Improving Access to Coastal Information: Metadata in the Marine Irish Digital Atlas

L. O'Dea, N. Dwyer, V. Cummins, D. Dunne, and É. Ó Tuama
Coastal and Marine Resources Centre, University College Cork,
Naval Base, Haulbowline, Cobh, Co. Cork, Ireland

Abstract

Throughout Ireland's past, its coast has played a vital role in its culture and the livelihood of its people. The need for integrated coastal management (ICM) grows with the increase in development pressures. ICM professionals cite the lack of access to coastal information and data as being one of the primary factors that impede progress. The Marine Irish Digital Atlas (MIDA) aims to meet these needs by developing an Internet-based atlas that provides a straightforward primary means to access and view Irish coastal information using maps, text, and imagery. The atlas is centred on a web GIS system, which enables users to visualise and query spatial data from a wide variety of organisations across Ireland. Users can also view metadata to determine data quality and obtain contact information. These information, mapping, and metadata components are woven together in a user-friendly web interface. An international metadata standard and open source database have been selected in anticipation of the future trends in data sharing in Ireland and the EU, thus opening the door for future collaborations.

Introduction

Ireland's inhabitants have always had a dependence on the island's coastal resources. At the dawn of the twenty-first century there are mounting pressures on the coastal zone from a wide variety of sources such as population growth, coastal development, resource use, and climate change. There is an ever-increasing need for integrated coastal management (ICM) in order to best manage Ireland's resources and coastal environments in a way that will balance human needs with conservation of sensitive habitats. While some efforts are being made in County Development Plans and River Basin Management Plans to address coastal management issues, much work still needs to be done in order to meet this need. This requires a complex integration of involvement from a diverse cross-section of subject areas as well as organisations (Connolly *et al.,* 2001; Bartlett, 1999). ICM professionals, in Ireland and internationally, cite the lack of access to coastal information and data as being one of the primary factors that impede progress (Commission of the European Communities, 2000; Connolly *et al.,* 2001; Bartlett 1999). In light of this, some of Ireland's key data holders are making efforts to improve accessibility to spatial data that they hold. With the assistance of the Internet, this accessibility varies according to the organisation, for example providing downloadable data (e.g., The Heritage Service in the Republic of Ireland, and The Environment and Heritage Service in Northern Ireland); giving access to an organisation's metadata database (e.g., The Marine Institute); and providing an interactive web GIS site that displays some of their data holdings (e.g., The Geological Survey of Ireland). These efforts, however, solely disseminate an organisation's own datasets and do not address the many valuable datasets locked within institutions across Ireland. There is currently no single source that attempts to bring these resources together and provide a collective approach to showing what and where coastal spatial data exist for the whole of Ireland.

The Marine Irish Digital Atlas (MIDA) project came about from the recognised need in the Irish coastal management and research community for a resource that would enable people to more easily find existing coastal spatial data. With the birth and improvement of web-enabled GIS over the last 5 years came the opportunity to create an accessible Internet-based tool that provides not only a means to search existing metadata, but to also visualise and query spatial datasets themselves. This atlas will provide a starting point for people who are searching for spatial data and information related to the Irish coast. Centred around a web GIS (Figure 1), this resource will provide a means of visualising available spatial data as well

