlor
Regis University

Follow this and additional works at: https://epublications.regis.edu/theses
Part of the Computer Sciences Commons

Recommended Citation
https://epublications.regis.edu/theses/138

This Thesis - Open Access is brought to you for free and open access by ePublications at Regis University. It has been accepted for inclusion in All Regis University Theses by an authorized administrator of ePublications at Regis University. For more information, please contact epublications@regis.edu.
Regis University
College for Professional Studies Graduate Programs
Final Project/Thesis

Disclaimer

Use of the materials available in the Regis University Thesis Collection ("Collection") is limited and restricted to those users who agree to comply with the following terms of use. Regis University reserves the right to deny access to the Collection to any person who violates these terms of use or who seeks to or does alter, avoid or supersede the functional conditions, restrictions and limitations of the Collection.

The site may be used only for lawful purposes. The user is solely responsible for knowing and adhering to any and all applicable laws, rules, and regulations relating or pertaining to use of the Collection.

All content in this Collection is owned by and subject to the exclusive control of Regis University and the authors of the materials. It is available only for research purposes and may not be used in violation of copyright laws or for unlawful purposes. The materials may not be downloaded in whole or in part without permission of the copyright holder or as otherwise authorized in the "fair use" standards of the U.S. copyright laws and regulations.
Implementation of the Metadata Elements of the INSPIRE Directive

by

Fiona Lawlor
fm_lawlor@yahoo.co.uk


School of Computer and Information Sciences
College for Professional Studies
Regis University
Denver, Colorado

Date 29th August, 2008
Certification of Authorship of Thesis/Practicum Work

Submitted to:

Student’s Name: Fiona Lawlor

Date of Submission: 29th August 2008

Title of Submission: Implementation of the Metadata Elements of the INSPIRE Directive

Certification of Authorship: I hereby certify that I am the author of this document and that any assistance I received in its preparation is fully acknowledged and disclosed in the document. I have also cited all sources from which I obtained data, ideas, or words that are copied directly or paraphrased in the document. Sources are properly credited according to accepted standards for professional publications. I also certify that this paper was prepared by me for the purpose of partial fulfilment of requirements for MScSIS.

Student’s Signature:

_______________________________
Fiona Lawlor
Date 29th August 08
Authorization to Publish Student Work

I, (Student’s Name), the undersigned student, in the School of Computer and Information Sciences hereby authorize Regis University to publish through a Regis University owned and maintained web server, the document described below ("Work"). I acknowledge and understand that the Work will be freely available to all users of the World Wide Web under the condition that it can only be used for legitimate, non-commercial academic research and study. I understand that this restriction on use will be contained in a header note on the Regis University web site but will not be otherwise policed or enforced. I understand and acknowledge that under the Family Educational Rights and Privacy Act I have no obligation to release the Work to any party for any purpose. I am authorizing the release of the Work as a voluntary act without any coercion or restraint. On behalf of myself, my heirs, personal representatives and beneficiaries, I do hereby release Regis University, its officers, employees and agents from any claims, causes, causes of action, law suits, claims for injury, defamation, or other damage to me or my family arising out of or resulting from good faith compliance with the provisions of this authorization. This authorization shall be valid and in force until rescinded in writing.

Title of Document(s) to be published: Implementation of the metadata elements of the INSPIRE Directive

29th August 2008

Student Signature

Check if applicable: Not applicable
The Work contains private or proprietary information of the following parties and their attached permission is required as well:

Name of Organization and/or Authorized Personnel
Abstract

The INSPIRE Directive entered into force in 2007. The INSPIRE (Infrastructure for Spatial Information in Europe) initiative seeks to set in place mechanisms for sharing spatial data across Europe. Metadata is one of these mechanisms: to enable discovery and to act as an information source about spatial data resources. This project examines the implementation of spatial data metadata in Ireland. Using an Irish public service organisation as a case study, the project examines how metadata creation and presentation could be optimised to provide the best value for the end user. This project also examines database tools that can be leveraged to place metadata within the spatial data maintenance workflow, while seeking to reduce the administrative burden on the domain experts who create the metadata.
Acknowledgements

I would like to acknowledge the kind assistance of all the facilitators in this MScSIS program, in particular Brad Blake who kindly offered to be my thesis advisor. I would like to thank my work colleague George McHugh who has often donated his considerable technical expertise to me with great patience. On a personal note I would like to acknowledge the considerable support of both my parents, Patrick and Christine, and my husband Alan who helped me all he could and more than he knew in completing all the modules in this MSc course.
# Table of Contents

Certification of Authorship of Thesis/Practicum Work ........................................... i  
Authorization to Publish Student Work .............................................................. ii  
Abstract ............................................................................................................. iii  
Acknowledgements ............................................................................................ iv  
List of Figures ................................................................................................... iv  
List of Tables ...................................................................................................... v  
Chapter 1 Introduction ....................................................................................... 1  
Chapter 2 Literature Review ............................................................................. 3  
Chapter 3 Metadata Database Definition and Design ...................................... 9  
Chapter 4 Display of metadata files ................................................................. 20  
Chapter 5 Economy of collection and update: using database tools to streamline the creation and update of metadata about spatial data resources ........................................................................... 27  
Chapter 6 Conclusions and Next Steps ............................................................. 40  
References .......................................................................................................... 45  
Appendix 1 ........................................................................................................... 47  
Appendix 2 ........................................................................................................... 50  
Appendix 3 ........................................................................................................... 85  
Glossary .............................................................................................................. 87
List of Figures

Figure 3.1 Initial Entity Relationship Diagram
Figure 3.2 Revised Entity Relationship Diagram
Figure 4.1 General Stylesheet
Figure 4.3 Simplified Frequently Asked Questions Style
Figure 4.2 Frequently Asked Questions Style
Figure 5.1 Existing databases and workflow
Figure 5.2 Existing databases and workflow
Figure 5.3 Creation of SpatialData1 table
Figure 5.4 MD_Records updated
Figure 5.5 Updated spatial data creation workflow
Figure 5.6 Populate trigger autopopulates selected fields
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 3.1</td>
<td>Table showing the INSPIRE Metadata profile elements</td>
</tr>
<tr>
<td>Table 4.1</td>
<td>Table showing the results of expert GIS user review of three stylesheets</td>
</tr>
<tr>
<td>Table 4.2</td>
<td>Table showing the results of expert scientific user review of three stylesheets</td>
</tr>
<tr>
<td>Table 4.3</td>
<td>Table showing the results of expert scientific and GIS user review of three stylesheets</td>
</tr>
<tr>
<td>Table 4.4</td>
<td>Table showing the results of general public user review of three stylesheets</td>
</tr>
<tr>
<td>Table 5.1</td>
<td>Table showing the INSPIRE Metadata profile elements, plotted against whether they could potentially be set as default or autpopulated values in the metadata database</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction

Section 1.1 Statement of Problem and Project aim

The Environmental Protection Agency (EPA) in Ireland, like many of Ireland's public service organisations, includes spatial data as a considerable element of its overall business data. Under the legislation that created the EPA in Ireland (the EPA Act of 1992) the EPA has a role as a data provider for environmental data. In my role of Geographical Information Systems (GIS) Officer for the EPA I have observed that to date, data management processes in the EPA have focused on the collection of data and its physical storage in Microsoft SQL Server databases: collecting and validating spatial data, creating the database and uploading the data. A significant amount of effort has been expended on quality assurance and validation work on data\(^1\). Similarly, while everyone in the organisation has access to the data, non-specialist users have no information available to them on the thematic content of the data. To summarise, the Irish EPA has expended a lot of effort in collecting and validating spatial data, and making this data available, but has not yet considered the importance of providing information to users about the data it holds. Most strikingly, the EPA cannot at any one time produce a list of its data holdings, and some sections of the EPA are not aware of data generated and held by other sections.

As outlined in Chapter Two, the INSPIRE Directive entered into force in March 2007. The key aim of this Directive is to put the mechanisms in place for creation of a spatial information infrastructure for the European Community. As a Member State of the European Community, public service organisations who generate spatial data in Ireland will be required to create metadata to comply with the INSPIRE Directive. Chapter Two summarises a review of the literature carried out to ascertain the role of metadata, contemporary experiences in the creation and maintenance of metadata repositories and the state of play in metadata implementation in Ireland at present.

Section 1.2 Content of the remaining Thesis chapters

In Chapter Three the process of converting the INSPIRE metadata profile into a logical and then physical database is described. Chapter Four includes a discussion of

\(^1\) 250 man-days in 2006 and 500 man-days to date in 2007, recorded through GIS Technician work records.
the role of XML in metadata exchange, and describes how XSL stylesheets can be used to control how metadata appears to users. Following the examination of potential barriers to metadata documented in Chapter Two, Chapter Five describes some potential methods to bring metadata into a typical spatial data creation and maintenance workflow.

Chapter Six describes some conclusions that can be drawn about the practicalities of implementing metadata for the INSPIRE Directive in the Irish EPA. Chapter Six includes an examination of the concepts described in Chapters Three to Five, and how some of them could be developed further.
Chapter 2: Literature Review

The common definition for metadata is “data about data” but the W3C (World Wide Web consortium) add to this definition by describing metadata as “machine understandable information for the web” (www.W3.org). Connolly and Begg (2005) describe metadata as data about the objects in the database. This type of Database Management Systems (DBMS) constrained metadata is more useful to the administrators and users of that DBMS, as it typically covers the data items, their relationships and the users that have access to this data. This project will not focus on this type of metadata; it will instead examine metadata in its role as information about a data resource. This scope will be limited to examination of metadata for spatial data resources, in the light of the recent entry into force of the INSPIRE Directive in the European Union (EU) and its impact on Irish spatial data providers in the public sector, taking environmental spatial data management in the Irish EPA as a case study.

The INSPIRE Directive entered into force in March 2007. The key aim of this Directive is to put the mechanisms in place for creation of a spatial information infrastructure for the European Community. This is based on the need for information including spatial information in the creation and implementation of policy. European-wide policy requires pan-European data and the INSPIRE Directive aims to put in place the mechanisms that will yield a co-ordinated approach to the creation of this data, namely spatial data infrastructures (SDI). The INSPIRE Directive identifies three main “streams” or themes of data, and lists them specifically. SDI are to be designed to ensure that the data is stored, made available and maintained to facilitate sharing at a European Union level, sharing across national public organisations and finally with conditions of use that are not overly restrictive.

Metadata is identified as a mandatory component of a spatial data infrastructure. This need stems from the time and resources that are spent in searching for data, which represent a key obstacle for the full exploitation of that data. To guide EU Member States in implementing the INSPIRE Directive, the Directive drafting team authored a set of Implementing rules, one of which covers metadata. Article 5 of the INSPIRE Directive stipulates that EU Member States must create metadata in accordance with the Implementing rules, for all the data themes that are specifically covered by the
INSPIRE Directive. The metadata must contain *conditions of use* of the data it describes, *limitations on its access* and information about the *quality and validity* of the data. Article 11 of the INSPIRE Directive stipulates that metadata must be served out via *network services*, to allow search and view of metadata files. A metadata profile (collection of metadata elements and XML tags that should be present in the metadata resource) for the INSPIRE Directive has been created in line with ‒ but not completely identical to ‒ the ISO19115 metadata profile.

The INSPIRE Directive sets out the legal requirements on EU Member States to create metadata. What about the functional requirements of metadata ‒ what is its real world application? Batcheller (2008) cites the value of metadata as a data discovery tool, with the emphasis being not so much on the content of the dataset but instead on the marketability information (how a user can get his hands on the data). The INSPIRE Directive places a similar importance on the function of metadata as discovery tool. Longley et al (2001) describe the single most important metadata element as being the telephone number of the last person who used the data. Qi (2004) goes so far as to state that *“data is useless without metadata”*. Perhaps the status of metadata in the Irish public sector (with reference to environmental spatial data at least) hints at the functional value of metadata. There are 5 major players in the generation of this data. The Environmental Protection Agency generates spatial data on land cover, water quality, air quality and sources of contamination. The Environmental Protection Agency hosts a GIS web mapping application (maps.epa.ie) on which they display all their data, with metadata links for the data. The site received 10,000 hits in a three-month period from September to December 2007, but none of these hits were on the metadata links. In the same period the contact point listed on the site, however, received 50 separate queries about the data, its content and how to access a copy of it. Of these 50 queries, 38 individuals received their own digital copies of up to thirteen of the spatial datasets featured on the EPA site: not one of the users received metadata with the data, and there was only one further onward query about the content of the data once the user received it.

---

2 The source for this information is a Google Analytics report for this application.
3 The source for this data is the written telephone records of contact person for this application: the records include the date of the call, the caller name and the nature of the query.
4 The source for this data is the EPA data dissemination database, a Microsoft Access database including records of all the data deliveries made externally to the organisation.
Of the 4 other players, 3 more host similar web mapping applications. The Geological Survey of Ireland and the National Parks and Wildlife Service host web mapping applications with links to metadata that report only the contact details to request data, and the scale and date of creation of the data. The national mapping agency – the Ordnance Survey of Ireland – generates mapping data for resale (their semi-state status requires them to generate a certain amount of their own revenue) but they do not have discovery level metadata on their public facing site. Metadata is delivered with the data on their e-commerce site, but you must already be a registered customer (and have a license agreement in place) to request this data.

If metadata is considered an important enough part of a spatial data infrastructure to necessitate it having a whole Article of the INSPIRE Directive dedicated to it, then why is it that in the Irish case at least, it is sketchily implemented at best? Other authors point at some possible reasons for this. Mathys (2004) postulates that the labour intensity (and monotony) of creating metadata overrides its value as a resource. The same author cites the tedium of creating metadata as the cause for it often being error prone.

Tsou (2002) notes that the two main metadata standards (FGDC and ISO 19115) both contain a high number of elements (300 and 400 respectively) and this high number of elements complicates the creation of metadata profiles and metadata itself unnecessarily.

