# Support to Building the Inter-American Biodiversity Information Network

**Trust Fund #TF-030388**

# Recommended Standards and Practices for sharing of GIS-based Information

**(Document 4)**

**July 2004**

**Support to Building IABIN (Inter-American Biodiversity Information Network) Project**

**Recommended Standards and Practices for sharing of GIS-based Information**

## Project Background

The World Bank has financed this work under a trust fund from the Government of Japan. The objective is to assist the World Bank in the completion of project preparation for the project 'Building Inter-American Biodiversity Information Network (IABIN)' and for assistance in supervision of the project. The work undertaken covers three areas: background studies on key aspects of biodiversity informatics; direct assistance to the World Bank in project preparation; and assistance to the World Bank in project supervision. The current document is one of the background studies.

The work has been carried out by Nippon Koei UK and the UNEP World Conservation Monitoring Centre.

**Table of Contents**

## Tables

## Report Summary

This report aims to highlight the main issues involved in managing and sharing Geographical Information Systems (GIS) based information. Initially, a brief description of GIS is given, with some of the issues and implications highlighted, with special reference given to data types and projections. This is then linked into an overview of Spatial Data Standards and their development, in terms of exchange interoperability and metadata standards. By considering and deploying these standards when designing a system to manage and disseminate GIS-based information, it is possible to create a framework that is both user-friendly and transparent to the outside world. This adds value to the data provided, as the user is quickly able to identify the dataset of relevance and include it in their own GIS packages. Particular attention should be paid to this when creating GIS systems that will rely on distributed data providers for their spatial information, as it will assist in compiling that data with a reduced workload on the central GIS manager.

# CHAPTER 1 INTRODUCTION

## 1.1 Introduction to GIS

GIS (Geographical Information Systems) is the name given to various computer systems available to collate, manage, analyse and present information with a geographical (locational) component.

### 1.1.1 Data Formats

GIS is able to store location information in one of two formats; Raster and Vector.

**Raster** data is stored as cells for a given area of the world; the best example of this is satellite imagery. Within Raster data, an area of the globe is characterised as a rectangle split into equal area squares (cells), each with a given value, depending on what is found at that location. For example, a raster layer of percentage forest canopy cover contains cells with a value of between 0 and 100 representing the average forest canopy cover for each cell. Raster data can be used to store information such as the altitude, or the vegetation type at a location. Raster data is also used to produce and display the outputs of modelling analysis. One such use is climate change predictions, where, for example, an entire area is analysed to produce values for its temperature, precipitation or sea-level change, under certain climatic conditions.

**Vector** data stores information as entities in the form of points, lines or polygons. This data is made up of nodes and lines joining those nodes.

Points represent a single node at a given location on the globe, with attribute information about that location attached. These features can be used to identify various types of information such as species observation locations taken from GPS receivers or the location of cities within a country.

Line features are two or more nodes joined together by connecting straight lines. Each feature has attribute information describing what that line is. Examples of line data can be rivers, roads or railways.

Polygon features can be described as special line features. They have three or more nodes joined together by connecting straight lines where the first node is at the same location as the last node. Polygon features can be used to display anything with an area such as Protected Area boundaries or city limits.

The above examples of points, lines and polygons are just a fraction of what can be represented by vector information. As well as this, features could be identified by all three types, depending on the presentation and analysis needs of the user.

For example, cities that are identified as points at a global or regional scale can be represented as polygons at a country or local level.

The way in which the data is stored has implications on the analysis that can be undertaken on those features, for example, it is possible to calculate the area and perimeter of individual polygons, whereas points and lines do not have this attribute.

### 1.1.2 Map Projections

Another issue that needs to be addressed when storing and analysing GIS data is the projection in which it is stored. A map projection is a mathematical formula used to convert the three-dimensional surface of the earth into a two-dimensional surface, such as a map. Because the earth's surface is curved, it must be squeezed or stretched to fit into the same area on a map's flat surface. The projection process always distorts shape, area, distance, and/or direction. There is a myriad of data projections for displaying and analysing GIS data, depending upon the location and analysis requirements of that data. If the user requires statistics and analysis in relation to shape, area, distance or direction of features, then they need to present the data in a projection that provides them with the best approximation of these values. All calculations in a GIS are an approximation of actual reality due to the shape and distortions of the globe. However, by being aware of the issues and using the most appropriate projection available, it is possible to reduce the error when undertaking analysis. More information on projections is available from the ESRI website at:

http://mapshop.esri.com/help/index.htm#concepts_projections.htm

## 1.2 Spatial Data Standards

Spatial data standards provide the architecture for the integration of GIS data from multiple sources and in multiple data formats, along with providing the ability to more fully integrate spatial data with non-spatial data. The ability to integrate data is as important within an organisation as it is between different organisations.