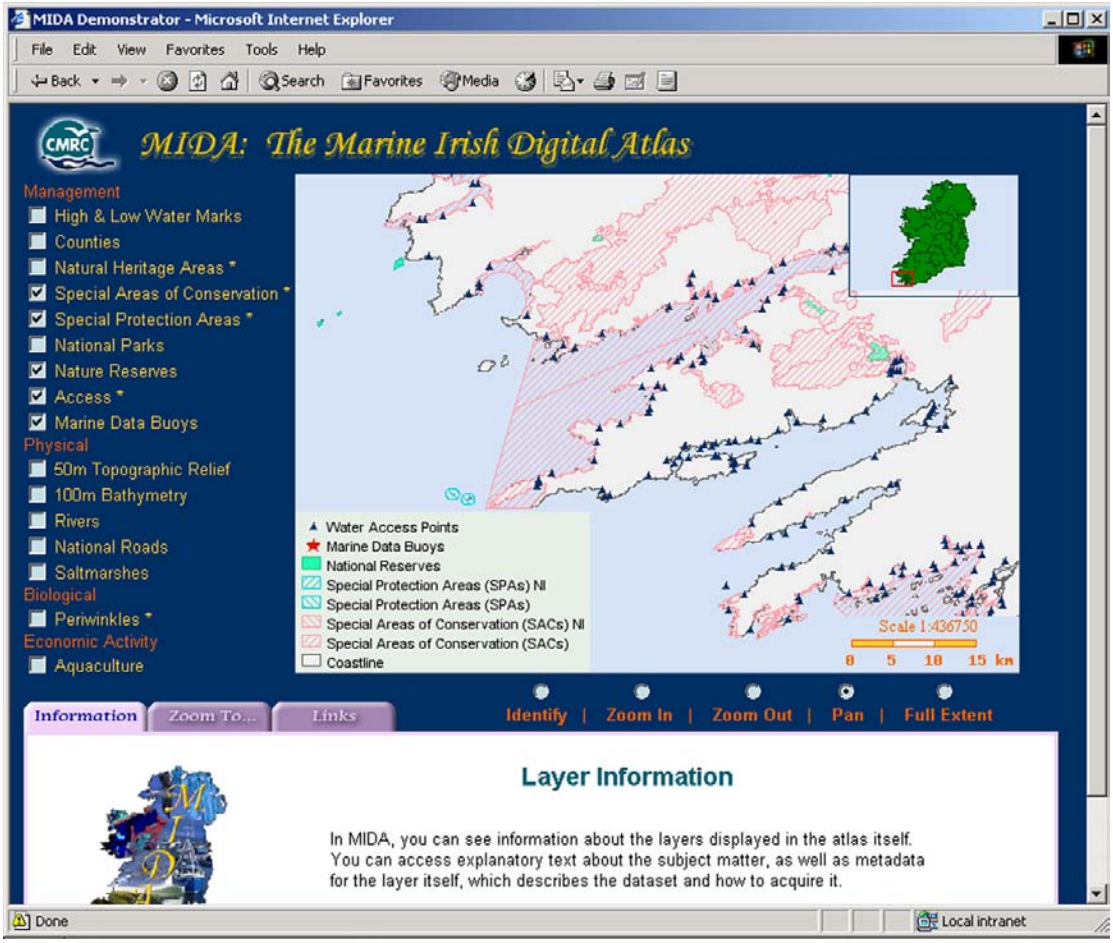

Figure 1: *The Marine Irish Digital Atlas Prototype, which allows users to select spatial data to visualise in the map frame, as well as see information about each theme displayed.*

as accessing related textual information about coastal issues and the various themes displayed (Dwyer *et al.,* 2003b).

Although not the focus of the atlas, metadata is a vital component of the MIDA. It will be available and easily accessible for all thematic layers displayed. This will enable atlas users to determine what data exist, if they will meet their needs and who to contact to acquire them. Complying with international metadata standards, an open source metadata database was designed in anticipation of the future trends in data sharing in Ireland and the EU.

This paper gives an overview of metadata, discussing standards and collection tools that were evaluated, as well as the current situation with metadata standards in Ireland. It then discusses the metadata component of the atlas, including the process of collecting and managing the metadata database, the creation of an ISO-compliant metadata input wizard, and the display and accessibility of metadata within the atlas.

Metadata Background

Metadata is the key textual information that describes a dataset, and contributes significantly to the data's value. Metadata consists of information such as when a dataset was created, who created it, as well as explanations of the dataset's contents. The SDI Cookbook, produced by a working group of the Global Spatial Data Infrastructure (GSDI), gives a good overview of what metadata is and why it is so important. The document emphasises that metadata is beneficial to data-producing organisations by providing a well documented record of their datasets, while also helping people within and outside of organisations to find spatial data that they need and determine how best to use them (GSDI, 2001). As the numbers of datasets proliferate and data exchange becomes the norm with improved digital connectivity,

there is a need to have accurate and consistent metadata. Without good quality metadata, datasets are rendered virtually useless and cannot be shared or used with confidence.

Current International Standards and Initiatives

Ideally, metadata structures and definitions should be referenced to a standard. This means that metadata should meet some sort of international quality level and should enable metadata exchange with other people and institutions with minimum effort. Historically, as institutions created geospatial datasets they created metadata, following organisational or national standards. When it came to sharing the data or even the metadata itself with other people and institutions, the shortcomings of these varying standards became apparent. This is still the case for many organisations and efforts are being made to work together to improve the situation (e.g., the Irish Spatial Data Infrastructure for governmental agencies).