Lagoze at al (2006) describe the experience of the creation of the National Digital Science Library metadata repository, which led them to conclude that implementation and maintenance of even the lowest level of standardisation of metadata are not as straightforward as they would have expected, and certainly not low in people-costs. They observed that very few of the contributor organisations to the library were able to provide the human resources required to provide good quality metadata. While the creation of metadata is viewed as a monotonous task, it is also a task that requires a high degree of expert knowledge: it requires domain expertise (knowledge of the relevance of the resource in its field) and metadata expertise (understanding of the metadata profile and the correct content of each metadata attribute). While metadata is seen as a data entry task it is more complex, and not a task that could be delegated to
entry-level administrative staff: it must be entered by the domain expert, implying a higher cost (Lagoze et al, 2006). Finally, the domain expert—the person most suited to creating the metadata—may not necessarily be a metadata expert, or indeed used to entering data into an electronic catalogue. Lagoze et al (2006) noted that the experience of the National Digital Science Library was that the support required to help domain experts become familiar with metadata profiles and metadata entry tools resulted in metadata taking several months to be fully captured.

Albertoni et al (2005) raise the question of how easy it is to discover a metadata resource. The authors cite a key problem with metadata being the difficulty in knowing the right query terms to use to find a resource, and the time it takes to sift through those query results. The implementation of metadata clearinghouses that act as a single search point accessing different metadata resource nodes helps to make discovery of data resources easier: the requirement of the INSPIRE Directive to publish metadata resources via network services supports the implementation of such single search points.

This review has focussed on the requirements of the INSPIRE Directive initially, followed by an examination of the implementation of metadata in Ireland. The issues associated with creating, maintaining and discovering metadata have been outlined. Despite these difficulties, metadata is given a high priority in the INSPIRE Directive, and this priority is justified when the advantages of truly valuable metadata are considered: to be valuable to a user metadata should contain currency, quality, validity and accessibility information. Tsou (2002) point about the extent and complexity of metadata profiles is valid, but well-formed metadata should be fit for purpose.

Stvilia and Gasser (2007) illustrate this point through his discussion on metrics of value and cost of information. Stvilia and Gasser use metrics developed by Taylor in 1986, and elaborate these, to define six categories of value that can be added to information:

- Increased ease of use
- Reduction of data "noise"
- Quality of data
- Adaptability of data to other applications
• Time savings
• Cost savings

Examining these metrics with respect to metadata allows the concept of valuable data to be defined more clearly. Valuable metadata should be easy to use. In the context of the metadata consumer, ease of use means the user can discover the metadata and also that the metadata resource is easy for the user to follow. Ideally, this would be achieved by the user being able to control the level of complexity of the metadata by drilling down into more complex parts of the metadata profile as they require, rather than viewing the whole metadata resource at once. As Stvilia and Gasser (2007) note: some metadata elements are more critical than others, and this can vary depending on the nature of the data resource being described. Limiting a metadata resource to those parts of a full metadata profile that are relevant has the effect of reducing data noise, the second of Stvilia and Gasser’s six categories.

Stvilia and Gasser (2007) describe the need to consider quality in the context of its relevance: increasing quality over the required level for a resource may not increase the value of that resource. In the context of metadata, good quality implies accuracy, clarity and relevance: all the absolutely necessary elements should be included but adding even high-quality unnecessary elements will decrease the value of the metadata by increasing the noise.

A primary use of metadata is to allow discovery of data resources. Valuable metadata should exist in a format that allows it to exist or be read by any number of discovery tools. In the context of the World Wide Web, which is the primary platform for discovering data resources, metadata should be accessible and searchable by web browsers, and so should exist in a format that facilitates this the ideal format for metadata is an XML file.

Following Lagoze et al’s (2006) conclusions on the time and cost associated with metadata creation, it is apparent that metadata value is a function of the time it takes to create and maintain it. The first consideration in optimising metadata creation time should be the selection of an appropriate metadata profile. While Tsou (2002) noted metadata profiles are large, not all elements of a profile are mandatory and metadata creators should agree to select the elements from a common profile that are most relevant to the data resources they work with. This selection should consider adaptability as well as relevance especially in the case of publishing.
metadata resources via network services, adherence to a standard profile means the resource can be consumed by other metadata catalogues.

The second consideration in optimising metadata creation time should be for the cost of maintaining a metadata resource once it has been created, so it maintains its value. To facilitate maintenance of metadata, update of metadata resources should be part of the same workflow as update of the data resource it describes, rather than a separate task at the end of often labour intensive data updates (Mathys, 2004). In the case of environmental spatial data, where specialist software is used to update data, complementary software tools could be developed that prompt a user to update certain metadata elements when they commit changes to the spatial database, as the user goes along.

As an EU Member State, Ireland is required to comply with the INSPIRE Directive and (among other things) to create a catalogue of metadata for a number of data resources. The metadata must adhere to the INSPIRE profile. But the metadata should also be valuable, and it should avoid the pitfalls noted from other approaches, in the interest of meeting the requirements of the INSPIRE Directive in a viable manner. As discussed, valuable metadata has a number of characteristics: chief among these are economy in both creation and maintenance costs, as well as in the choice of metadata elements to use from the INSPIRE profile. Finally and in this author’s opinion most importantly the metadata should be valuable to the consumer easy to discover and acting as a clear, authoritative and current source of information about a data resource.
Chapter 3 Metadata Database Definition and Design

Chapter Two (Literature Review) introduced the concept of data value, and characteristics of valuable metadata.

To create a database of valuable, INSPIRE-compliant metadata it is necessary to do the following:

1. Review the INSPIRE profile in order to create a logical – and then physical – model of the metadata database.
2. Create a catalogue within this database that automatically updates data resources in the database so that their associated metadata can be tracked.
3. Create accessible views of the metadata files that are appropriate to the user audience of that particular view.
4. Research and suggest a mechanism for keeping metadata up to date within the workflow of data creation and maintenance.

The first item in the list will be discussed in this chapter.

Section 3.1 A review of the INSPIRE metadata profile.

The primary input for the metadata database design, the INSPIRE metadata profile, is a selection (profile) of the International Organisation for Standardisation (ISO) spatial metadata profile, ISO19115 (and also the spatial data services metadata profile, ISO 19119, which for INSPIRE has elements included in an overall INSPIRE profile).

The Draft Guidelines for INSPIRE metadata implementing rules (Drafting Team and European Commission, 2007) set out the details of each metadata element: rather than transpose this document into this chapter, it is preferable to create a matrix showing each metadata element identifying whether it is a mandatory or optional element - its relevance and its intended. The details given in the Draft Guidelines (Drafting Team and European Commission, 2007) will be used again later on in this chapter to create the model for the metadata database.

First, it is necessary to define how we intend to assess the relevance of each field: this will determine whether it is a required field in the database and also determine the display mechanisms to be discussed in Chapter Four. It is possible to review the metadata profile in the context of the expected users of metadata. As we have
established in Chapter Two, there are two main applications of metadata. The first is as a resource for discovery of data resources. The second is as an information source about a data resource.

The type of data resource determines the potential users of metadata. The EPA data requests database was used to provide information about the users that have made queries related to data, or requested data, from the EPA since 2003. This database is a Microsoft Access application that was developed within the EPA in 2003 to record the type of queries that are received about spatial data, and to record the contact details of all users who requested electronic copies of spatial data. This database was analysed to determine where the queries originated, and to track what types of queries were most common. The analysis reveals four categories of users with reasonably distinct query patterns:

- **General public as browsers.** This group is categorised by being from any type of professional background, seeking only to access environmental data about an area of interest, and seeking to browse the data resource via the World Wide Web rather than access electronic copies of the data. Assessing the types of queries these users make shows this group are typically interested where the data comes from and how up to date it is.

- **Expert commercial users.** This group is comprised of professional consultancy services (Civil and Environmental Engineers) who wish to access data to use in provision of professional, for profit services on behalf of private sector or public sector clients. Typically, this group are concerned with how up to date the data is, how it can be accessed and if there are restrictions on its use. It is usual for this group to make a small volume of very specific queries about the data content (units of measurement, data collection methods etc). This group are the most likely to request regular updates of this data.

- **Expert non-commercial users.** This group is made up Non-Government Organisations and third level education post graduate and academic staff. Typically, this group request a wide range and high volume of data often relating to a specific location for once off research projects. This group make the highest volume of queries about the data. Particularly in the case of the post graduate researchers, who are using the data as an education resource, requests either about the collection methods for the data, or how the data can be applied (spatial data requires GIS software to access it).
• **Intra-organisational users.** EPA staff made a number of requests for data, and their requests are similar to both the expert commercial users who had specific technical details about the content of the data, and the source of data in cases where it was sourced outside of the EPA.

Returning to the potential users of metadata, it is clear that all of the user-groups mentioned above could potentially use metadata as a discovery resource. This application is independent of the type of audience because it fulfils a common function for all users: discovery-level metadata describes the data content and provides information about where and how the data can be accessed. It is up to the user to determine if the data is suitable for their requirements by investigating the data resource further.

In terms of an information resource, the analysis summarised above shows that certain user groups tend to place more importance on certain elements of the data. To this end, a matrix of INSPIRE profile elements versus intended audience was created (Table 3.1) plotting each field of the profile against its perceived usefulness to each user group. The nature and content of environmental spatial data itself has a bearing on the relevance of elements of the metadata profile. For time-series monitoring data particularly related to air and water pollutants information about the currency of the data is crucial. In the case of interpolated data, generated by modelling techniques, the underlying interpolation method and degrees of confidence become more important. There is no absolute rule of thumb for creating a perfect metadata profile, and elements that are mandatory for one environmental spatial dataset may not be required at all for another. Plotting each field against the relevance for each audience group leads to an observed overall relevance that may or may not match the obligation imposed by the INSPIRE implementing rules (see Table 3.1).

The highest perceived obligation (the Obligation (User Requirements) field of the matrix, (Table 3.1) is based on the highest obligation perceived for an element against any group of users. This matrix is an attempt to point towards a database structure for the metadata, and also for how the data from this database will be shown to the users, so it leads to two different types of information.
Where an element is mandatory for any user group or for conformance to the INSPIRE Directive then it must be included in the database so the data for that element can be captured. However, if it is perceived that the element is not required for a user group then ideally it should not be included in the metadata resource for that group to reduce clutter.
Table 3.1: INSPIRE metadata profile elements plotted against obligation

<table>
<thead>
<tr>
<th>Element Groups</th>
<th>Element</th>
<th>Purpose</th>
<th>Obligation (INSPIRE)</th>
<th>General Public</th>
<th>Expert Commercial Users</th>
<th>Expert non-commercial</th>
<th>Intra-organisational users</th>
<th>Obligation (User Requirements)</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ResourceTitle</td>
<td>Name of data resource</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>All elements that are required for Discovery should be considered mandatory</td>
</tr>
<tr>
<td></td>
<td>ResourceAbstract</td>
<td>Brief summary description of the data resource</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>See above</td>
</tr>
<tr>
<td></td>
<td>ResourceType</td>
<td>Defines whether resource is a dataset, series of datasets or a web service</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>See above</td>
</tr>
<tr>
<td></td>
<td>ResourceLocator</td>
<td>URL address for online access: conditional on resource having this type of access available</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>Conditional only on the data having this type of access</td>
</tr>
<tr>
<td></td>
<td>UniqueResourceIdentifier</td>
<td>Value that uniquely defines a data resource (particularly within a series of datasets)</td>
<td>M</td>
<td>O</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>All but the general public may want to access a specific copy of the resource</td>
</tr>
<tr>
<td></td>
<td>CoupledResource</td>
<td>Relevant to metadata about web services: this element describes the datasets the service is based on</td>
<td>C</td>
<td>O</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>Conditional on the item being a web service. Also, in this case the General Public are considered to only want to access the data to view: they would not be expected to have the specialist software required to access the data.</td>
</tr>
<tr>
<td></td>
<td>Identification</td>
<td>ResourceLanguage</td>
<td>M</td>
<td>O</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>Required for Discovery metadata</td>
</tr>
<tr>
<td></td>
<td>Classification</td>
<td>TopicCategory</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>Required to make the metadata a &quot;searchable&quot; resource</td>
</tr>
<tr>
<td>Field</td>
<td>Description</td>
<td>Mandatory (M)</td>
<td>Optional (O)</td>
<td>M</td>
<td>O</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>---------------</td>
<td>--------------</td>
<td>---------------</td>
<td>---------------</td>
<td>---------------</td>
<td>---------------</td>
<td>---------------</td>
<td>---------------</td>
</tr>
<tr>
<td>SpatialDataServiceType</td>
<td>Type as per a list of possible service types</td>
<td></td>
<td></td>
<td>M</td>
<td>O</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>KeywordValue</td>
<td>Words that can be used to describe - and so search for - a resource. They are taken from a defined vocabulary.</td>
<td></td>
<td></td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Keyword</td>
<td>Keywords that can be used to describe - and so search for - a resource. They are taken from a defined vocabulary.</td>
<td>C</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>ThesaurusName</td>
<td>Identifies the defined vocabulary that is used to define the keyword</td>
<td></td>
<td></td>
<td>C</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>C</td>
</tr>
<tr>
<td>WestBoundLongitude</td>
<td>One of four elements that define a bounding box for the geographical extent of the data resource</td>
<td>M</td>
<td>O</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>EastBoundLongitude</td>
<td>See above</td>
<td>M</td>
<td>O</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>SouthBoundLongitude</td>
<td>See above</td>
<td>M</td>
<td>O</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>NorthBoundLongitude</td>
<td>See above</td>
<td>M</td>
<td>O</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>TemporalExtent</td>
<td>Period of time that the data resource covers</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>DateOfPublication</td>
<td>Date on which the resource was published</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>DateOfLastRevision</td>
<td>Date of the last revision to the dataset, since the publication date</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>TemporalReference</td>
<td>DateOfCreation</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Lineage</td>
<td>Brief statement about the origin and revision history of the dataset</td>
<td>M</td>
<td>O</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Quality &amp; Validity</td>
<td>SpatialResolution</td>
<td>C</td>
<td>O</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Degree</td>
<td>Conformance of resource with INSPIRE implementing rules, True or False</td>
<td>M</td>
<td>O</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>O</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Conformity</td>
<td>Specification</td>
<td>M</td>
<td>O</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>O</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>LimitationsOnPublicAccess</td>
<td>Limitations on access to data, related to Intellectual Property, Privacy etc</td>
<td>M</td>
<td>O</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Access Constraints</td>
<td>ConditionsOfUse</td>
<td>M</td>
<td>O</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
</tbody>
</table>

Again, optional to the public as this is not considered relevant to them.

Required to make the metadata a "searchable" resource.

This is relevant to only very specialist users.

As the public would be viewing the data, these co-ordinates are not immediately relevant to them - or indeed obvious.