As the pervasiveness of GIS has grown within organisations over the past 30 years, the development of spatial data standards has had to grow with it to allow for the full integration of GIS data with other data holdings. The expansion of GIS data standards has developed in the three main ways listed below, all of which are inter-related but can be analysed separately.

### 1.2.1 GIS information exchange standards

These standards allow for the exchange of data between different organisations or departments, so that the same information can be viewed with other data held on different software platforms and in different formats. To fulfil the need to distribute data to individuals and organisations working on different software packages, it has been important to develop standards so that the data structures are easily readable.

### 1.2.2 GIS information inter-operability standards

These standards allow for the integration of data held in an organisation or department, with other data held in a different organisation or department, which may or may not be working on the same GIS software platform. The data would not be moved away from the owner, but would be linked interactively through either the organisation's internal network or through the internet.

### 1.2.3 Spatial metadata standards

These standards allow for the dissemination of metadata on GIS data, both internal and external to an organisation. Metadata standards allow the data holdings of an organisation to be discovered, interrogated and incorporated within other networks.

## 1.3 GIS as a tool for linking datasets in a meaningful way

The value of GIS lies in the fact that it can be used to create meaningful outputs of the relationships between spatial information. Principally, these outputs are maps, but can often be statistics, graphs or tables. When presenting large amounts of data in reports, books, papers or posters, it is often more effective to include a map or graph to get the message across quickly and effectively. Where information has a geographical component, it is possible to display that data on a map. The information does not necessarily have to be geographical in content in order for it to be displayed cartographically, for example the signatories to a convention can be better viewed when countries are shaded on a map than listed in a long table.

When designing databases and other methods of storing information, it is important to consider possible linkages to other datasets held internally and externally to an organisation; this relates to non-spatial as well as GIS data. Various data standards have been designed to allow these linkages, such as the ISO ALPHA-3 digit country codes, although as data on specific subjects is not always covered by international standards, it is important to consult with possible collaborators and institutions to assess the best way to proceed.

# CHAPTER 2 OVERVIEW OF EXISTING STANDARDS AND PRACTICES

## 2.1 GIS information exchange standards

Data needs to be transferable between different GIS software packages to be truly shareable. Since the early stages in the development of GIS as a concept and process, methods of converting data from one system to another has been a priority. Initially, the inter-operability took the form of specific extensions to GIS software packages that allow the users to convert data from one file format to another. ESRI ArcGIS 8.3 (ArcINFO version) now supports data conversion from over 15 different file formats.

Although data converters allow GIS data to be viewed and used within different software packages, they still add a level of complexity that slows down inter-operability. To bypass this problem, various standard interchange and open file formats were developed, which allows data to be viewable within multiple software systems without the need for data converters. Probably the most significant of these file formats is GML (Geography Markup Language). GML is a specialised form of XML (Extensible Markup Language), which uses custom nested tags to define different properties of a layer. In simple terms, a tag is used to open and close a layer, within that layer there are tags which define each record (individual points, lines or polygons) and within each record there are tags which define the attributes relating to that record. GML has developed with a set of OGC (Open GIS Consortium) standards for the names of each individual tag that allows layers stored in GML to be viewed through multiple systems.

Standard interchange formats can be used to provide data to, as well as distributing it from, the IABIN system. However, as with any data transfer and distribution system, care needs to be taken with regard to maintaining the accuracy and currency of data provided. If data supplied is held locally on the IABIN system but maintained on the local systems of the contributing organisations, then mechanisms and processes need to be in place to allow data to be constantly updated with the minimal effort on both parts. If this doesn’t occur, data held within the IABIN system will loose accuracy and therefore usefulness. The same problem also occurs if data is distributed in this way, as the end-users will always be working on a dataset that is held on their local system, that is separate from that held within the IABIN system, and is therefore out of date.

## 2.2 Interoperability standards and practices

Although the information exchange methods mentioned in the previous section allow for a level of interoperability, they are never truly interoperable due to the fact that the user is always working with a static copy of the original data.