Over the last few years there has been a push to try and standardise metadata for geospatial data at the international level. A few of the key programmes include the U.S. Federal Geographic Data Committee's (FGDC) Content Standard for Digital Geospatial Metadata, first created in 1994; the European Committee for Standardization (CEN) Pre-standard CEN/TC 287, adopted in 1998; the Australia and New Zealand Spatial Information Council's (ANZLIC) Australian Spatial Data Infrastructure, developed in 1996; the International Standard Organisation's (ISO) ISO 19115: Metadata Standards for Geographic Information, reformatted in 2003; and the Infrastructure for Spatial Information in Europe (INSPIRE) which is currently in development. A number of international projects are also involved in data exchange and metadata issues for non-spatial data, one of the most popular being the Dublin Core Metadata Initiative. While not a standard, it is a consensus driven definition of a set of elements that are required in order to be considered core. The initiative has been quite successful because of the simplicity of the set of elements required.

The Situation in Ireland

As of yet there is no independent Irish-defined standard for metadata. Neither is there a national policy advocating best practise for the collection and dissemination of metadata, although the Ordnance Survey of Ireland notes that, "the issue of SDI at a policy level is now well recognised at the heart of government and is explicitly mentioned in the latest Information Society Action Plan" (OSI, 2002, p.5). The Department of the Environment and Local Government has been charged with leading up the development of a policy document for an Irish Spatial Data Infrastructure (ISDI) (McCormack, 2003). However, as things currently stand, there is still an uncoordinated approach to data documentation, and individual organisations devise their own solutions (Dwyer *et al.,* 2003a). Some organisations have been very lax about metadata, while others (e.g., the Marine Institute, the Environmental Protection Agency) have done thorough research to select and work towards an appropriate standard.

IRLOGI is the umbrella organisation for the geographical information industry in Ireland, and is the nearest thing to a coordinating body that exists. Its mission is to stimulate the development and effective use of geographic information. IRLOGI's Geo-ID initiative attempts to coordinate, harmonise and present metadata from organisations in Ireland that wish to display information in the system. The metadata elements adhere to the CEN/TC 287 standard, which is being aligned to the ISO 19115 standard. About 25 public and private organisations publicise their metadata via the Geo-ID portal. However, update and maintenance of this portal has been erratic since its creation. As part of the ISDI there is interest in renewing the Geo-ID site to meet ISDI standards, this time with financial and organisational support to maintain and update the metadata database on a regular basis (Bartlett, 2003). A paper by Rybazuk *et al.* (1999) reviews the state of metadata in Ireland for IRLOGI and recommends Geo-ID as a system around which coordinated development of a national data infrastructure could develop as part of a larger European wide initiative.

There appears to be a general trend among organisations in Ireland who are implementing metadata standards. Independent research by organisations such as the Marine Institute with their data inventory, the EPA in relation with the Water Framework Directive, the Department

of the Environment in relation to the Irish Spatial Data Infrastructure, and the Coastal and Marine Resources Centre as part of the Marine Irish Digital Atlas Project have all resulted in the decision to use the ISO 19115 standard. This standard is optimal because many popular national standards such as FGDC, CEN, and ANZLIC are reformatting their standards to make them compliant with ISO 19115. These decisions point to ISO 19115 as the dominant international metadata standard, and as more organisations and SDIs work towards this standard it will become easier to share data around the world.

Selecting a Standard Profile

The number of elements required for complete metadata can be overwhelming. For this reason many organisations have found ways to work around the challenges of full metadata by creating subsets or profiles of a selected standard. A profile consists of a selection of mandatory and optional elements of a standard that meet an organisation's needs, thus making metadata easier to manage and understand. This is an internationally recognised way of adhering to a standard while not incorporating all of its elements. The SDI Cookbook provides some guidance on this by identifying three levels, namely:

> *'Discovery* metadata - What datasets hold the sort of data I am interested in? This enables organisations to know and publicise what data holdings they have.
> *Exploration* metadata - Do the identified datasets contain sufficient information to enable a sensible analysis to be made for my purposes? This is documentation to be provided with the data to ensure that others use the data correctly and wisely.
> *Exploitation* metadata – What is the process of obtaining and using the data that are required? This helps end users and provider organisations to effectively store, reuse, maintain and archive their data holdings.'
>
> - SDI Cookbook (GSDI, 2001, p.28)

Other organisations such as the National Geospatial Data Framework in the U.K. and ANZLIC in Australia and New Zealand also support the use of metadata levels. While different terminology may be used, they usually resemble the above three levels quite closely (Dwyer *et al.,* 2003a).