All the "time" attributes are conditional. One should be considered mandatory but it depends on the nature of the data resource.

Required for expert users.

Required for expert users.

Required for GIS experts, but not any other type of user.

Required for GIS experts, but not any other type of user.

Mandatory where electronic copies of the data are sought - so not relevant to the public or internal users.

See above.
<table>
<thead>
<tr>
<th>Metadata on Metadata</th>
<th>MetadataDate</th>
<th>Date that the metadata was created</th>
<th>M</th>
<th>M</th>
<th>M</th>
<th>M</th>
<th>M</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metadata on Metadata</td>
<td>MetadataLanguage</td>
<td>Language used in creating the metadata file</td>
<td>M</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>M</td>
</tr>
<tr>
<td>MetadataContact</td>
<td>Point of contact for information about the metadata</td>
<td>M</td>
<td>O</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td></td>
</tr>
<tr>
<td>Responsible Party</td>
<td>ResponsiblePartyRole</td>
<td>The function performed by the responsible party identified above</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>ResponsibleParty</td>
<td>Contact details of organisation and party responsible for the data resource</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>MetadataDate</td>
<td>Date that the metadata was created</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>MetadataLanguage</td>
<td>Language used in creating the metadata file</td>
<td>M</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>M</td>
<td></td>
</tr>
</tbody>
</table>

**Responsible Party**

Contact details of organisation and party responsible for the data resource

**Responsible Party Role**

The function performed by the responsible party identified above

**MetadataContact**

Point of contact for information about the metadata

**MetadataDate**

Date that the metadata was created

**Metadata on Metadata**

Date that the metadata was created

**MetadataLanguage**

Language used in creating the metadata file

**Should form part of the Discovery level metadata**

See above

**Required for expert users**

Considered mandatory because it shows how up to date the information about a resource is kept, even if the resource itself is old data

**Should be obvious from the content of the file**
Section 3.2 Metadata Database Design

The matrix (Table 3.1) and the INSPIRE guidance material (Drafting Team and European Commission, 2007) provide the basis for the initial Entity Relationship Diagram.

There are two entities: Data Resources and the Metadata Records that describe them. The elements form the attributes of the metadata records, but straightaway it becomes apparent that transposing the complete list of elements to the single metadata record entity is not desirable. There is potential for update anomalies through the existence of free-text fields (free-text prescribed by the INSPIRE guidance material (Drafting Team and European Commission, 2007)) when drop-down lists could be defined to enforce consistency. Secondly, these potentially large free text fields can be reduced to two or three character codes in the metadata record, linking to look-up tables with the full text values, so improving database performance.

The database was therefore normalised as per the method outlined in both Rob and Coronel (2004), and Connolly and Begg (2005). The primary key of the metadata
table must be the unique identification number for the associated resource dataset, each resource dataset having a one-to-one relationship with a metadata record.

There are a number of potential repeating groups in the metadata record. First of all, the ResponsibleParty and MetadataContact fields require entry of full contact details, and will probably refer to a small group of people within a single organisation. The contact details may even refer to the same person, so repetition of data is limited by having a single contacts table.

As highlighted above, there are a number of free-text fields that will contain identical records, and so will potentially lead to more repeated data:

- Language (ResourceLanguage and MetadataLanguage) could be replaced by 2-character country codes to link to the full language text
- TopicCategory originates from a defined list that could be given a code
- SpatialDataServiceType is one of three possible variables
- Specification references a legislative or other type of specification that conformance of the data resource is measured against: this is expected to be a set group of specifications
- LimitationsPublicAccess could be grouped into codes that represent a fixed set of access types, from full access to no access, with degrees in-between
- Similarly, ConditionsOfUse can be a defined list of conditions
- There is an Organisation field in the contacts table: this can be defined list of organisation contacts, maintained separately to limit repeating address fields in the contacts table.

Revising the data model to convert the free text fields to coded values yields the following:
As this project is using a real organisation as the case study, with an existing SQL Server 2005 installation, the database was created in SQL Server 2005. Ten typical metadata record entries (from the environmental spatial data domain) were made in order to validate the database design. The script used to generate the database tables is included in Appendix 1.

The tables were populated directly but it is envisioned a practical implementation of this database would have a web-based data entry form as the data capture interface. Creation of this form is outside the scope of this project but is discussed further in the considerations of metadata workflow, in Chapter Five.

Entry of the data validated that the database design was adequate for storage of the data, and that the use of the lookup tables maintained good data consistency. The next
stage in the project was to create a mechanism for users to view the metadata. This is discussed in Chapter Four.
Chapter 4: Display of metadata files

Chapter Two (Literature Review) introduced the concept of data value, and characteristics of valuable metadata.

Chapter Three discussed the design of an INSPIRE-compliant metadata database. This chapter discusses the potential for a method to retrieve the data from this database, and then present it in a way that assures its value to the user.

Section 4.1 XML as a metadata file format

The INSPIRE Directive itself does not specify the file format that metadata must be saved and exchanged in. Guidance documents relating to INSPIRE (Drafting Team Metadata, 2007) do not explicitly specify XML file formats, but it is noted in this document that metadata data repositories should be accessible, and that XML format is an ideal way to achieve this. The ISO19115 standard for metadata on which the INSPIRE metadata profile is based similarly espouses XML as the ideal file format for metadata (International Organisation for Standardisation, 2003). Carey (2007) describes some of the features of XML that make it such an ideal format for metadata files: XML documents are supported in a wide variety of applications, XML is easily usable over the Internet and XML documents are relatively easy to create. These advantages have lead to XML becoming a de facto standard for metadata exchange.

Section 4.2 Retrieving XML from a SQL Server 2004 database

As the Irish EPA was to be used as the case study for this project, all activities relating to this project were constrained to using the corporate standard software in this organisation. The corporate RDBMS is Microsoft SQL Server 2005. An advantage of this RDBMS is the increased support provided for XML in the 2005 version.

XML data can be extracted directly from a SQL Server 2005 database via "FOR XML" statements in addition to standard SQL statements to create XML files. The exact syntax of the query used to select a metadata record from the database is shown in Appendix 2 (GetMetadataXML).

The query selects each field of the database and writes it as an element in an XML file, with the table name as the root element. A sample of the output can be seen in Appendix 2. The FOR XML query is effective to extract the data from the database,
but refining this process and allowing users to write XML metadata files for selected data resources on-the-fly is not covered in this project (see Chapter Six for more discussion).

**Section 4.3 Using Stylesheets to control the user’s view of a single metadata resource**

Once the data is extracted from the database it must be presented to the user in a user-readable format. One way to achieve this is by applying a stylesheet to the XML file to control the appearance of the output. Chapter Two discussed the value of metadata, and noted that it was important to make the metadata an accessible resource for all users. In Chapter Three, a matrix of perceived value of the various metadata elements to the four general spatial data user groups was created (Table 3.1). The use of stylesheets has the potential to tailor metadata files to suit the particular user audience. To prove this concept, three metadata stylesheets were created. Their syntax is shown in Appendix 2.

The first stylesheet was based on the existing stylesheet used in the EPA metadata application introduced in Chapter Two (Styleheet1_General.xsl). This stylesheet extracts all the elements from the related xml file and applies the INSPIRE names (or full versions of the names) to each element. This stylesheet is very much aimed at the expert user, as the sample below shows:

**EPA Spatial Data Resources**

**Data Resource Identification**

**Data Resource Title:** Soils map  
**Data Resource Abstract:** National soils map prepared by Teagasc, 2006  
**Data Resource Locator:** See responsible party contact details  
**Unique Identifier for this data resource:** n/a  
**Coupled Resource:** n/a  
**Resource Language:** English

**Data Resource Classification**

**Topic Category:** Geoscientific Information  
**Spatial Data Service Type:** Geographic viewer

**Data Resource Keywords**

**Keyword(s):** soil  
**Keyword(s) Thesaurus:** n/a

**Data Resource Location**
West Longitude: 0
East Longitude: 0
North Longitude: 0
South Longitude: 0

Data Resource Temporal Reference

Start of Temporal Range: 2004-03-31T00:00:00
End of Temporal Range: 2006-05-31T00:00:00
Publication Date: 2006-05-31T00:00:00
Revision Date: 2006-05-31T00:00:00
Creation Date: 2004-03-31T00:00:00

Data Resource Quality and Validity

Lineage: Details given in report, teagasc_class_combined.doc, available on request

Data Resource Conformity

Degree of Conformance: Not Evaluated
Specification: Not Evaluated

Data Resource Access Information

Limitations on Public Access: Open access to public
Conditions Of Use: Conditions apply

Data Resource Responsible Party

Responsible Contact Surname: O'Rourke
Responsible Contact First Name: Fiona
Responsible Contact Phone: 00353539170729
Responsible Organisation: Environmental Protection Agency
Responsible Address Line 1: PO Box 3000
Responsible Organisation Address Line 2: Johnstown Castle Estate
Responsible Organisation Address Line 3: Wexford
Responsible Organisation Website: www.epa.ie
Responsible Organisation Role: Distributor

Information about this Metadata

Metadata Contact Surname: Jones
Metadata Contact First Name: Joe
Metadata Contact Phone: 00353531234567
Metadata Organisation: Environmental Protection Agency
Metadata Address Line 1: PO Box 3000
Metadata Organisation Address Line 2: Johnstown Castle Estate
Metadata Organisation Address Line 3: Wexford
Metadata Organisation Website: www.epa.ie
Metadata Language: English
Last Update: 2008-08-11T00:00:00

Figure 4.1 General Stylesheet

The second stylesheet (Stylesheet2_FAQAll.xsl) changes all the element names to questions, in order to achieve a "Frequently Asked Questions" approach. This may be
more accessible to the non-expert user but as noted in Chapters Two and Three, there are still a number of fields that the non-expert user may not need. This increases the data ìnoiseî which in turn limits the usefulness of the metadata as an information resource about the data:

EPA Spatial Data Resources

Some Identification details for this data:

What is the common name of this dataset? Soils map
Can I have a brief description of this data? National soils map prepared by Teagasc, 2006
Is there a place online that I can access this data from? See responsible party contact details
Is there a name that uniquely identifies this data? n/a
What data is this service based on? (for web services only) n/a
What language is this data written in? English

Data Resource Classification

What data category does this data belong to? Geoscientific Information
What type of data service is this? (for web services only) Geographic viewer

Data Resource Keywords

What key words describe this data? soil
Which thesaurus do these keywords come from? n/a

Data Resource Location

What is the western bounding co-ordinate for this data? 0
What is the eastern bounding co-ordinate for this data? 0
What is the northern bounding co-ordinate for this data? 0
What is the southern bounding co-ordinate for this data? 0

Data Resource Temporal Reference

From what starting date is this data relevant? 2004-03-31T00:00:00
To what end date is this data relevant? 2006-05-31T00:00:00
What date was this data published? 2006-05-31T00:00:00
What date was this data revised? 2006-05-31T00:00:00
What date was this data created? 2004-03-31T00:00:00

Data Resource Quality and Validity

What is the history of this dataset? Details given in report, teagasc_class_combined.doc, available on request

Data Resource Conformity

Does this data conform to INSPIRE? Not Evaluated
What requirements does this dataset adhere to? Not Evaluated

Data Resource Access Information
Are there limitations on access to this data? Open access to public
Are there conditions for using this data? Conditions apply

Who do I contact about this data?

Data contact Surname O’Rourke
Data contact First Name: Fiona
Data contact Phone: 00353539170729
Data contact Organisation: Environmental Protection Agency
Data contact organisation address line 1: PO Box 3000
Data contact organisation address line 2: Johnstown Castle Estate
Data contact organisation address line 3: Wexford
Organisation Website: www.epa.ie
Organisation Role: Distributor

Information about this Metadata

Metadata Contact Surname: Jones
Metadata Contact First Name: Joe
Metadata Contact Phone: 00353531234567
Metadata Organisation: Environmental Protection Agency
Metadata Address Line 1: PO Box 3000
Metadata Organisation Address Line 2: Johnstown Castle Estate
Metadata Organisation Address Line 3: Wexford
Metadata Organisation Website: www.epa.ie
Metadata Language: English
Last Update: 2008-08-11T00:00:00

Figure 4.2 Frequently Asked Questions Style

The final sample stylesheet (Stylesheet3_FAQPublic.xsl) uses the same Frequently Asked Questions style, but the number of elements are limited to only those fields identified as mandatory in the matrix presented in Chapter Three (Table 3.1). This is an attempt to reduce the potential for data noise and so increase the usefulness and accessibility of this data to non-expert users.

EPA Spatial Data Resources

What is the common name of this dataset? Soils map
Can I have a brief description of this data? National soils map prepared by Teagasc, 2006
Is there a place online that I can access this data from? See responsible party contact details
What data category does this data belong to? Geoscientific Information

What key words describe this data? soil

From what starting date is this data relevant? 2004-03-31T00:00:00
To what end date is this data relevant? 2006-05-31T00:00:00

Who do I contact about this data?

Data contact Surname O’Rourke
Data contact First Name: Fiona
Data contact Phone: 00353539170729
A sample metadata file in each of the three formats was shown to a small group of eight potential users in order to get their feedback on the usefulness of each file. The group comprised of two expert spatial data users, two scientific experts with limited GIS experience but who access spatial data over the EPA Intranet WebGIS, one expert GIS user who managed a scientific data theme (water quality) and three non-expert users outside of the EPA who were to act to represent the general public users.

Each user was presented with a very brief (Microsoft Excel) score sheet and asked to provide a score on their perception of the usefulness of the metadata files in each of the three formats. The scores, along with any comments the users made, are presented in Appendix 2. Specifically, the users were asked to score the metadata files on a scale of one to ten (ten being the highest) under the following criteria:

- Relevance of the information presented
- Clarity of the information (was it clear what each entry in the file represented?)
- Format of the information
- Completeness of the information (were any essential pieces of information missing from the file?)

Finally the users were asked to comment if assuming improvements were made to the metadata where required the metadata files were useful all the users replied that they were a useful resource.