With increased access to the Internet and improved connection speeds it has been possible to develop software systems to display and manage GIS data over the web, allowing organisations to be truly interoperable, both internally and externally.

Software specific interoperability has developed to allow users of the same GIS software to have access to the same data over the web. An example of this is ESRI's ArcSDE, which is used to manage and control access to GIS data stored on a database such as SQL Server or Oracle. Data managed in this form can then be connected to, viewed and edited through other ESRI products such as ArcGIS, ArcExplorer or ArcView from anywhere connected to the Internet. A database manager controls connection rights to the data and individual users can be given rights to either view the data or have edit access on a user specific basis.

Oracle has taken this a step further with their Oracle Spatial software. Oracle Spatial allows the manager to store GIS data in an Oracle database and allow access from multiple different GIS software types. The data can also be accessed using GML for integration with GIS software that is not intrinsically supported.

The development of web-based technologies has allowed for the evolution of Interactive Mapping Services (IMS). These developments allow the user to view GIS data in an Internet browser environment. In its simplest form, an IMS creates and displays a map as an image to the user's specifications. However, it can also allow the user to change the extent of the map, view different layers at the same time, and query the underlying data stored with the mapped features. Information on the IMS technologies employed within UNEP-WCMC can be viewed in report 2.10 'UNEP-WCMC Interactive Map Service'.

As this technology has developed, it has become important to create standards, so that it is possible for these IMSs to interact with each other and with other GIS packages; the OGC has been instrumental in this process. Today there are two different standards developed for the interoperability of IMSs. These are standards for WMS (Web Map Services) and WFS (Web Feature Services); they relate to the Internet Map Server and Internet Data Server respectively. Both the Internet Map Server and Internet Data Server are explained fully in section 3.2.

These standards provide a uniform language for the sending and receiving of requests from a Map Server. These requests are the backbone of an IMS as they control what is displayed in the final map.

## 2.3 Spatial Metadata Standards

To provide full interoperability between spatial data it is important to provide the user with as much information as possible about the underlying data structure and the systems on which it is managed and displayed. Metadata is simply data about data and can be constructed for multiple information sources including published resources and Web pages, as well as GIS data. Standards for metadata have developed in line with the advances in technology and data formats. Today, a wide range of metadata standards is available covering different data types and user needs. However ISO have developed a standard called ISO 19115 that has combined attributes from several other metadata standards to create a universal standard for the storage and distribution of metadata.

Metadata for spatial data is used to enhance the interoperability of that data by allowing the user to search on-line databases and clearing-houses for individual layers and GIS systems. For example, the FGDC (Federal Geographic Data Committee) has developed an on-line geographical information clearing-house that allows the user to interrogate the metadata catalogues of various different organisations to provide them with information on the sources of specific GIS layers. This facility can search on various themes and keywords, geographical extents, and individual sources, allowing the users to specify precisely what they are looking for. Where available, this metadata also allows the users to incorporate the layers of interest into a WMS, allowing them to view the data with that of other organisations as well as with background layers. The metadata also provides the user with information on the ownership of the data and any copyright issues.

Other examples of metadata catalogues are listed in the table below

Table 1 Examples of Metadata Catalogues

| Metadata Catalogue | URL |
|---|---|
| FGDC | http://www.fgdc.gov/clearinghouse/clearinghouse.html |
| UNEP.Net Search engine | http://www.unep.net |
| FAO GeoNetwork | http://www.fao.org/geonetwork/srv/en/main.search |

Metadata can be stored in many different ways, such as in databases or as static pages. However, to be fully interactive and able to be interrogated from other systems, it needs to comply with one of the standards. These standards control the output of the information in the form of XML. The XML used to define metadata uses nested tags in the same way as the GML explained in section 2.1. However, each page in metadata XML is used to define a different data source. Within each metadata page, there are various separate tags that define different elements of that data source.

## 2.4 Multiple Sources

The very nature of IABIN and information networks in general requires the input of data obtained from multiple sources. As mentioned before, it is important to discuss with collaborating organisations when deciding on the standards and practices to employ for the integration and use of GIS-based information. Different organisations will have different ways of storing and distributing their data. To encourage them to provide that data to IABIN, standards and best practices need to be put in place to enable them to easily integrate their information within the IABIN system.