Metadata Collection and Tools

An organisation can collect its own metadata using one of a range of available tools, from simple text editors to more sophisticated database-connected forms. Factors that should be taken into consideration when evaluating metadata tools include whether the output format conforms rigorously to an agreed standard in order to enable easy exchange of metadata with other organisations; whether the tool provides the freedom to select a profile of metadata elements to suit needs; and whether a quick and inexpensive solution is necessary, or time, money, and resources can be spent in designing a reliable tool to meet specific needs.

While researching options for the MIDA in 2002, a number of commercial and freeware tools for the compilation of metadata were studied. The Wisconsin State Government website proved to be a useful resource (http://wisclinc.state.wi.us/metadata/mtools.html), listing 13 metadata editors and suites as well as a number of other utilities. Most of these were written to comply with the FGDC standard. At the end of 2002 three collection tools, Enraemed, M3Cat and ArcCatalog, could handle the ISO 19115 standard and were evaluated for the MIDA. All of the tools displayed both useful and bothersome features. As a result of testing those tools, an open source collection tool and database were developed within the Coastal and Marine Resources Centre (CMRC) for the collection, management and display of metadata in the MIDA (Dwyer *et al.,* 2003a).

Metadata in the MIDA

Collection and Management

Metadata forms an integral part of the MIDA by consistently documenting each dataset held. Current standards and implementations of metadata both nationally and internationally were reviewed. The ISO 19115 standard was chosen as the MIDA's metadata standard because many key organisations that handle spatial data in Ireland (e.g., Marine Institute, Environmental Protection Agency with the Water Framework Directive, The Department of the Environment with the ISDI) as well as organisations developing standards internationally (e.g., FGDC, ANZLIC, NGDF, CEN, INSPIRE) intend to adopt it. Because of the diversity and number of organisations that will be contributing datasets to the MIDA, a profile of selected metadata elements was created rather than attempt to conform to the complete ISO standard. A total of 55 fields were selected to be stored as the Discovery Metadata for each layer in the atlas (see Appendix A). This metadata is saved in Extensible Markup Language (XML) using the ISO-defined XML tags. A web-enabled data entry wizard was designed and implemented in order to ease the task of metadata collection (Figure 2). As the data contributor fills in the elements in the wizard, those elements are stored with the appropriate ISO XML tag. Through the wizard the database records can be recalled and edited if necessary. Existing records can be used as a template for entering new layers if someone is entering multiple records with similar information. This wizard is currently available as a stand-alone tool within the CMRC. Eventually it will be available from the MIDA homepage and be publicly available to those who wish to provide metadata. The data contributor will provide full metadata as a file with the dataset, if it exists.