The average scores returned were as follows:

<table>
<thead>
<tr>
<th>Stylesheet</th>
<th>Relevance</th>
<th>Clarity</th>
<th>Format</th>
<th>Completeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-General</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 4.1: Results from two Spatial Data experts (non-scientist)
Table 4.2: Results from two Scientists (non-GIS experts)

<table>
<thead>
<tr>
<th>Stylesheet</th>
<th>Relevance</th>
<th>Clarity</th>
<th>Format</th>
<th>Completeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-General</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>2-FAQ_All</td>
<td>4</td>
<td>7</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>3-FAQ_Public</td>
<td>3</td>
<td>7</td>
<td>7</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 4.3: Results from one scientist and spatial data expert

<table>
<thead>
<tr>
<th>Stylesheet</th>
<th>Relevance</th>
<th>Clarity</th>
<th>Format</th>
<th>Completeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-General</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>2-FAQ_All</td>
<td>3</td>
<td>8</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>3-FAQ_Public</td>
<td>2</td>
<td>7</td>
<td>6</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 4.4: Results from three general public users

<table>
<thead>
<tr>
<th>Stylesheet</th>
<th>Relevance</th>
<th>Clarity</th>
<th>Format</th>
<th>Completeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-General</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>2-FAQ_All</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>3-FAQ_Public</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
</tbody>
</table>

The specific comments relating to relevance and completeness revealed that users noted some items were missing from the metadata profile itself: the scientists noted that there was no mention of the individual data fields, or any form of dictionary to explain what they meant. It was commented that the responsible party contact details were perhaps the most valuable element of the metadata. Note that some of the comments could be attributed to the quality of the metadata itself: real rather than test metadata would be expected to have richer details added to at least the Lineage field.

The general public users noted that the field names on the full profile were difficult to decipher, and while the FAQ style was better it could still be improved in terms of user friendliness of format and improved clarity. It was also noted that some metadata items could have fuller descriptions, for example the west, east, north and south longitudes should have been replaced with a text description of the geographic extent (e.g. “Covers the whole of the island of Ireland”).

Other conclusions are discussed in Chapter Six.
Chapter 5: Economy of collection and update: using database tools to streamline the creation and update of metadata about spatial data resources

In Chapter Two (Literature Review) Lagoze et al’s 2006 study of the National Digital Science Library metadata repository noted some key difficulties with metadata collection. Specifically, this included the high people costs associated with requiring the domain expert to handle the creation and update of the specific data layers within their domain of expertise.

Also in Chapter Two, a review of the state of play with respect to spatial metadata resources across the key player organisations in Ireland revealed metadata was not widely implemented and where it was the metadata files may not have been updated for some time.

It seems that are potential risks for the implementation of the INSPIRE Directive in Ireland in both the initial capture and update of metadata. Initially, it can be expected that there will, of necessity, be a large investment of domain expert time in creating metadata in order to meet the requirements of the Directive, but this investment will be nullified somewhat if the metadata is not kept up to date after this initial capture exercise.

Using the Irish Environmental Protection Agency’s GIS unit as the case study, it was noted that the workflow of spatial data creation and maintenance did not include metadata as an intrinsic part. Spatial databases were set up to serve intranet and internet WebGIS applications, but the publishing of data to these tools did not include the task of creating metadata. It seems there is potential for including metadata capture in the workflow of data creation and maintenance, but attempting to enforce this by documenting new Standard Operating Procedures may not the most efficient way to achieve this. Instead, it may be more beneficial to examine the potential of automating certain aspects of this workflow to include automatic capture of the metadata where possible, and where this is not possible to at least include automatic prompts and checks along the workflow to make metadata maintenance a more regular part of the overall spatial data maintenance workflow.
Section 5.1 Examining the existing workflow and the potential for database triggers

The existing spatial database set-up and workflow for the Irish EPA is shown below (Figure 5.1). The spatial data resides in a SQL Server 2005 database with a spatial data add-on called ArcSDE (Arc Spatial Database Engine), added to allow the SQL Server database handle and store spatial data. EPA users can access spatial data in two ways: expert GIS users access it through GIS software installed on the desktop, and non-specialist users access it through a customised EPA WebGIS.

There are two key databases: AppsDB serves existing intranet WebGIS applications, and DesktopDB serves the more specialist desktop GIS users. Data in the AppsDB is generally a copy of the data in the DesktopDB database, but is kept separate from the everyday DesktopDB. This is because the WebGIS applications that run on the AppsDB must be refreshed when data is updated to this database and this refresh makes these applications temporarily available to users. This AppsDB cannot be subject to constant data editing: publishing data updates to this database must be done in a structured batch, preferably out of normal working hours, to limit disruption to the up to 250 staff who may need to use the WebGIS applications. Editing and creation of new spatial data is carried out by specialist users, on the DesktopDB database. Some other specialist users have read-only access to this database. Most significantly in term of GIS, there is no metadata stored in either the DesktopDB or the AppsDB. Capture of metadata is not part of the current workflow.

Note that the actual server and database names are not shown.
In Chapter Two some of the potential obstacles to metadata capture and maintenance were identified from contemporary literature on the topic. Mathys (2004) noted the importance of placing metadata within the data update workflow. This would have two beneficial effects: it would familiarise all the data custodians with metadata and the elements of the metadata profile, and it would also mean that the initial investment involved in creating metadata was not negated by failing to keep that metadata up to date.

To place metadata in the EPA spatial data workflow, it is necessary to first implement the metadata database discussed in Chapter Three. Ideally, this should be used to describe all the data published to the DesktopDB, the AppsDB being a subset of the DesktopDB (Figure 5.2).

It is the Write Metadata actions to the Metadata DB that require examination: beyond implementing the metadata database, what can be done to enforce the position of metadata capture and maintenance within an established workflow?
All metadata capture events should be linked to a data creation event. This can be accomplished through the use of a database tool (e.g. a trigger) that would record this event and remind the user to enter the metadata for the new database table.

To prove this concept a set of triggers were created on a test database. The triggers operated to:

1. Log all Create_table events in an event log
2. Select the table name and the user name from the event log and
3. Populate the table name and user name to the metadata database (md_records table), in order to create a blank metadata record for the new table
4. Email a reminder to the user, based on their user name, to prompt them to enter a metadata record

The syntax of each of the triggers is included in Appendix 3, but the following is a synopsis of the operation of the triggers. First, a user enters a new table (SpatialData1)
in the database (test_data). A trigger (UpdateMD_records)\(^6\) activates to update the event_log is updated (Figure 5.3).

Figure 5.3. Creation of SpatialData1 table

The same trigger selects the table name and user name from the Event_log table, and populates it to the MD_Records table of the metadata database (Figure 5.4).

Figure 5.4. MD_Records updated

\(^6\) See appendix 3 for syntax
On insertion of the new metadata record, a second trigger (EMailReminder) is activated in order to remind the user that there is a new metadata record waiting to be updated. Similarly, the triggers could be used to open a metadata entry form: if the user wasn’t ready to enter the metadata at that point the form would be dismissed (and potentially forgotten about). The benefit of an email reminder is that it can remain in the user’s inbox until (s)he has the time to attend to it.

Despite this reminder there is still potential for the user to overlook the creation of the metadata. Here it is the role of the DBA to check for blank metadata records (which (s)he can cross check with the dates recorded in the event_log). A simple query (QryBlankMetadata) can be used to isolate these records for follow up.

Similar triggers can be implemented for edit events on the table (for example, INSERT data manipulation commands) to prompt the user to update the metadata on data updates.

Spatial data in particular can be complex to build. It may take many iterations of editing and updating before a table is finished. For example the generation of statistics on noise population exposure statistics for a major road (to be reported to the European Commission) involves four input datasets, elements of which are joined to a dataset in separate steps in order to create a final dataset. In this data creation, three temporary database tables are created before the final table is ready, each step in the process yielding a new output. This typical workflow reveals a potential problem with the triggers: none of the temporary tables are publishable data and will be deleted. However, the triggers would still have created new metadata records for them.

There are two potential workarounds: the first is to reserve the DesktopDB and the AppsDB as “published-data-only” databases and to implement the triggers on the DesktopDB. This would require creation of a “staging” or “working” database that can be used for preparation of data.

The second potential approach is to use naming conventions to name database tables that would flag the temporary tables (e.g. prefix them with TMP) and allow them to be omitted from the part of the trigger that creates the new metadata record (only updates records to the md_records table that have names not like TMP).

---
7 See Appendix 3 for syntax
The first option is preferable to the EPA case study: this is because of the potential volume of temporary tables that can be created in a complex data update task, so making the naming convention cumbersome: particularly for tables that will not be published and will be deleted from the database. The second factor is the relatively small group of expert users who actually write to the spatial database compared to the larger number of data readers. Users of the DesktopDB should only see publication-ready data, rather than be confused by the presence of temporary tables, so it is more correct to include a staging "working" database and only create metadata for those tables considered fit for publication to the spatial data repository.

The triggers should be placed in the DesktopDB (AppsDB is a copied subset of this data, so it is appropriate to catalogue DesktopDB only). This yields a new database configuration and workflow (Figure 5.5)

![Figure 5.5. Updated spatial data creation workflow](image)

**Section 5.2 Examining the potential for automatic capture of metadata elements**

Applying automatic tools that help to place metadata capture in a workflow will help to ensure that metadata is captured and maintained: in Chapter Two a number of potential obstacles to maintaining metadata were identified that relate to the labour intensity (Lagoze et al, 2006) and potential monotony of creating metadata (Mathys, 2004). Similarly, Stvilia and Gasser’s (2007) metrics for quantifying the value of
metadata note that ease of use, time savings (and by extension cost savings) are all factors that relate to the value of a metadata resource. The metadata capture and maintenance workflow must be designed with ease and speed of use in mind. Batcheller’s (2008) paper on the potential for automating metadata capture provides some useful ideas in this regard.

The INSPIRE metadata profile matrix created in Chapter Three (Table 3.1) describes the content of each of the metadata fields. The creation of the metadata database, also described in Chapter Three, proved that a number of the fields had a limited number of possible entries and so could be linked to look-up tables: this will assist to streamline data entry. The nature of the dataset itself limits the potential contents of the fields even further: for example, certain types of data will belong in a single INSPIRE metadata data topic category such as Farming or Inland Water.

The content of the data itself can be another source of metadata information: for example, the geographical bounding box co-ordinates (the four metadata elements relating to Longitude) are contained in the data itself. And finally, the context of the data: within Ireland, and within the EPA data management workflow and procedures, will inform other elements such as Access Limitation and Conditions of Use.

Finally, outside of the data, default values can be auto-populated to fields such as the Creation Date or Publication Date field, which the user can then have the option to update should (s)he choose.

A second matrix can be created to note the potential for automatic capture of data, either from the context, nature or content of the data itself (Table 5.1).
Table 5.1 Table plotting INSPIRE metadata elements against potential to be default or autopopulated values

<table>
<thead>
<tr>
<th>Element Groups</th>
<th>Element</th>
<th>Purpose</th>
<th>Autopopulate Code Key:</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ResourceTitle</td>
<td>Name of data resource</td>
<td>N</td>
<td>This is a unique identifier per dataset</td>
</tr>
<tr>
<td></td>
<td>ResourceAbstract</td>
<td>Brief summary description of the data resource</td>
<td>N</td>
<td>This is unique information per dataset</td>
</tr>
<tr>
<td></td>
<td>ResourceType</td>
<td>Defines whether resource is a dataset, series of datasets or a web service</td>
<td>D</td>
<td>A default value of “dataset” can be applied for the metadata author to accept or overwrite: the majority of the EPA’s data holdings are single datasets</td>
</tr>
<tr>
<td></td>
<td>ResourceLocator</td>
<td>URL address for online access: conditional on resource having this type of access available</td>
<td>D</td>
<td>For the EPA this would be a single url</td>
</tr>
<tr>
<td></td>
<td>UniqueResourceIdentifier</td>
<td>Value that uniquely defines a data resource (particularly within a series of datasets)</td>
<td>D</td>
<td>The database table name for this resource, already captured via the triggers explained in section 6.2</td>
</tr>
<tr>
<td>Identification</td>
<td>CoupledResource</td>
<td>Relevant to metadata about web services: this element describes the datasets the service is based on</td>
<td>N</td>
<td>This could be a drop down list of all the data objects</td>
</tr>
<tr>
<td>Classification</td>
<td>ResourceType</td>
<td>Defines whether resource is a dataset, series of datasets or a web service</td>
<td>D</td>
<td>All EPA datasets are created in English, this can be a default value</td>
</tr>
<tr>
<td></td>
<td>SpatialDataServiceType</td>
<td>Type as per a list of possible service types</td>
<td>Y</td>
<td>Information about the topic can be included as part of the data table name by applying a convention</td>
</tr>
<tr>
<td></td>
<td>KeywordValue</td>
<td>Words that can be used to describe - and so search for - a resource. They are taken from a defined vocabulary.</td>
<td>N</td>
<td>Not possible as the list is complex</td>
</tr>
<tr>
<td>Keyword</td>
<td>ThesaurusName</td>
<td>Identifies the defined vocabulary that is used to define the keyword</td>
<td>N</td>
<td>This relates to the topic code and so could be captured from the data table name by applying a convention</td>
</tr>
<tr>
<td></td>
<td>WestBoundLongitude</td>
<td>One of four elements that define a bounding box for the geographical extent of the data resource</td>
<td>Y</td>
<td>Thesauri are not implemented at present: when they are, they could potentially be linked to the keyword types</td>
</tr>
<tr>
<td></td>
<td>EastBoundLongitude</td>
<td>see above</td>
<td>Y</td>
<td>This is an attribute of the data table itself</td>
</tr>
<tr>
<td></td>
<td>SouthBoundLongitude</td>
<td>see above</td>
<td>Y</td>
<td>See above</td>
</tr>
<tr>
<td></td>
<td>NorthBoundLongitude</td>
<td>see above</td>
<td>Y</td>
<td>See above</td>
</tr>
<tr>
<td>Field</td>
<td>Description</td>
<td>Default/Alternative Value</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>----------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TemporalExtent</td>
<td>Period of time that the data resource covers</td>
<td>D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DateOfPublication</td>
<td>Date on which the resource was published</td>
<td>D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DateOfLastRevision</td>
<td>Date of the last revision to the dataset, since the publication date</td>
<td>D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temporal Reference</td>
<td>Date the resource was created</td>
<td>D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lineage</td>
<td>Brief statement about the origin and revision history of the dataset</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality &amp; Validity</td>
<td>Date of the last revision to the dataset, since the publication date</td>
<td>D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SpatialResolution</td>
<td>The cartographic scale (as a denominator) that the data would represent</td>
<td>A default of 50,000 can be applied: this is the general operating scale of most of the EPA's spatial datasets</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degree</td>
<td>Conformance of resource with INSPIRE implementing rules, True or False</td>
<td>D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conformity</td>
<td>References the requirement under which conformance (above) is stated</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LimitationsOnPublicAccess</td>
<td>Limitations on access to data, related to Intellectual Property, Privacy etc</td>
<td>D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Access Constraints</td>
<td>ConditionsOfUse</td>
<td>D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ResponsibleParty</td>
<td>Contact details of organisation and party responsible for the data resource</td>
<td>D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ResponsiblePartyRole</td>
<td>The function performed by the responsible party identified above</td>
<td>D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MetadataContact</td>
<td>Point of contact for information about the metadata</td>
<td>D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MetadataDate</td>
<td>Date that the metadata was created</td>
<td>D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metadata on Metadata</td>
<td>Language used in creating the metadata file</td>
<td>D</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The end date could be set to default to the date for that day, the start date cannot be automated.