Different systems will need to be put in place depending upon the level of integration required to fulfil the aims and objectives of IABIN. The level of integration required depends on how it is envisaged that the end users will make use of the GIS data available.

There are two main ways of integrating GIS data from multiple data sources into the IABIN site, as follows:

Firstly and most simply, this can be as a link from the IABIN website through to the download page or map on the collaborating organisation's site. This level of integration can be achieved by querying metadata held elsewhere through a clearing-house held on the IABIN site; this would allow the user to search for information on a specific set of data within the IABIN website and link through to that data. This method of integration has advantages in that the data owners remain in control of their data as, if any changes and improvements are made, the link from the site is automatically linking through to the most up-to-date data. However, this requires the collaborating organisation to have a website, a metadata catalogue and the ability to distribute data and maps over the web. As well as this, the data will not be integrated with any other datasets held in IABIN or other collaborating organisations and will therefore only cover that organisation's area of interest.

The second way in which the GIS data can be integrated, is to store it centrally in the IABIN system. This has advantages in that the IABIN data managers would have full control of how the data is presented, where it is stored and from where it is retrieved. However, as the data is moved away from the original data owners, it is more likely to become out of date, unless there are agreements and mechanisms for the owners to up-date their data on the IABIN system.

# CHAPTER 3 GIS OVER THE INTERNET

## 3.1 OpenGIS (OGC)

The Open GIS Consortium, Inc. is a consortium of companies, government agencies and universities participating in a process to develop publicly available interface specifications. OpenGIS specifications support interoperable solutions that enable users to make better use of geographical information through the Web and mainstream information technologies. The specifications allow technology developers to make complex spatial information and services accessible and useful with all kinds of applications.

Through the development of standards for GML, WMS and WFS, the OGC has improved the interoperability of GIS information over the web. These standards allow developers and users to better distribute their data to other organisations and the general public through the internet and bespoke applications; an example of this is the FGDC Clearinghouse.

## 3.2 Internet map and data servers

Internet map and data servers define the way in which GIS information is passed to the user over the web.

There are two ways in which users can gain access to GIS data over the Internet: either through an Internet Map Server or an Internet Data Server.

The Internet Map Server relies on the server to undertake all the processing and map creation. In this situation, the user sends a request over the Internet to the server. This request contains information on the extent, layers and display properties needed for the map. The server then collates the data needed for the map and creates an image using the information from the request. This image is then placed in a location where the user can access it. A response is then sent to the users telling them where the image can be obtained.

An Internet Data Server relies on the user to create the map. A request is sent from the user to the server containing information on the extent and layers to be displayed in the map. The server then collates the data needed to create the map and sends it as a response through to the user. The Internet software on the user's computer then creates the map with defined colours, extents and layers.

There are various criteria for choosing either the Internet Map server or Internet Data server for delivering maps over the web. These largely depend on the speed of map delivery and the amount of functionality the user needs. Internet Map Servers provide a higher speed of map delivery to the client in the initial stages, as all that is being transferred over the web is small map requests and responses, with the bulk of the processing being undertaken on the client's machine. However, there is a limit to the functionality that is allowed through the service, with the user only able to interrogate and display the data. With an Internet Data Server the download of information is greater and ultimately relies on the speed of the user's computer to display the maps. However, as the data is processed locally, the user has a greater degree of control over functionality. This allows users to edit the data and perform more complicated queries and GIS data manipulation functions.

### 3.3 Directory services

Directory services allow for the querying and dissemination of metadata over the web. In relation to GIS data, this includes the projection, bounding boxes and information about additional attributes as well as citation information, keywords, abstract and the purpose of a dataset.

By providing the metadata held within an organisation over the Internet, it is possible for third party users to gain a greater understanding of the data holdings of an organisation. This has advantages in that it reduces the number of simple data explanation queries that come in to the data managers, and can increase the awareness of potential users and clients to products and services.

When providing data for display or download, it is straightforward to include a metadata record with it. This can be displayed either as a static XML page displayed over the Internet, included with the downloadable file, or as a dynamic page that is created from a query of the metadata catalogue. Either way of displaying the data is acceptable, but it is important to note that metadata is only of value if it is up to date and accurate in relation to the data it is explaining.