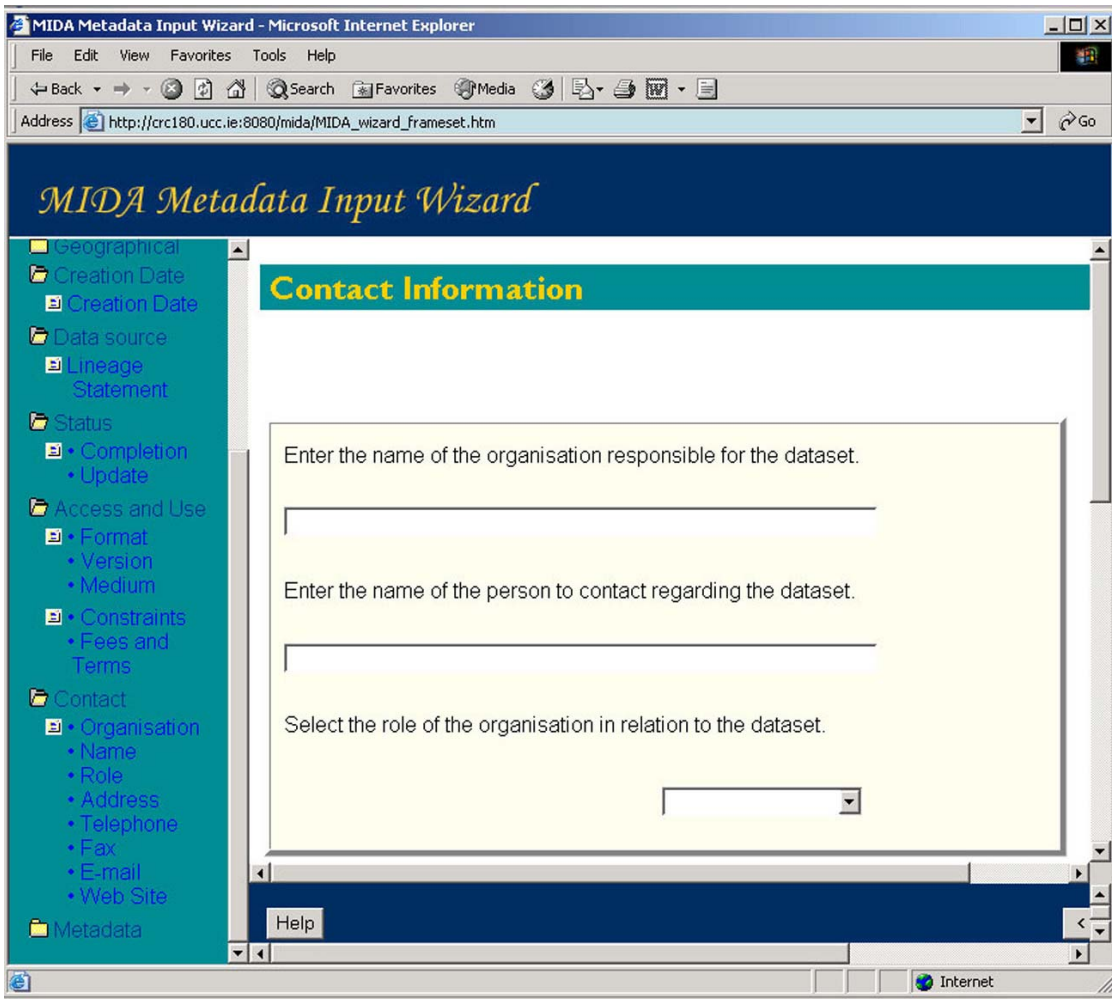

Figure 2: *The MIDA Metadata Input Wizard is web-enabled and stores metadata in ISO compliant XML format.*

ESRI's ArcCatalog is used to help the MIDA team with management of the datasets used in the atlas. This system is separate from the Discovery Metadata collected and is displayed in the atlas itself. ArcCatalog provides a way of recording details vital towards data management and update, such as when and how data from external organisations were acquired and what modifications were made in order to display those datasets in the atlas. Currently only a limited profile of ISO elements is available in the ArcCatalog Metadata Wizard, which will improve in future versions of ArcGIS. Because of the completeness of FGDC elements in ArcCatalog, CMRC metadata is recorded in FGDC format. ESRI is developing a translator tool that will easily transform full FGDC metadata into full ISO metadata.

A separate XML database will be used to store the Discovery Metadata, which will be the metadata accessible via the MIDA website. XML was chosen as the appropriate technology because it is an open standard, platform independent and easily exchangable with other organisations. The XML tags that store each metadata element are compliant with ISO standards, thus enabling easy exchange with organisations using the same standard. Metadata actually displayed and queried in the MIDA will be supplied by this XML database (Dwyer *et al.,* 2003a).

Presentation and Accessibility

In addition to reviewing national and international metadata standards, a number of online atlases were reviewed with their metadata presentation in mind. Various examples of metadata profiles that are used internationally were evaluated and referenced to determine what would be best for the MIDA. As a result of this information gathering exercise, three levels of metadata were chosen to make available in the MIDA:

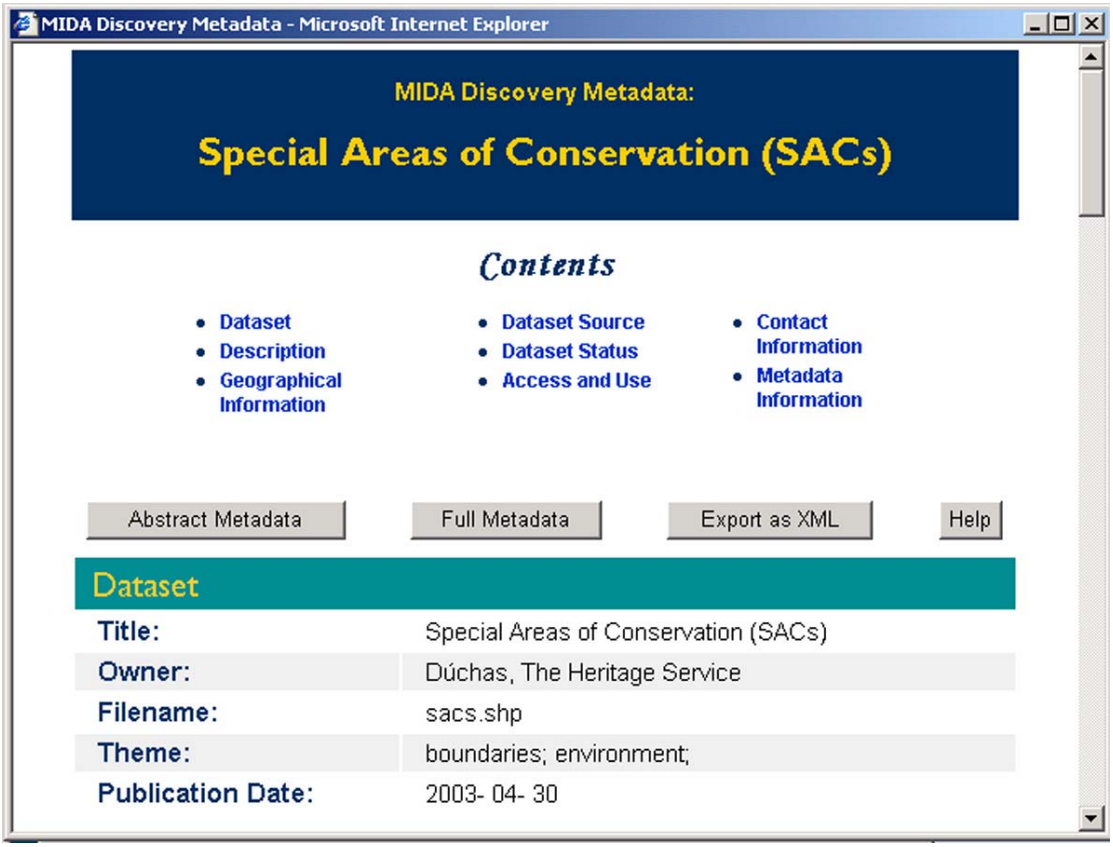

Figure 3: *Three levels of metadata are accessible in the atlas in order to meet the needs of a diverse user group. Discovery Metadata makes up the core database.*

- *Abstract Metadata:* This consists of a simple description or explanation of each dataset. It is the first level that a user sees when they choose to view the

metadata for a particular layer. When acessed, the Abstract Metadata is simply extracted from the Discovery Metadata database and displayed. There is a link that the user can select if they wish to see more information in the Discovery Metadata.

- *Discovery Metadata:* This contains the main metadata fields that are vital in defining each dataset (Figure 3). It serves as the backbone to the atlas and will be stored as XML in a database, which is simple to display and query online. Appendix A lists and defines the elements that form the Discovery Metadata within the atlas. All datasets displayed in the MIDA will have this Discovery Metadata. If the user wishes to see more information, a link to Full Metadata is available.
- *Full Metadata:* This file displayed in the atlas is supplied by the data owner who contributes the dataset. Its availability depends on whether or not it exists, and its quality is the responsibility of the data owner. Due to the variability of metadata among data owners, Full Metadata may meet a variety of standards, if any at all. The metadata is displayed as the owner provides it, with slight modifications for layout purposes, and is not stored in the MIDA XML database (Dwyer et al., 2003a).

Metadata is accessible from several places within in the atlas. The primary path is from the main atlas page (Figure 1). When a layer is selected to view in the atlas, relevant information and links appear in the information pane below the map (Figure 4), including access to the metadata relevant to the layer displayed. Elsewhere in the MIDA website, the database will be available for anyone to view, query and download Discovery Metadata for layers displayed in the atlas. Query results will provide not only links to the three levels of metadata, but also a link to the main atlas page showing the selected layer in the map.