Partially: as above, the system date can be used to set a default value.

This is unique information per dataset.

A default of 50,000 can be applied: this is the general operating scale of most of the EPA's spatial datasets.

Enter a default value of “not evaluated”, which can be changed to “conformant” as the implementation of the INSPIRE Directive progresses.

This relates to the topic code and so could be captured from the data table name by applying a convention.

A default value of “Open Access to Public” can be applied, as this applies to most of the EPA’s spatial datasets.

A default value of “No conditions apply” can be applied, as this applies to most of the EPA’s spatial datasets.

A default contact can be applied to all data.

A default value of “Owner” can be applied to all data.

A default contact can be applied to all data.

The system date can be used as a default value.

All EPA metadata is created in English.
The default values are easily applied to the table at the point of table creation (see database creation scripts in Appendix 1). While many of the values have to be confirmed by the metadata author there is potential for time saving, the check action being quicker than selecting a value for the field each time (especially for the creation of multiple metadata files).

Four of the auto-populated fields can be captured from the content of the dataset: the four longitude fields. They are contained within the spatial dataset: they could be extracted via a query created and run by the user but this would not be the optimum solution for streamlining the process (see Chapter 6 for more discussion of this).

The UniqueIdentifier for the resource is the dataset name as it is stored in the database: this is already being captured via the UpdateMD_Records trigger discussed above. The remaining three auto-populated fields can be populated in relation to a naming convention that is related to the context of the data. This context will allow the values of the TopicCategory, KeywordValue and Specification fields to be predicted. Datasets that have a common theme will typically be searchable by the same keywords: for example, the EPA grants operating licenses to industries and waste facilities, so all the datasets relating to these licensed locations would have Licensed and Regulated as keywords. They may have more keywords that the metadata author can add to the list as they update the metadata, but these two additions will at least make the metadata a feasibly discoverable resource. The datasets would also belong to the same topic category, and would fall under the same data specification (of which there are only three in the lookup in any case).

Auto-population of these fields would require a set of rules that allow the data to update where certain conditions are satisfied. One mechanism for this is the use of a prefix on all table names that links to a set of keywords, and so on to topic categories and specification values. For example, applying the prefix Lic to the data relating to licensed industries and waste facilities would identify them as a particular data theme.

A new trigger can be created that copies the DBOBJECTNAME to the UniqueResID field and that uses a LIKE statement to evaluate the keyword, topic category and
specification value that go with the table name, now recorded in the UniqueResID. The syntax of this trigger (called Populate) is shown in Appendix 3. The result of this trigger is to update the four specified fields after the initial insert of the new blank metadata record (Figure 5.6).

As well as allowing for auto-population of fields (albeit allowing a low level of automation) applying a naming convention organises tables of data in logical groups: expert GIS users may not be from a traditional database or IT background and the advent of geospatial databases is a recent phenomenon. Most expert GIS users are more used to storing the data as the native GIS format known as a shapefile, so applying this kind of naming convention to make the data tables have the appearance of common groups would also help expert GIS users locate relevant tables in the database.

![Figure 5.6 Populate trigger autopopulates selected fields](image-url)

As a quick proof of concept, two metadata records were re-entered: the first in the standard database and the second after default values and the autopopulation queries
discussed above had been run. There was a time saving of 45 seconds on the second file. This is not a huge time saving but with improved autopopulation methods or in the case of a large volume of metadata creation it does represent a worthwhile time saving. It may also reduce the perception of metadata entry as a monotonous task (Mathys, 2004).
Chapter 6: Conclusions and Next Steps

This project has examined the INSPIRE metadata profile and sought to examine some of the concepts that will be involved in the implementation of this Directive by the Irish EPA. This project only examined metadata: there are many other aspects of the INSPIRE Directive that could potentially be covered by similar projects.

This chapter is a review of the concepts introduced albeit in some cases at a rudimentary level in the preceding chapters. This chapter summarises key conclusions that can be drawn from this project and suggests areas of further development of some of the concepts discussed.

Section 6.1 Discussion and Conclusions

Chapter Two was a review of the literature on metadata and metadata projects. Several authors cited the potential problems that arise in implementing metadata projects.

Conclusion One: Metadata projects require a significant investment of domain expert time, and there is potential for metadata projects to experience difficulties that lead to metadata going out of date. There is a risk of similar problems with implementing the metadata requirements of the INSPIRE Directive.

This project focussed on the INSPIRE metadata profile, and the metadata database described in Chapter Three was designed according to this profile. The review of the metadata stylesheets in Chapter Four revealed a flaw with this approach: the expert users noted that key elements of the metadata they felt were essential were missing (from the INSPIRE metadata profile and the database), namely explanation of each field of the data resource and some explicit data quality parameters.

Conclusion Two: The INSPIRE metadata profile is not completely fit for purpose in the context of environmental spatial data.

This leads to an immediate second conclusion:

Conclusion Three: Metadata profiles should be checked against actual user needs, rather than adhering only to published standards.
Once the database design (described in Chapter Three) was underway it became apparent that the data types prescribed by the INSPIRE implementing rules for metadata did not guarantee a sound database design. The frequency of free text fields in the profile had the potential to allow data inconsistencies and a large amount of data redundancy by allowing repeated large sized free text fields. As well as database design considerations, the use of free text could not be expected to lead to uniform metadata across an organisation: different users could potentially use different text (and terms) to describe essentially the same data resources. In Chapter Two Lagoze et al (2006) were cited for their description of the difficulties some domain experts had in being accustomed to creating metadata, and what the different elements of a metadata profile described. Refining free text fields to instead implement a range of possible values for a metadata element ï€ where this is practical and correct ï€ will support the metadata author in becoming familiar with metadata elements. It could viably lead to better quality metadata as an end result. Using ranges of values also supports better normalised database design. Finally, it could offer increased protection to the organisation from malicious attacks such as SQL injection, which are potentially easier to carry out on free text fields (in cases where the metadata data entry form is implemented on Internet web applications).

**Conclusion Four:** The INSPIRE metadata profile data types should be reviewed to include ranges of values rather than free text fields, where possible.

Chapter Four included a brief, low level review of three metadata stylesheets. The stylesheets were used to view the same input XML file, but each of the styles was scored differently by the users. This points to the potential value of customising the way the same input metadata is viewed by different user groups. All the users commented that the metadata files were useful ï€ or potentially useful ï€ to them.

**Conclusion Five:** Presentation of metadata can affect its value and relevance to users: metadata output format should be customised to suit the audience rather than applying a single standard approach. Metadata is viewed as a useful resource by all users.

Chapter Five examined the potential for making metadata creation and maintenance a more natural part of the overall data management workflow. The potential use of database tools such as triggers was examined. The triggers that were implemented as
part of this project were simplistic, but showed that there is potential for using them to at least help track and monitor metadata creation.

**Conclusion Six:** A metadata database should be extended beyond being just a repository for metadata: all potential database tools for implementing proactive metadata workflows should be used to their full potential.

The potential for auto population of certain metadata elements in the metadata database was briefly discussed in Chapter Five. Entry of even the default values allowed for a small time saving in the entry of metadata.

**Conclusion Seven:** Use of default values and the auto population of metadata fields - where possible - has the potential to save time in metadata creation, and reduce the monotony of metadata creation tasks.

---

**Section 6.2 Future Development**

Following on from the second conclusion (above), it is apparent that the metadata profile must be revisited in order to ensure that it meets user requirements as well as the INSPIRE Directive requirements. To this end, Table 3.1 should be revisited to include all required elements. This should be done in consultation with users, as well as by reviewing other established metadata standards and profiles outside of the INSPIRE Directive.

No metadata data entry forms were implemented for this database: they should be implemented, and be presented in a "wizard" like workflow style that groups metadata elements into logical groupings. (For an example of this, see the GIS software ESRI ArcCatalog and its metadata wizards).

A simple FOR XML statement was used to extract the metadata. This should be improved to allow better output of metadata and the potential for users to create metadata files for selected data resources on the fly. Data extraction should include a "data catalogue" page which includes the data resource titles and abstracts so that users can select more detail about the data resources they are interested in.
As a result of conclusion five (above) the XML extraction could be improved to include different stylesheet references (potentially making identical copies of the same XML with different stylesheet references depending on the output location). For example, XML files destined for the general public could apply a brief, “Frequently Asked Questions” style whereas domain experts may favour a different style. Similarly, more research is needed on actual, good quality metadata to determine the optimum styles for each audience. The stylesheet design could be combined with other web presentation formats (e.g. HTML or ASP) that would allow the stylesheets to have a “snappier” and more visually pleasing format. Hyperlinks to allow users to drill down through a full metadata profile (e.g. “See More” links) should be included to reduce data noise or clutter that could result from having the whole metadata profile appear on a single document.

Conclusion six notes that there is potential for the use of database tools such as triggers and stored procedures to assist implementation and automation of metadata workflow. The triggers discussed in Chapter Five were simplistic and there is considerable scope for improving these triggers, particularly in the case of data resource edits (the triggers used in Chapter Five favoured data resource creation). No data entry tools were created for the metadata database and there is scope for prompting forms to the user even with only a selection of the most important metadata elements in order to ensure that metadata is created and maintained.

Chapter Five also discussed the potential for automating the process of metadata creation. Conclusion seven notes this has the potential to save time and effort. There is considerable scope to improve and develop the methods outlined in Chapter Five. First of all, a stored procedure or trigger could be developed to autopopulate fields for new tables by extracting certain data (e.g. geographic extent co-ordinates) from the data resource itself. The query used for this is shown below, but is not optimum because the user must enter the table name each time:

```sql
Select e, n, w, s
INTO TMPbounds
UPDATE md_records
SET EastLong = e,
    NorthLong = n,
    SouthLong = s,
    WestLong = w
FROM TMPbounds
```
The potential to add default values could also be developed. Simple rules such as user-lead naming conventions for data resources were used in Chapter Five to examine the potential value of this concept. There is scope for improvement of these rules and also for applying more complex solutions that use the content of the underlying data resource itself to populate metadata fields: perhaps using Artificial Intelligence methods similar to those used in expert systems has potential here.

**Section 6.3 Concluding Remarks**

The conclusions and potential future developments discussed above result from examination of metadata profiles, metadata database design, metadata presentation and metadata maintenance workflow at a reasonably simple level. What is apparent particularly from the literature reviewed and discussed in Chapter Two is that metadata projects have suffered difficulties and that there is considerable risk of them failing. The INSPIRE Directive must be implemented by all EU Member States: it can be a shallow "tick the box" exercise or participating organisations in EU Member States can strive to make it more than this.

In Chapter Two it was noted that metadata is not well implemented by public service organisations in Ireland. The literature reviewed in Chapter Two pointed at some of the reasons for this. This leads to an overall conclusion that metadata projects must be planned carefully so they don’t become "tick the box" exercises that create minimal metadata in order to satisfy the requirements of the INSPIRE Directive. That all users saw the potential usefulness of metadata is encouraging, but points to the need for metadata to be a complete and accurate information resource. Once good quality metadata becomes available, its apparent value and increased use will provide the impetus for metadata to become a more inherent part of data management workflows.

This project has attempted to examine areas were threats to metadata implementation can be identified and circumvented by more design and planning for metadata presentation and data management workflows: the INSPIRE Directive must be implemented, but this should be viewed as an opportunity rather than solely an administrative overhead.
References


Appendix 1
Database Creation Scripts

CREATE TABLE dbo.Contact
    (ContactID int NOT NULL,
     ContactLName nvarchar(20) NULL,
     ContactFName nvarchar(20) NULL,
     ContactOrgID nvarchar(3) NULL,
     ContactEmail nvarchar(20) NULL,
     ContactPhone nvarchar(20) NULL)

CREATE TABLE dbo.DegreeConformance
    (ConformanceCd nvarchar(3) NULL,
     ConformanceFull nvarchar(20) NULL)

CREATE TABLE dbo.LanguageCodes
    (LanguageCd nvarchar(3) NULL,
     LanguageName nvarchar(20) NULL)

CREATE TABLE dbo.LimitationCodes
    (LimitationCd nvarchar(3) NULL,
     LimitationTypeFull nvarchar(200) NULL)

CREATE TABLE dbo.MD_Records
    (MetadataID int IDENTITY(1,1) NOT NULL,
     DBObjectName nvarchar(30) NULL,
     DBUserName nvarchar(60) NULL,
     ResTitle nvarchar(50) NULL,
     ResAbstract nvarchar(max) NULL,
     ResLocator nvarchar(100) NULL,
     UniqueResID nvarchar(100) NULL,
     CoupledRes nvarchar(100) NULL,
     ResLanguageCD nvarchar(3) NULL,
     TopicCd nvarchar(3) NULL,
     ServiceCd nvarchar(3) NULL,
     Keyword nvarchar(20) NULL,
     Thesaurus nvarchar(20) NULL,
     WestLong int NULL,
     EastLong int NULL,
     NorthLong int NULL,
     SouthLong int NULL,
     TempExtStart smalldatetime NULL,
     TempExtEnd smalldatetime NULL,
     PubDate smalldatetime NULL,
     RevDate smalldatetime NULL,
     CreateDate datetime NULL,
     Lineage nvarchar(max) NULL,
     SpResolution int NULL,
     Degree nvarchar(3) NULL,
     SpecificationCd nvarchar(3) NULL,
     LimitationCd nvarchar(3) NULL,
     ConditionCd nvarchar(3) NULL,
     RespContactID int NULL,
     RespRoleCd nvarchar(3) NULL,
     MDContactID int NULL,
     MDDate smalldatetime NULL,
     MDLanguageCd nvarchar(3) NULL)

CREATE TABLE dbo.Organisations
(OrgCD int NOT NULL,
  OrgName nvarchar(50)  NULL,
  OrgAddr1 nvarchar(50)  NULL,
  OrgAddr2 nvarchar(50)  NULL,
  OrgAddr3 nvarchar(50)  NULL,
  OrgURL nvarchar(50)  NULL)

CREATE TABLE dbo.ResourceType
  (ResourceTypeCd nvarchar(3)  NULL,
   ResourceTypeFull nvarchar(30)  NULL)

CREATE TABLE dbo.ResponsibleRoleCodes
  (RespRoleCd nvarchar(3)  NULL,
   RespRoleFull nvarchar(100)  NULL)

CREATE TABLE dbo.ServiceTypes
  (ServiceCd int NULL,
   ServiceTypeFull nvarchar(100)  NULL)

CREATE TABLE dbo.SpecificationCodes
  (SpecificationCd int NOT NULL,
   SpecificationTypeFull nvarchar(200)  NULL)

CREATE TABLE dbo.TopicCodes
  (TopicCd nvarchar(3)  NULL,
   TopicName nvarchar(50)  NULL)

CREATE TABLE dbo.UseConditionCodes
  (ConditionCd nvarchar(3)  NULL,
   ConditionTypeFull nvarchar(20)  NULL)

**Addition of Default Values**

Following the discussion on default values in Chapter 5, the scripts could be modified to add a default value at time of table creation, e.g.