Metadata management software has been developed on multiple platforms, often in specific applications designed for individual organisations. For example, UNEP.Net and FAO have created their own metadata catalogues and management tools to store the data and create the XML outputs (see table 2). An example of a metadata management tool specifically designed to manage GIS data can be found in ESRI's ArcCatalog. The advantage of this system is that the metadata is always stored with the data. This allows some of the information such as the bounding box and projection to be completed automatically from the dataset. It also means that when the data is exported, the metadata is exported along with it. ESRI's ArcIMS also contains a metadata server that allows the user to directly search and query the metadata held in ArcCatalog over the Internet.

Various other software systems have been developed for the dissemination of metadata over the Internet. An example of this is the Z39.50 server that allows the user to query a metadata catalogue returning the XML page of the requested result. Z39.50 servers are used within the collaborating partners of the FGDC to supply their metadata to the clearinghouse.

# CHAPTER 4 EXAMPLES OF GIS INFORMATION AVAILABLE OVER THE INTERNET

## 4.1 Introduction

There are many different ways to provide GIS information over the Internet in many different formats such as maps, graphs, tables or raw data. Table 2 shows some of the different publicly available GIS data available on the web. Below are two examples of how available technologies have been used to provide access to GIS data, either on sites open to the public or through mechanisms that provide access to the data over the Internet but through private connections.

## 4.2 UNEP World Conservation Monitoring Centre (UNEP-WCMC)

UNEP-WCMC has made heavy use of ESRI's ArcIMS, ArcGIS and ArcSDE software to provide remote access to its GIS data, both publicly and privately.

The various interactive maps available publicly on the UNEP-WCMC website combine together into the brand name IMapS. These IMapS provide varying levels of functionality depending on the needs of the specific project. This ranges from fully interactive maps that allow the user to view and query a large number of layers with links to the other databases held within the Centre, through to simple static maps created 'on the fly' to display species and protected area distribution information in conjunction with other non-spatial data. More information on the IMapS available from UNEP-WCMC and the technologies used can be found in Annex 4**.**

Although these systems are currently reliant on ESRI software to display and query GIS data, the same functionality is also available through other software. Particular attention should be made to products that make use of open source technologies such as Web Map Services (WMS) and Web Feature Services (WFS). Various map viewers are freely available over the Internet that make use of these technologies, such as:

Intergraph WMS Viewer (http://imgs.intergraph.com/interop/wmsviewer.asp).

It is also easy to develop bespoke web pages to view and query GIS data. The link below provides various links to 'cookbooks' created to assist developers in constructing their own WMS or WFS.

http://www.opengis.org/resources/?page=cookbooks

UNEP-WCMC has also made use of ESRI software to provide external organisations with read access to its raw GIS data. This functionality allows a restricted user group to include the most up-to-date GIS data from UNEP-WCMC within their own GIS systems. Access is only provided to a restricted user group at present, and users can only view the information and not make alterations. However, the system would allow for remote write access to datasets held at UNEP-WCMC if desired. This has particular importance for IABIN where it may be desirable to allow contributing organisations to maintain GIS data held on the IABIN system. This would allow these organisations to maintain control and ownership of their data whilst providing them to IABIN.

### 4.3 The European Environment Agency (EEA)

The GIS department of the EEA is responsible for the development and maintenance of improved routines and systems for handling the flow of geographical information, partly within the EEA and partly between the EEA, the European Topic Centres (ETCs) and the outer world.

In late 2002, Information Technology and Services (ITS) – the GIS operations group of the EEA - published a project proposal for 'Implementing Spatial Data Infrastructure at the EEA'. This proposal outlines the concept, architecture and data management infrastructure that needs to be put in place to fully support the needs of the EEA and wider user groups. The proposal was made in line with the recommendations of the Global Spatial Data Infrastructure (GSDI) Group which has developed a recommended approach for the development of a spatial data infrastructure.

To develop the infrastructure needed to support the EEA, the ETCs and the outside world, the GIS department has made use of ESRI products but has also developed its own bespoke applications for managing the metadata. These applications are used to allow the contributing organisations to manage remotely the metadata held within the EEA system. This has allowed the contributing organisations to remain in control of their information, whilst providing it to the centralised system. This functionality has not yet been developed for the actual GIS data, which is currently supplied to the EEA via FTP, email or on CD, but it is envisaged that this will be developed under the next phase of the infrastructure development project.

More information on GIS within the EEA is available in Annex 5.