Figure 4: *Metadata is accessible to users in multiple locations in the atlas, including the information pane of the main atlas page when a layer is selected.*

Conclusions

It is important that any efforts made to collect data from disparate sources, such as the MIDA, be developed in a way that is open to the possibility of data sharing. With technological advances in computer networks and open source technology, the potential for sharing data internationally through distributed networks is becoming a reality (Ó Tuama, 2003). There is good potential for this within the international coastal and marine community, as well as within relevant European environmental initiatives. Widely adopted international standards like ISO 19115 will make these efforts more seamless and will eliminate one of the major hurdles that prevent this type of data sharing becoming a reality.

As Longhorn (2003) states, providing access to data over the web does not solve problems, it simply gives access to data of varying quality and currency. What the MIDA aims to do is improve access to Irish coastal data by providing an easily accessible primary means of seeing what data exist. The ability for users to visualise and query the data via the web GIS

as well as metadata in the database allow them to make their own decisions on data quality and relevance to their purposes. It is also hoped that MIDA will allow managers to identify where gaps in spatial data exist, where data quality needs to be enhanced, and where energies should be focused to improve Ireland's coastal data.

Acknowledgements

This work is funded for the period September 2002 – August 2005 by the Higher Education Authority (HEA) of Ireland under the PRTLI 3 Programme as part of the National Development Plan 2002 – 2006.

References

Australia and New Zealand Spatial Information Council (ANZLIC). *Australian Spatial Data Infrastructure.* <http://www.anzlic.org.au/infrastructure.html>.

Bartlett, D.J. 1999. Working on the Frontiers of Science: Applying GIS to the Coastal Zone. In *Marine and Coastal Geographical Systems*, edited by Wright, D.J. and Bartlett, D.J. Taylor & Francis: London, pp.11-24.

Bartlett, D.J. 2003. Personal Communication.

Commission of the European Communities. 2000. Communication to the Council and the European Parliament on "*Integrated Coastal Zone Management: A Strategy for Europe*." COM(2000) 547 of 17 September, Brussels: European Commission.

Connolly, N., Buchanan, C., O'Connell, M., O'Mahony, C., Kay, D. and Buckley, S. 2001. *Assessment of Human Activity in the Coastal Zone.* Maritime Ireland/Wales INTERREG Report No. 9. Marine Institute, Dublin.

The Dublin Core Data Initiative. <http://dublincore.org/>.

Dwyer, N., O'Dea, L., and Cummins, V. 2003a. Work Package 2: Metadata Review and Recommendations in the MIDA.

Dwyer, N., O'Dea, L., and Cummins, V. 2003b. *The Marine Irish Digital Atlas: a Web Portal to Coastal and Marine Data in Ireland*. Proceedings from CoastGIS 2003 in Genoa, Italy, October. <http://mida.ucc.ie/assets/documents/MIDA_CoastGIS2003.pdf>

Environment and Heritage Service. <http://www.ehsni.gov.uk/>.

European Committee for Standardization (CEN). *Pre-standard CEN/TC 287*. <http://www.cenorm.be/cenorm/index.htm>.

Federal Geographic Data Committee (FGDC). *Content Standard for Digital Geospatial Metadata*. <http://www.fgdc.gov/metadata/contstan.html>.

Geo-ID: The Geo-Spatial Information Directory. <http://www.gis-ireland.com>.

The Geological Survey of Ireland. <http://www.gsi.ie/>.

Global Spatial Data Infrastructure (GSDI). 2001. *The SDI Cookbook, Version 1.1*. Ed. D.D. Nebert. <www.gsdi.org/pubs/cookbook/cookbook0515.pdf>.

The Heritage Service. <http://www.heritagedata.ie/>.

Infrastructure for Spatial Information in Europe (INSPIRE). <http://inspire.jrc.it/home.html>.

International Standard Organisation (ISO). *ISO 19115: Metadata Standards for Geographic Information*. <http://www.iso.org>.

Longhorn, R. 2003. *Coastal Spatial Data Infrastructure as Part of National/Regional SDI*. Proceedings from CoastGIS 2003 in Genoa, Italy, October. <http://www.gisig.it/coastgis/papers/Longhorn.htm>.

The Marine Institute. <http://www.maps.marine.ie/>.

McCormack, B. 2003. *Irish Spatial Data Infrastructure*. Proceedings from GIS Ireland 2003 in Dublin, 14 October. <http://www.irlogi.ie/pdf/ISDI_IRLOGI_2003.pdf>.

Ordnance Survey Ireland. May 2002. *Developing Ireland's National Mapping – A Progress Report*. <http://www.osi.ie/pdf/infopaperupdatemay02.pdf>.