CREATE TABLE dbo.MD_Records
  (MetadataID int IDENTITY(1,1) NOT NULL,
   DBObjectName nvarchar(30)  NULL,
   DBUserName nvarchar(60)  NULL,
   ResTitle nvarchar(50)  NULL,
   ResAbstract nvarchar(max)  NULL,
   ResType nvarchar(3)  DEFAULT 'DAT',
   ResLocator nvarchar(100)  DEFAULT 'maps.epa.ie',
   UniqueResID nvarchar(100)  NULL,
   CoupledRes nvarchar(100)  NULL,
   ResLanguageCD nvarchar(3)  DEFAULT 'ENG',
   TopicCd nvarchar(3)  NULL,
   ServiceCd nvarchar(3)  NULL,
   Keyword nvarchar(20)  NULL,
   Thesaurus nvarchar(20)  NULL,
   WestLong int NULL,
   EastLong int NULL,
   NorthLong int NULL,
   SouthLong int NULL,
   TempExtStart smalldatetime NULL,
   TempExtEnd smalldatetime NULL DEFAULT GETDATE(),
   PubDate smalldatetime NULL DEFAULT GETDATE(),
   RevDate smalldatetime NULL DEFAULT GETDATE(),
```
CreateDate datetime NULL DEFAULT GETDATE(),
Lineage nvarchar(max) NULL,
SpResolution int NULL DEFAULT 50000,
Degree nvarchar(3) NULL DEFAULT 'NEV',
SpecificationCd nvarchar(3) NULL,
LimitationCd nvarchar(3) NULL DEFAULT 'OPE',
ConditionCd nvarchar(3) NULL DEFAULT 'NON',
RespContactID int NULL DEFAULT 1,
RespRoleCd nvarchar(3) NULL DEFAULT 'OWN',
MDContactID int NULL DEFAULT 1,
MDDate smalldatetime NULL DEFAULT GETDATE(),
MDLanguageCd nvarchar(3) NULL DEFAULT 'ENG')
```
Appendix 2
FOR XML query
XML output sample
XSL stylesheets
Results of Stylesheet Review

1. FOR XML query
use metadata
GO
select ResTitle, ResAbstract, ResLocator, UniqueResID, CoupledRes,
(select LanguageName FROM LanguageCodes where md_records.ResLanguageCd =
LanguageCodes.LanguageCd) AS ResLanguage,
(select TopicName FROM TopicCodes where md_records.TopicCd =
TopicCodes.TopicCD) AS TopicName,
(select ServiceTypeFull FROM ServiceTypes where md_records.ServiceCd =
ServiceTypes.ServiceCd) AS ServiceType,
Keyword, Thesaurus, WestLong, EastLong, NorthLong, SouthLong, TempExtStart,
TempExtEnd, PubDate, RevDate,
CreateDate, Lineage, SpResolution,
(select ConformanceFull FROM DegreeConformance where md_records.Degree =
DegreeConformance.ConformanceCd) AS DegreeConformance,
(select SpecificationTypeFull FROM SpecificationCodes where md_records.SpecificationCd =
SpecificationCodes.SpecificationCd) AS Specification,
(select LimitationTypeFull FROM LimitationCodes where md_records.LimitationCd =
LimitationCodes.LimitationCd) AS AccessLimitation,
(select ConditionTypeFull FROM UseConditionCodes where md_records.ConditionCd =
UseConditionCodes.ConditionCd) AS ConditionOfUse,
(select ContactLName FROM Contact where md_records.RespContactID =
Contact.ContactID) AS ResponsibleContactLName,
(select ContactFName FROM Contact where md_records.RespContactID =
Contact.ContactID) AS ResponsibleContactFName,
(select ContactEmail FROM Contact where md_records.RespContactID =
Contact.ContactID) AS ResponsibleContactEmail,
(select ContactPhone FROM Contact where md_records.RespContactID =
Contact.ContactID) AS ResponsibleContactPhone,
(select OrgName FROM Organisations where Organisations.OrgCD =
Contact.ContactOrgID) AS ResponsibleOrganisationName,
(select OrgAddr1 FROM Organisations where Organisations.OrgCD =
Contact.ContactOrgID) AS ResponsibleOrganisationAddr1,
(select OrgAddr2 FROM Organisations where Organisations.OrgCD =
Contact.ContactOrgID) AS ResponsibleOrganisationAddr2,
(select OrgAddr3 FROM Organisations where Organisations.OrgCD =
Contact.ContactOrgID) AS ResponsibleOrganisationAddr3,
(select OrgURL FROM Organisations where Organisations.OrgCD =
Contact.ContactOrgID) AS ResponsibleOrganisationURL,
(select RespRoleFull FROM ResponsibleRoleCodes where md_records.RespRoleCD =
ResponsibleRoleCodes.RespRoleCD) AS ResponsibleOrganisationRole,
(select ContactLName FROM Contact where md_records.MDContactID =
Contact.ContactID) AS MetadataContactLName,
(select ContactFName FROM Contact where md_records.MDContactID = Contact.ContactID) As MetadataContactFName,
(select ContactEmail FROM Contact where md_records.MDContactID = Contact.ContactID) As MetadataContactEmail,
(select ContactPhone FROM Contact where md_records.MDContactID = Contact.ContactID) As MetadataContactPhone,
(select OrgName FROM Organisations where Organisations.OrgCD = Contact.ContactOrgID) As MetadataOrganisationName,
(select OrgAddr1 FROM Organisations where Organisations.OrgCD = Contact.ContactOrgID) As MetadataOrganisationAddr1,
(select OrgAddr2 FROM Organisations where Organisations.OrgCD = Contact.ContactOrgID) As MetadataOrganisationAddr2,
(select OrgAddr3 FROM Organisations where Organisations.OrgCD = Contact.ContactOrgID) As MetadataOrganisationAddr3,
(select OrgURL FROM Organisations where Organisations.OrgCD = Contact.ContactOrgID) As MetadataOrganisationURL,
MDDate,
(select LanguageName FROM LanguageCodes where md_records.MDLanguageCd = LanguageCodes.LanguageCd) AS MetadataLanguage
FROM md_records, Contact
WHERE ResTitle = 'soils map'
FOR XML AUTO, ELEMENTS

2. SAMPLE XML OUTPUT
<?xml version="1.0" encoding="UTF-8" ?>
<md_records>
  <ResTitle>Soils map</ResTitle>
  <ResAbstract>National soils map prepared by Teagasc, 2006</ResAbstract>
  <ResLocator>See responsible party contact details</ResLocator>
  <UniqueResID>n/a</UniqueResID>
  <CoupledRes>n/a</CoupledRes>
  <ResLanguage>English</ResLanguage>
  <TopicName>Geoscientific Information</TopicName>
  <ServiceType>Geographic viewer</ServiceType>
  <Keyword>soil</Keyword>
  <Thesaurus>n/a</Thesaurus>
  <WestLong>0</WestLong>
  <EastLong>0</EastLong>
  <NorthLong>0</NorthLong>
  <SouthLong>0</SouthLong>
  <TempExtStart>2004-03-31T00:00:00</TempExtStart>
  <TempExtEnd>2006-05-31T00:00:00</TempExtEnd>
  <PubDate>2006-05-31T00:00:00</PubDate>
  <RevDate>2006-05-31T00:00:00</RevDate>
  <CreateDate>2006-05-31T00:00:00</CreateDate>
  <Lineage>Details given in report, teagasc_class_combined.doc, available on request</Lineage>
  <SpResolution>50000</SpResolution>
  <DegreeConformance>Not Evaluated</DegreeConformance>
  <Specification>Not Evaluated</Specification>
3. XSL stylesheets
Note: The stylesheet reference had to be inserted into copies of the xml file before the root element e.g.
<?xml version="1.0" encoding="UTF-8" ?>
<?xml-stylesheet type="text/xsl" href="Stylesheet1.xsl" ?>
<md_records>
This would have to be entered into the file at the point of writing the XML from the database in order for the process to be streamlined (see Chapter 6).

Styleheet1_General.xsl

<?xml version='1.0'?><xsl:stylesheet xmlns:xsl="http://www.w3.org/TR/WD-xsl"
xmns="http://www.w3.org/TR/REC-html40" result-ns=""">
  <xsl:template match="/">
    <TABLE>
      <TR ALIGN="center" VALIGN="center">
        <TD COLSPAN="2">
          <FONT COLOR="#000000" SIZE="2" FACE="Arial">
            <TABLE>
              <TR ALIGN="center" VALIGN="center">
                <TD COLSPAN="2">
                  <FONT COLOR="blue" SIZE="2"><B>EPA Spatial Data Resources</B></FONT>
                </TD>
              </TR>
              <table width="100%" border="0" bgcolor="blue">
                <tr>
                  <td></td>
                </tr>
              </table>
              <BR/>
              <!-- Heading-->
              <TR ALIGN="center" VALIGN="center">
                <TD COLSPAN="2" FONT COLOR="blue" SIZE="2"><B>Data Resource Identification</B></TD>
              </TR>
              <BR/>
              <BR/>
              <!-- show Resource Title -->
              <xsl:if test="md_records/ResTitle[text()]">
                <TR ALIGN="center" VALIGN="center">
                  <TD COLSPAN="2" FONT COLOR="blue" SIZE="2"><B>Data Resource Title:</B>
                  <xsl:for-each select="md_records/ResTitle[text()]">
                    <xsl:value-of />
                  </xsl:for-each>
                </TD>
              </TR>
            </TABLE>
          </FONT COLOR="#000000" SIZE="2" FACE="Arial">
        </TD>
      </TR>
    </TABLE>
  </xsl:template>
</xsl:stylesheet>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Resource Classification</td>
<td>&lt;B&gt;Data Resource Classification&lt;/B&gt;</td>
</tr>
<tr>
<td>Topic Category:</td>
<td>&lt;xsl:for-each select=&quot;md_records/TopicName[text()]&quot;&gt; &lt;xsl:value-of /&gt; &lt;/xsl:for-each&gt;</td>
</tr>
<tr>
<td>Spatial Data Service Type:</td>
<td>&lt;xsl:for-each select=&quot;md_records/ServiceType[text()]&quot;&gt; &lt;xsl:value-of /&gt; &lt;/xsl:for-each&gt;</td>
</tr>
<tr>
<td>Data Resource Keywords</td>
<td>&lt;B&gt;Data Resource Keywords&lt;/B&gt;</td>
</tr>
<tr>
<td>Data Resource Location</td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>West Longitude:</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>East Longitude:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Resource Temporal Reference</td>
</tr>
<tr>
<td>----------------------------------</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Start of Temporal Range:</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>South Longitude:</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>North Longitude:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lineage:</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>Spatial Resolution:</td>
</tr>
<tr>
<td>Degree of Conformance:</td>
</tr>
<tr>
<td>Specification:</td>
</tr>
</tbody>
</table>