Below is a table highlighting some organisations where GIS information has been presented over the Internet in different formats. More information is available on those systems available in Japan in the accompanying report on GIS in Japan.

Table 2. Examples of GIS information available over the Internet

| Organisation | Website | GIS Information | Availability | Geographical Extent |
|---|---|---|---|---|
| Natureserve | www.natureserve.org | Downloadable species distribution maps in ArcView shapefile format | Public Access to general data but not for commercial use – Additional services available for a fee | North and Latin America |
| Conabio-REMIB | www.conabio.gob.mx | Downloadable Coverages Shapefiles, DXF files and satellite images with metadata | Access after registration through the website | Mainly Mexico and the USA |
| INBio | www.inbio.ac.cr | Database information on the sightings of individual species | Public Access for non-commercial use | Costa Rica |
| CRIA | www.cria.org.br | Map Service showing species distributions with background layers | Public Access to view and query the data | Brazil |
| JIBIS (Japanese Integrated Biodiversity Information System) | http://www.biodic.go.jp/english/J-IBIS.html | Distribution maps available for individual species | Public Access | Japan |
| GBIF | www.gbif.org | Database information on | | Global |

| Organisation | Website | GIS Information | Availability | Geographical Extent |
|---|---|---|---|---|
| | | the collection locations of individual species | | |
| UNEP.Net | www.unep.net | Datasets, Map Services and Metadata | Public Access | Global |
| BirdLife International | www.birdlife.org | Point data on the locations of Important Bird Areas downloadable as text files. | For non-commercial use as long as it is cited back to BirdLife. | Global |
| European Environment Agency | www.eea.eu.int | Downloadable datasets in ArcInfo export format. | Open access to a limited number of datasets, access to other datasets through their owners. | Europe |
| FAO GeoNetwork | www.fao.org/geonetwork/srv/en/main.search | Downloadable datasets, map services, imagery, applications and metadata | Downloadable after agreeing to terms and conditions. | Global |
| DNLIS Digital National Land Information System | http://nlftp.mlit.go.jp/cgi-bin/ksj/dls/_kategori_view.cgi | Data available for download in ASCII format | Downloadable after agreeing to terms and conditions. | Japan |

# CHAPTER 5 CONCLUSIONS AND RECOMMENDATIONS

## 5.1 Overview

A large number of GIS data standards and practices has been developed since the inception of GIS as a concept and process. Most of these have evolved through the needs of different organisations to share their data both internally and externally, reflecting those needs and the technologies available at the time. Today, there is a wide range of software technologies to support data exchange, interoperability and metadata services, with most organisations using a mixture of software to support their needs. When deciding on the standards and practices to adopt for sharing GIS-based information, the main point to take on board is the needs and capabilities of the contributors and end-users. If the technologies and standards employed do not allow easy integration with contributors and users, assist them in their programme of work, or provide them with the information they require, then their use is of little value.

Special consideration also needs to be given to the development of policies for the management and distribution of information provided by partners and third parties. Data distribution policies need to be developed to allow these providers to include their data in the network, whilst maintaining ownership and control over how the data is distributed and used within other systems. This will also have an effect on the type of system that is employed to store and manage the data within IABIN.

## 5.2 GIS Information Exchange Standards

In an ideal world GIS data would be entirely interoperable, meaning that the information would remain in the possession of the original owners and incorporated into other organisations through various web applications. If data is constantly being updated and is being exchanged on an *ad hoc* basis, it is very easy for the end-user to be working on an outdated copy of the dataset. However, due to Internet connection speed issues and end-user capacity, information exchange is often the most viable and convenient option for the sharing of data. To decide on the best way for an organisation to share data with others, it is important to assess the end-users' capabilities. It will probably be best to distribute data in a number of formats such as shapefiles or DXF (AutoCAD Digital Exchange Format), although to be truly open GML is the best route to take. GML has advantages in that it is a text-based standard and not reliant on any specific software to be displayed. However, it can cause problems when dealing with large datasets, as the file size increases dramatically when either a large number of records is stored or each record contains intricate lines or polygons.

### 5.3 Interoperability Standards

To integrate GIS data stored in different formats in different locations across the web, access must be provided through an Internet map or data server. To be fully interoperable between software systems, this server needs to provide information that complies with the OGC standards for WMS or WFS. A service that complies with WMS or WFS standards can be viewed in many software packages such as ArcIMS, but also has the advantage that it can be imported into different applications running on different softwares such as the OGC compliant map viewer used in the FGDC data clearinghouse. UNEP-WCMC provides maps over the internet using ESRI's ArcIMS technology, this has an additional WMS and WFS compliant connector that allows the same data to be viewed through other OGC compliant systems.