Ó Tuama, É., Best, C., and Hamre, T. 2003. *A Web-Based Distributed Architecture for Coastal Zone Management.* Proceedings from CoastGIS 2003 in Genoa, Italy, October. <http://www.gisig.it/coastgis/papers/o tuama.htm>.

Rybazuk, K., Ovington, M., and Prendergast, W.P. 1999. *Towards a Metadata Base of Digital Spatial Data for Ireland.* <http://indigo.ie/~issprs/pdfs/si16_4.pdf>.

** Websites were last accessed 29 April 2004.*

Appendix A. MIDA Discovery Metadata Elements

This table displays all of the Discovery metadata elements selected to make up the MIDA metadata database. These 55 elements were selected from the ISO 19115 spatial metadata standards. Elements with an asterisk (*) are not part of the ISO standard and will be defined as part of the profile in the implementation. Elements shown in italics are of fixed value and not input by the user. They are included to meet ISO standard compatibility rules. *Note: These elements were taken from the draft ISO 19115 elements so may vary slightly from the final version, released in May 2003.*

Shortname	Definition
Dataset	
ResTitle	Name by which the cited resource is known
ResAltTitle	Short name or other language name by which the cited information is known
IdAbs	Brief narrative summary of the content of the resource
ResRefDate	Reference date for the cited resource
RefDateType	Event used for reference date
Description	
TpCat	Main theme(s) of the dataset
IdStatus	Status of the resource
Keyword	Commonly used word… used to describe the subject
GraphOver	A graphic that illustrates the resource
MaintFreq	Frequency with which changes and additions are made to the resource after the initial resource is completed
DataLang	*Language used within the dataset*
DataChar	*Full name of the character coding standard used for the dataset*
EquScale	Level of detail expressed as the scale of a comparable map or chart
ScaleDist	Ground sample distance
SpatRpType	Method used to spatially represent geographic information
Geographic Extent	
GeoDesc	Description of the geographic area within which data is available
ExDesc	Spatial and temporal extent for the referring object
RefSysId	Name of reference system used
WestBL	Western-most coordinate of the limit of the dataset extent, expressed in longitude in decimal degrees (positive east)
EastBL	Eastern-most coordinate of the limit of the dataset extent, expressed in longitude in decimal degrees (positive east)
southBL	Southern-most coordinate of the limit of the dataset extent, expressed in latitude in decimal degrees (positive north)
northBL	Northern-most coordinate of the limit of the dataset extent, expressed in latitude in decimal degrees (positive north)
*WestEx	Western-most coordinate of the limit of the dataset extent, expressed in local coordinates
*EastEx	Eastern-most coordinate of the limit of the dataset extent, expressed in local coordinates
*SouthEx	Southern-most coordinate of the limit of the dataset extent, expressed in local coordinates
*NorthEx	Northern-most coordinate of the limit of the dataset extent, expressed in local coordinates
Data Creation Date	
TM_CalDate	Date for the content of the dataset
Data Source	
DqScope	*The specific data to which the data quality information applies*
Lineage Statement	General explanation of the data producer's knowledge about the lineage of a dataset
Access and Use	
FormatName	Name of the data transfer format
FormatVer	Version of the format (date, number, etc)
MedName	Name of the medium on which the resource can be received
ResFes	Fees and terms for retrieving the resource, including monetary units
OthConsts	Other restrictions and legal prerequisites for accessing and using the resource
Contact Information	
RpIndName	Name of the responsible person – surname, given name, title separated by a delimiter
RpOrgName	Name of the responsible organisation
Role	Function performed by the responsible party
DelPoint	Address line for the location
City	City of the location

AdminArea	County
Country	Country of the physical address
VoiceNum	Telephone number
FaxNum	Fax number
EMailAdd	Email
Linkage / orName / orDesc	Location for on-line access to additional info.
Metadata Information	
MdFileID	Unique identifier for this metadata file
MdLang	*Language used for documenting metadata*
MdChar	*Full name of the character coding standard used for the metadata set*
MdDateSt	Date that the metadata was created
MdStanName	*Name of the metadata standard (including profile name used)*
mdStanVer	*Version (profile) of the metadata standard used*
RpOrgName	Name of the responsible organisation for metadata
Role	Function performed by the responsible party
SuppInfo	Supplementary information regarding the dataset
CI_Contact – OnLine Resource	On line linkage to further information