Data Resource Conformity

29th August 2008
<table>
<thead>
<tr>
<th>Data Resource Access Information</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Limitations on Public Access:</strong></td>
</tr>
<tr>
<td>&lt;xsl:if test=&quot;md_records/AccessLimitation[text()]&quot;&gt;</td>
</tr>
<tr>
<td>&lt;TR ALIGN=&quot;center&quot; VALIGN=&quot;center&quot;&gt;</td>
</tr>
<tr>
<td>&lt;TD COLSPAN=&quot;2&quot;&gt;</td>
</tr>
<tr>
<td>&lt;B&gt; Limitations on Public Access:&lt;/B&gt;</td>
</tr>
<tr>
<td>&lt;xsl:for-each select=&quot;md_records/AccessLimitation[text()]&quot;&gt;</td>
</tr>
<tr>
<td>&lt;xsl:value-of /&gt;</td>
</tr>
<tr>
<td>&lt;/xsl:for-each&gt;</td>
</tr>
<tr>
<td>&lt;/TD&gt;</td>
</tr>
<tr>
<td>&lt;/TR&gt;&lt;/xsl:if&gt;</td>
</tr>
<tr>
<td>&lt;xsl:if test=&quot;md_records/ConditionOfUse[text()]&quot;&gt;</td>
</tr>
<tr>
<td>&lt;TR ALIGN=&quot;center&quot; VALIGN=&quot;center&quot;&gt;</td>
</tr>
<tr>
<td>&lt;TD COLSPAN=&quot;2&quot;&gt;</td>
</tr>
<tr>
<td>&lt;B&gt; Conditions Of Use:&lt;/B&gt;</td>
</tr>
<tr>
<td>&lt;xsl:for-each select=&quot;md_records/ConditionOfUse[text()]&quot;&gt;</td>
</tr>
<tr>
<td>&lt;xsl:value-of /&gt;</td>
</tr>
<tr>
<td>&lt;/xsl:for-each&gt;</td>
</tr>
<tr>
<td>&lt;/TD&gt;</td>
</tr>
<tr>
<td>&lt;/TR&gt;&lt;/xsl:if&gt;</td>
</tr>
<tr>
<td>&lt;BR/&gt;</td>
</tr>
<tr>
<td>&lt;!-- Heading--&gt;</td>
</tr>
<tr>
<td>&lt;TR ALIGN=&quot;center&quot; VALIGN=&quot;center&quot;&gt;</td>
</tr>
<tr>
<td>&lt;TD COLSPAN=&quot;2&quot;&gt;</td>
</tr>
<tr>
<td>&lt;B&gt;Data Resource Responsible Party&lt;/B&gt;</td>
</tr>
<tr>
<td>&lt;xsl:if test=&quot;md_records/ResponsibleContactLName[text()]&quot;&gt;</td>
</tr>
<tr>
<td>&lt;TR ALIGN=&quot;center&quot; VALIGN=&quot;center&quot;&gt;</td>
</tr>
<tr>
<td>&lt;TD COLSPAN=&quot;2&quot;&gt;</td>
</tr>
<tr>
<td>&lt;B&gt; Responsible Contact Surname:&lt;/B&gt;</td>
</tr>
<tr>
<td>&lt;xsl:for-each select=&quot;md_records/ResponsibleContactLName[text()]&quot;&gt;</td>
</tr>
<tr>
<td>&lt;xsl:value-of /&gt;</td>
</tr>
<tr>
<td>&lt;/xsl:for-each&gt;</td>
</tr>
<tr>
<td>&lt;/TD&gt;</td>
</tr>
<tr>
<td>&lt;/TR&gt;&lt;/xsl:if&gt;</td>
</tr>
<tr>
<td>&lt;xsl:if test=&quot;md_records/ResponsibleContact FName[text()]&quot;&gt;</td>
</tr>
<tr>
<td>&lt;TR ALIGN=&quot;center&quot; VALIGN=&quot;center&quot;&gt;</td>
</tr>
<tr>
<td>&lt;TD COLSPAN=&quot;2&quot;&gt;</td>
</tr>
<tr>
<td>&lt;B&gt; Responsible Contact First Name: &lt;/B&gt;</td>
</tr>
<tr>
<td>&lt;xsl:for-each select=&quot;md_records/ResponsibleContact FName[text()]&quot;&gt;</td>
</tr>
<tr>
<td>&lt;xsl:value-of /&gt;</td>
</tr>
<tr>
<td>&lt;/xsl:for-each&gt;</td>
</tr>
<tr>
<td>&lt;/TD&gt;</td>
</tr>
</tbody>
</table>
| </TR></xsl:if>
<table>
<thead>
<tr>
<th>Responsible Contact FName</th>
<th>Responsible Contact E-Mail</th>
<th>Responsible Contact Phone</th>
<th>Responsible Organisation</th>
<th>Responsible Address Line 1</th>
<th>Responsible Organisation Address Line 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
<xsl:for-each select="md_records/ResponsibleOrganisationAddr2[text()]">
    <xsl:value-of />
</xsl:for-each>
</TD>
</TR><BR/>
</xsl:if>
<xsl:if test="md_records/ResponsibleOrganisationAddr3[text()]">
    <TR ALIGN="center" VALIGN="center">
        <TD COLSPAN="2">
            <B> Responsible Organisation Address Line 3:</B>
            <xsl:for-each select="md_records/ResponsibleOrganisationAddr3[text()]">
                <xsl:value-of />
            </xsl:for-each>
        </TD>
    </TR><BR/>
</xsl:if>
<xsl:if test="md_records/ResponsibleOrganisationURL[text()]">
    <TR ALIGN="center" VALIGN="center">
        <TD COLSPAN="2">
            <B> Responsible Organisation Website:</B>
            <xsl:for-each select="md_records/ResponsibleOrganisationURL[text()]">
                <xsl:value-of />
            </xsl:for-each>
        </TD>
    </TR><BR/>
</xsl:if>
<xsl:if test="md_records/ResponsibleOrganisationRole[text()]">
    <TR ALIGN="center" VALIGN="center">
        <TD COLSPAN="2">
            <B> Responsible Organisation Role:</B>
            <xsl:for-each select="md_records/ResponsibleOrganisationRole[text()]">
                <xsl:value-of />
            </xsl:for-each>
        </TD>
    </TR><BR/>
</xsl:if>

<!-- Heading -->

<TR ALIGN="center" VALIGN="center">
    <TD COLSPAN="2"><FONT COLOR="blue" SIZE="2"><B>Information about this Metadata</B></FONT></TD>
</TR>
</TR>
</TD>
</TR><BR/>
</xsl:if>

<BR/>

<xsl:if test="md_records/MetadataContactLName[text()]">
    <TR ALIGN="center" VALIGN="center">
        <TD COLSPAN="2">
            <B> Metadata Contact Surname:</B>
            <xsl:for-each select="md_records/MetadataContactLName[text()]">
                <xsl:value-of />
            </xsl:for-each>
        </TD>
    </TR>
</xsl:if>
| B | Metadata Organisation Address Line 2:
| <xsl:for-each select="md_records/MetadataOrganisationAddr2[text()]">  
| <xsl:value-of />  
| </xsl:for-each> |
| B | Metadata Organisation Address Line 3:
| <xsl:for-each select="md_records/MetadataOrganisationAddr3[text()]">  
| <xsl:value-of />  
| </xsl:for-each> |
| B | Metadata Organisation Website:
| <xsl:for-each select="md_records/MetadataOrganisationURL[text()]">  
| <xsl:value-of />  
| </xsl:for-each> |
| B | Metadata Language:
| <xsl:for-each select="md_records/MetadataLanguage[text()]">  
| <xsl:value-of />  
| </xsl:for-each> |
| B | Last Update:
|  
| <xsl:if test="md_records/MetadataOrganisationAddr2[text()]">  
| <TR ALIGN="center" VALIGN="center">  
| <TD COLSPAN="2">  
| Metadata Organisation Address Line 2:<B>  
| <xsl:for-each select="md_records/MetadataOrganisationAddr2[text()]">  
| <xsl:value-of />  
| </xsl:for-each>  
| </TD>  
| </TR><BR/>  
| </xsl:if>  
| <xsl:if test="md_records/MetadataOrganisationAddr3[text()]">  
| <TR ALIGN="center" VALIGN="center">  
| <TD COLSPAN="2">  
| Metadata Organisation Address Line 3:<B>  
| <xsl:for-each select="md_records/MetadataOrganisationAddr3[text()]">  
| <xsl:value-of />  
| </xsl:for-each>  
| </TD>  
| </TR><BR/>  
| </xsl:if>  
| <xsl:if test="md_records/MetadataOrganisationURL[text()]">  
| <TR ALIGN="center" VALIGN="center">  
| <TD COLSPAN="2">  
| Metadata Organisation Website:<B>  
| <xsl:for-each select="md_records/MetadataOrganisationURL[text()]">  
| <xsl:value-of />  
| </xsl:for-each>  
| </TD>  
| </TR><BR/>  
| </xsl:if>  
| <xsl:if test="md_records/MetadataLanguage[text()]">  
| <TR ALIGN="center" VALIGN="center">  
| <TD COLSPAN="2">  
| Metadata Language:<B>  
| <xsl:for-each select="md_records/MetadataLanguage[text()]">  
| <xsl:value-of />  
| </xsl:for-each>  
| </TD>  
| </TR><BR/>  
| </xsl:if>  
| <xsl:if test="md_records/MDDate[text()]">  
| <TR ALIGN="center" VALIGN="center">  
| <TD COLSPAN="2">  
| Last Update:<B>  
| <xsl:for-each select="md_records/MDDate[text()]">  
| <xsl:value-of />  
| </xsl:for-each>  
| </TD>  
| </TR><BR/>  
| </xsl:if>
Stylesheet2_FAQAll.xsl

<?xml version='1.0'?>
<xsl:stylesheet xmlns:xsl="http://www.w3.org/TR/WD-xsl"
    xmlns="http://www.w3.org/TR/REC-html40" result-ns=""">

    <xsl:template match="/">

        <TABLE>
            <TR ALIGN="center" VALIGN="center">
                <TD COLSPAN="2"></TD>
                <TD WIDTH="450" ALIGN="CENTER"><FONT SIZE="3" ><B>EPA Spatial Data Resources</B></FONT></TD>
            </TR>
        </TABLE>

        <table width="100%" border="0" bgcolor="blue">
            <tr>
                <td></td>
            </tr>
            <tr>
            </tr>
        </table>
        <BR/>
    <!-- Heading -->
    <TR ALIGN="center" VALIGN="center">
        <TD COLSPAN="2"><FONT COLOR="blue" SIZE="2"><B>Some Identification details for this data:</B></FONT></TD>
    </TR>
    <BR/>
    <BR/>

    <!-- show Resource Title -->
    <xsl:if test="md_records/ResTitle[text()]">
        <TR ALIGN="center" VALIGN="center">
            <TD COLSPAN="2">
                <B>What is the common name of this dataset?:</B>
            </TD>
        </TR>
    </xsl:if>
    <xsl:for-each select="md_records/ResTitle[text()]">
        <xsl:value-of />
    </xsl:for-each>

    <xsl:if test="md_records/ResAbstract[text()]">
        <TR ALIGN="center" VALIGN="center">
            <TD COLSPAN="2">
                <B>Can I have a brief description of this data?:</B>
            </TD>
        </TR>
    </xsl:if>
    <xsl:for-each select="md_records/ResAbstract[text()]">
        <xsl:value-of />
    </xsl:for-each>

    </xsl:template>
</xsl:stylesheet>
<table>
<thead>
<tr>
<th>Is there a place online that I can access this data from?</th>
<th>Resource Locator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is there a name that uniquely identifies this data?</td>
<td>Resource ID</td>
</tr>
<tr>
<td>What data is this service based on? (for web services only)</td>
<td>Coupled Resource</td>
</tr>
<tr>
<td>What language is this data written in?</td>
<td>Resource Language</td>
</tr>
</tbody>
</table>
<TR ALIGN="center" VALIGN="center">
    <TD COLSPAN="2">
        <B>Data Resource Classification</B>
    </TD>
</TR>

<xsl:if test="md_records/TopicName[text()]">
    <TR ALIGN="center" VALIGN="center">
        <TD COLSPAN="2">
            <B>What data category does this data belong to?</B>
            <xsl:for-each select="md_records/TopicName[text()]">
                <xsl:value-of />
            </xsl:for-each>
        </TD>
    </TR>
</xsl:if>

<xsl:if test="md_records/ServiceType[text()]">
    <TR ALIGN="center" VALIGN="center">
        <TD COLSPAN="2">
            <B>What type of data service is this? (for web services only)</B>
            <xsl:for-each select="md_records/ServiceType[text()]">
                <xsl:value-of />
            </xsl:for-each>
        </TD>
    </TR>
</xsl:if>

<xsl:if test="md_records/Keyword[text()]">
    <TR ALIGN="center" VALIGN="center">
        <TD COLSPAN="2">
            <B>What key words describe this data?</B>
            <xsl:for-each select="md_records/Keyword[text()]">
                <xsl:value-of />
            </xsl:for-each>
        </TD>
    </TR>
</xsl:if>
<table>
<thead>
<tr>
<th>Data Resource Location</th>
<th>WLongitude</th>
<th>ELongitude</th>
<th>NLongitude</th>
</tr>
</thead>
</table>

**Which thesaurus do these keywords come from?**

**What is the western bounding co-ordinate for this data?**

**What is the eastern bounding co-ordinate for this data?**

**What is the northern bounding co-ordinate for this data?**
<table>
<thead>
<tr>
<th>What is the southern bounding co-ordinate for this data?</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;xsl:value-of /&gt;</td>
</tr>
<tr>
<td>&lt;xsl:for-each select=&quot;md_records/SouthLon[text()]&quot; /&gt;</td>
</tr>
</tbody>
</table>

Data Resource Temporal Reference

<table>
<thead>
<tr>
<th>From what starting date is this data relevant?</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;xsl:value-of /&gt;</td>
</tr>
<tr>
<td>&lt;xsl:for-each select=&quot;md_records/TempExtStart[text()]&quot; /&gt;</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>To what end date is this data relevant?</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;xsl:value-of /&gt;</td>
</tr>
<tr>
<td>&lt;xsl:for-each select=&quot;md_records/TempExtEnd[text()]&quot; /&gt;</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Publication Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;xsl:value-of /&gt;</td>
</tr>
<tr>
<td>&lt;xsl:for-each select=&quot;md_records/PubDate[text()]&quot; /&gt;</td>
</tr>
<tr>
<td>What date was this data published?</td>
</tr>
<tr>
<td>----------------------------------</td>
</tr>
<tr>
<td>&lt;/TD&gt;&lt;BR/&gt;</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

**Data Resource Quality and Validity**

<table>
<thead>
<tr>
<th>What is the history of this dataset?</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;/TD&gt;&lt;BR/&gt;</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

29th August 2008
<table>
<thead>
<tr>
<th>What is the cartographic scale this data represents?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Does this data conform to INSPIRE?</td>
</tr>
<tr>
<td>What requirements does this dataset adhere to?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data Resource Conformity</th>
</tr>
</thead>
<tbody>
<tr>
<td>B Data Resource Conformity</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data Resource Access Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>B Data Resource Access Information</td>
</tr>
</tbody>
</table>

29th August 2008
<table>
<thead>
<tr>
<th>Are there limitations on access to this data?</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;B&gt; Are there conditions for using this data?</td>
</tr>
<tr>
<td>--------------------------------------------</td>
</tr>
<tr>
<td>Data contact Surname:</td>
</tr>
<tr>
<td>Data contact First Name:</td>
</tr>
<tr>
<td>Data contact E-Mail:</td>
</tr>
</tbody>
</table>
<TR ALIGN="center" VALIGN="center">
    <TD COLSPAN="2">
        <B>Data contact Phone:</B>
        <xsl:for-each select="md_records/ResponsibleContactPhone[text()]">
            <xsl:value-of />
        </xsl:for-each>
    </TD>
</TR>

<xsl:if test="md_records/ResponsibleOrganisationName[text()]">
    <TR ALIGN="center" VALIGN="center">
        <TD COLSPAN="2">
            <B>Data contact Organisation:</B>
            <xsl:for-each select="md_records/ResponsibleOrganisationName[text()]">
                <xsl:value-of />
            </xsl:for-each>
        </TD>
    </TR>
</xsl:if>

<xsl:if test="md_records/ResponsibleOrganisationAddr1[text()]">
    <TR ALIGN="center" VALIGN="center">
        <TD COLSPAN="2">
            <B>Data contact organisation address line 1:</B>
            <xsl:for-each select="md_records/ResponsibleOrganisationAddr1[text()]">
                <xsl:value-of />
            </xsl:for-each>
        </TD>
    </TR>
</xsl:if>

<xsl:if test="md_records/ResponsibleOrganisationAddr2[text()]">
    <TR ALIGN="center" VALIGN="center">
        <TD COLSPAN="2">
            <B>Data contact organisation address line 2:</B>
            <xsl:for-each select="md_records/ResponsibleOrganisationAddr2[text()]">
                <xsl:value-of />
            </xsl:for-each>
        </TD>
    </TR>
</xsl:if>

<xsl:if test="md_records/ResponsibleOrganisationAddr3[text()]">
    <TR ALIGN="center" VALIGN="center">
        <TD COLSPAN="2">
            <B>Data contact organisation address line 3:</B>
            <xsl:for-each select="md_records/ResponsibleOrganisationAddr3[text()]">
                <xsl:value-of />
            </xsl:for-each>
        </TD>
    </TR>
</xsl:if>
<table>
<thead>
<tr>
<th>Responsible Organisation Address</th>
<th>Responsible Organisation URL</th>
<th>Responsible Organisation Role</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Information about this Metadata

Metadata Contact Surname: 

Metadata Contact First Name:
<table>
<thead>
<tr>
<th>Metadata Contact E-Mail:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Metadata Contact Phone:</td>
<td></td>
</tr>
<tr>
<td>Metadata Organisation:</td>
<td></td>
</tr>
<tr>
<td>Metadata Address Line 1:</td>
<td></td>
</tr>
<tr>
<td>Metadata Organisation Address Line 2:</td>
<td></td>
</tr>
</tbody>
</table>
**EPA Spatial Data Resources**

<table>
<thead>
<tr>
<th>Resource Title</th>
<th>Resource Abstract</th>
<th>Resource Locator</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the common name of this dataset?</td>
<td>Can I have a brief description of this data?</td>
<td>EPA Spatial Data Resources</td>
</tr>
</tbody>
</table>

---

29th August 2008
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Is there a place online that I can access this data from?</strong></td>
<td><strong>What data category does this data belong to?</strong></td>
</tr>
<tr>
<td>![XSL:for-each select=&quot;md_records/ResLocator[text()]&quot;</td>
<td>![XSL:for-each select=&quot;md_records/TopicName[text()]&quot;</td>
</tr>
<tr>
<td>![XSL:value-of]</td>
<td>![XSL:value-of]</td>
</tr>
<tr>
<td><strong>What key words describe this data?</strong></td>
<td><strong>From what starting date is this data relevant?</strong></td>
</tr>
</tbody>
</table>
| ![XSL:for-each select="md_records/Keyword[text()]" | ![XSL:for-each select="md_records/TempExtStart[text()]"
| ![XSL:value-of] | ![XSL:value-of] |

---

**August 29th 2008**
**To what end date is this data relevant?**

<table>
<thead>
<tr>
<th>B</th>
<th>Data contact Surname</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B</th>
<th>Data contact First Name:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B</th>
<th>Data contact E-Mail:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Data contact Phone:</td>
<td>&lt;xsl:for-each select=&quot;md_records/ResponsibleContactPhone[text()]&quot;&gt; &lt;xsl:value-of /&gt; &lt;/xsl:for-each&gt;</td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>Data contact Organisation:</td>
<td>&lt;xsl:for-each select=&quot;md_records/ResponsibleOrganisationName[text()]&quot;&gt; &lt;xsl:value-of /&gt; &lt;/xsl:for-each&gt;</td>
</tr>
<tr>
<td>Data contact organisation address line 1:</td>
<td>&lt;xsl:for-each select=&quot;md_records/ResponsibleOrganisationAddr1[text()]&quot;&gt; &lt;xsl:value-of /&gt; &lt;/xsl:for-each&gt;</td>
</tr>
<tr>
<td>Data contact organisation address line 2:</td>
<td>&lt;xsl:for-each select=&quot;md_records/ResponsibleOrganisationAddr2[text()]&quot;&gt; &lt;xsl:value-of /&gt; &lt;/xsl:for-each&gt;</td>
</tr>
<tr>
<td>Data contact organisation address line 3:</td>
<td>&lt;xsl:for-each select=&quot;md_records/ResponsibleOrganisationAddr3[text()]&quot;&gt; &lt;xsl:value-of /&gt; &lt;/xsl:for-each&gt;</td>
</tr>
<tr>
<td>Organization Website:</td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td></td>
</tr>
<tr>
<td>xsl:for-each</td>
<td></td>
</tr>
<tr>
<td>xsl:if test=&quot;md_records/ResponsibleOrganisationURL[text()]&quot;</td>
<td></td>
</tr>
<tr>
<td>TR ALIGN=&quot;center&quot; VALIGN=&quot;center&quot;</td>
<td></td>
</tr>
<tr>
<td>TD COLSPAN=&quot;2&quot;</td>
<td></td>
</tr>
<tr>
<td>B Organisation Website:</td>
<td></td>
</tr>
<tr>
<td>xsl:for-each select=&quot;md_records/ResponsibleOrganisationURL[text()]&quot;</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Organisation Role:</th>
</tr>
</thead>
<tbody>
<tr>
<td>xsl:for-each</td>
</tr>
<tr>
<td>xsl:if test=&quot;md_records/ResponsibleOrganisationRole[text()]&quot;</td>
</tr>
<tr>
<td>TR ALIGN=&quot;center&quot; VALIGN=&quot;center&quot;</td>
</tr>
<tr>
<td>TD COLSPAN=&quot;2&quot;</td>
</tr>
<tr>
<td>B Organisation Role:</td>
</tr>
<tr>
<td>xsl:for-each select=&quot;md_records/ResponsibleOrganisationRole[text()]&quot;</td>
</tr>
</tbody>
</table>

When was this metadata last updated?

This metadata was last updated on:

When was this metadata last updated?
### 4. Results of Stylesheet Review

#### GIS Expert 1: Scores

<table>
<thead>
<tr>
<th>Stylesheet</th>
<th>Relevance</th>
<th>Clarity</th>
<th>Format</th>
<th>Completeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-General</td>
<td>8</td>
<td>6</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>2-FAQ_All</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>3-FAQ_Public</td>
<td>6</td>
<td>8</td>
<td>7</td>
<td>5</td>
</tr>
</tbody>
</table>

Comment: FAQ_public is too short.
Is Metadata Useful? Yes, but time consuming to create.

#### GIS Expert 2: Scores

<table>
<thead>
<tr>
<th>Stylesheet</th>
<th>Relevance</th>
<th>Clarity</th>
<th>Format</th>
<th>Completeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-General</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>2-FAQ_All</td>
<td>6</td>
<td>8</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>3-FAQ_Public</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>5</td>
</tr>
</tbody>
</table>

Comment: FAQ_public is a good view for the Internet WebGIS.
Is Metadata Useful? Yes

#### Scientist 1: Scores

<table>
<thead>
<tr>
<th>Stylesheet</th>
<th>Relevance</th>
<th>Clarity</th>
<th>Format</th>
<th>Completeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-General</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>2-FAQ_All</td>
<td>3</td>
<td>8</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>3-FAQ_Public</td>
<td>2</td>
<td>8</td>
<td>8</td>
<td>3</td>
</tr>
</tbody>
</table>

Comment: The FAQ style is clearer, I didn’t favour the field names in the other style at all. If this were scientific data there would be a lot missing: we need the field names and explanations of units of measurement used etc. These are crucial.
Is Metadata Useful? Yes, but this is too limited.

#### Scientist 2: Scores

<table>
<thead>
<tr>
<th>Stylesheet</th>
<th>Relevance</th>
<th>Clarity</th>
<th>Format</th>
<th>Completeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-General</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>2-FAQ_All</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>3-FAQ_Public</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>1</td>
</tr>
</tbody>
</table>

Comment: I like the FAQ style in terms of readability, but there is more to do and there is no explanation of the fields in the data at all.
Is Metadata Useful? Yes

#### Scientist/GIS Expert: Scores

<table>
<thead>
<tr>
<th>Stylesheet</th>
<th>Relevance</th>
<th>Clarity</th>
<th>Format</th>
<th>Completeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-General</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>2-FAQ_All</td>
<td>3</td>
<td>8</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>3-FAQ_Public</td>
<td>2</td>
<td>7</td>
<td>6</td>
<td>3</td>
</tr>
</tbody>
</table>
Comment: A list of the fields of the data is missing, and the lineage section is too general: I’d like to see users forced to enter more quality information about tests done on this data. The FAQ style is clearer.
Is Metadata Useful? Yes but this format is too short on details.

General public user 1 Scores:

<table>
<thead>
<tr>
<th>Stylesheet</th>
<th>Relevance</th>
<th>Clarity</th>
<th>Format</th>
<th>Completeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-General</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2-FAQ_Alt</td>
<td>5</td>
<td>4</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>3-FAQ_Public</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

Comment: I couldn’t follow the first sheet that well, I like the Questions and Answers style better. The third file seemed a bit short.
Is Metadata Useful? It’s good to know where the data comes from but I think the format could be better.

General public user 2 Scores:

<table>
<thead>
<tr>
<th>Stylesheet</th>
<th>Relevance</th>
<th>Clarity</th>
<th>Format</th>
<th>Completeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-General</td>
<td>2</td>
<td>6</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>2-FAQ_Alt</td>
<td>2</td>
<td>4</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>3-FAQ_Public</td>
<td>5</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
</tbody>
</table>

Comment: I just think it is enough to know where the data is coming from but I don’t think there is enough information in any of the sheets about how up to date this data is.
Is Metadata Useful? Yes.

General public user 3 Scores:

<table>
<thead>
<tr>
<th>Stylesheet</th>
<th>Relevance</th>
<th>Clarity</th>
<th>Format</th>
<th>Completeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-General</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>2-FAQ_Alt</td>
<td>5</td>
<td>7</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>3-FAQ_Public</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>

Comment: I didn’t like the first sheet, it wasn’t clear what it was telling me. I liked the third sheet, it was enough for me to know who to ring if I had a question i.e. that’s all I need.
Is Metadata Useful? I suppose: but if I knew who to ring with a question I’d be happy.
Appendix 3
Database Triggers and Queries

1. create an audit table that tracks the events

Use test_data
GO
CREATE TABLE dbo.EventLog
(EventID INT PRIMARY KEY IDENTITY,
EventInstance XML NOT NULL)

2. UpdateMD_records trigger

CREATE TRIGGER UpdateMD_Records
ON database
AFTER CREATE_TABLE
AS
INSERT INTO test_data.dbo.EventLog (EventInstance)
VALUES (EVENTDATA())
drop table dbo.NewDataObjects
Select EventInstance.value('(/ObjectName)[1]',
'nvarchar(30)') AS ObjectName,
EventInstance.value('(/LoginName)[1]',
'nvarchar(30)') AS UserName,
EventInstance.value('(/EventType)[1]',
'nvarchar(30)') AS EventType
INTO test_data.dbo.NewDataObjects
FROM test_data.dbo.EventLog
INSERT INTO metadata.dbo.MD_Records(DBObjectName, DBUserName)
select ObjectName, USerName from NewDataObjects Left Join
metadata.dbo.md_records ON md_records.DBObjectName =
NewDataObjects.ObjectName
WHERE (metadata.dbo.Md_records.DBObjectName IS Null AND ObjectName NOT
LIKE 'NewDataObjects')

3. EMailReminder trigger

USE metadata
GO
CREATE TRIGGER EmailReminder
ON metadata.dbo.md_records
AFTER INSERT
AS
SELECT DBUserName FROM MD_Records WHERE DBUserName LIKE
'EPAFLawlor' AND ResTitle IS NULL
EXEC msdb.dbo.sp_send_dbmail
@recipients = 'fiona_lawlor@epa.ie',
@body = 'You created a new table in the EPA GIS Database. There is a blank
metadata record waiting to be populated for this new table in the Metadata Database.',
@subject = 'Metadata Reminder';
4. **QryBlankMetadata**

   ```
   SELECT *
   FROM md_records
   WHERE ResTitle IS NULL AND DBUserName IS NOT NULL
   ```

5. **Populate trigger**

   ```
   use metadata
   go
   CREATE TRIGGER Populate
   ON md_records
   after insert
   AS
   Update md_records
   Set UniqueResID = DBOBJECTNAME,
       keyword = 'Licensed, Regulated',
       TopicCd = 'ENV',
       SpecificationCd = '1'
   Where DBOBJECTNAME like 'Lic%'
   ```
Glossary
Acronyms and Abbreviations used in this Thesis

ArcSDE  Arc Spatial Database Engine. "Arc" is the prefix used to identify the Environmental and Social Research Institute's (ESRI) suite of Geographical Information Software products. Arc Spatial Database Engine is a product that allows standard corporate relational database management systems (such as Microsoft SQL Server and Oracle) to handle spatial data.

DBMS  Database Management System

EC  European Commission

EU  European Union

EPA  Environmental Protection Agency: in this project, this is used to refer to the Irish Environmental Protection Agency.

FGDC  Federal Geographic Data Commission

GIS  Term used interchangeably to refer to Geographical Information Systems, Geographical Information Science and/or Geographical Information Software. In general this term relates to the management and analysis of spatial (geographical) data.

INSPIRE  Infrastructure for Spatial Information in Europe: the INSPIRE Directive is a piece of European Commission legislation that sets out requirements for the sharing and cataloguing of spatial data in Europe.

ISO  International Organisation for Standardisation

SDI  Spatial Data Infrastructure

SQL  Structured Query Language. This thesis also refers to Microsoft SQL Server 2005, which is a relational database management system.

URL  Unique resource locator

W3C  World Wide Web Consortium

WebGIS  Term used to describe a web application that provides GIS functionality, i.e. presents a map to the user that (s)he can browse and query.

XML  Extensible mark-up language

XSL  Extensible style sheet language