### 5.4 Spatial Metadata Standards

The construction and maintenance of metadata is critical for the effective sharing of GIS-based information. To make full use of metadata capabilities it needs to comply with accepted international standards. To publish metadata in any external clearinghouse, that metadata needs to comply with the standards of the site in question. At present many metadata clearinghouses have adopted the FGDC metadata standard, however, with the recent development and adoption of the ISO 19115 spatial metadata standard it is envisaged that clearinghouses will adopt this. Depending on the application used to manage metadata it is often possible to export it so that the output XML file complies with various different standards. For example, ESRI's metadata manager in ArcCatalog can export data in both FGDC and ISO standards as well as allowing the user to define their own export design.

To allow metadata to be fully interoperable with external organisations it needs to be searchable through the Internet. Many clearinghouses use the Z39.50 protocol to provide interoperable content between data providers. This can be implemented through various software packages.

## 5.5 Data Distribution Policy

Wherever possible it is recommended that an open access data policy should be employed in relation to all data outputs from IABIN. However, where information is provided with restrictions, this needs to be adhered to and access restrictions put in place.

To encourage partners to provide data to the system, a comprehensive security policy needs to be employed. This would allow the administrator to control access to data by requiring a login for users to access certain areas of the site. For example, certain areas of the database and simplified views on the data can be provided to the public, whereas more detailed information and restricted datasets can be provided to a more limited set of users such as the IABIN focal points or in-country biodiversity practitioners.

An area where this type of security could be important is with information on the distribution of endangered species. Access to detailed information on the actual location of species could be provided to a restricted user group, whereas the publicly accessible locations can be simplified to a certain level (see Table 3).

Table 3 An example of data location simplification

| Species | Detailed Location | Simplified Location |
|---|---|---|
| Yellow-eared Parrot (*Ognorhynchus icterotis*) | 1.105875 N / 77.612564 W | 1.10 N / 77.61 W |

## 5.6 Multiple Data Sources.

When constructing the IABIN system it is important to discuss fully with the contributing organisations the planned processes and mechanisms that need to be put in place. All possible contributors need to be consulted on the mechanism to ensure that it can be easily integrated within their current systems.

With this in mind, IABIN will also need to decide on the best way to include data from multiple sources into their system. As mentioned before, data can be incorporated into the system in two different ways: either by linking to information held within different organisations, or by incorporating that information into the IABIN system. Both have advantages and disadvantages with regard to accuracy, infrastructure within the collaborating centres and the currency of data provided. It is recommended that the IABIN system is constructed in such a way that it is possible to include data held both locally and remotely, to allow as many organisations to collaborate as possible.

# CHAPTER 6 ANNEXES

## 6.1 ANNEX 1 - Key Contacts

Peter Hall – Director of Biodiversity Information Systems, UNEP-WCMC

Phillip Fox - Head, Internet and GIS Department, UNEP-WCMC

Ian May - GIS Developer, UNEP-WCMC

## 6.2 ANNEX 2 - Acronyms and Abbreviations

| | |
|---|---|
| ESRI | Environmental Systems Research Institute (www.esri.com) |
| ETC | European Topic Centre |
| FAO | Food and Agriculture Organisation (www.fao.org) |
| FGDC | Federal Geographic Data Committee (www.fgdc.gov) |
| GIS | Geographical Information Systems |
| GML | Geography Markup Language |
| ISO | International Organization for standardisation (www.iso.org) |
| OGC | Open GIS Consortium (www.opengis.org) |
| UNEP | United Nations Environment Programme (www.unep.org) |
| UNEP-WCMC | UNEP World Conservation Monitoring Centre (www.unep-wcmc.org) |
| WFS | Web Feature Service |
| WMS | Web Map Service |
| XML | Extensible Markup Language |

## 6.3 ANNEX 3 - Glossary of Terms Used

| *Term* | *Definition* |
|---|---|
| ArcExplorer | ESRI’s Data viewing software |
| ArcGIS | ESRI’s GIS Software used to perform complicated GIS Processes |
| ArcSDE | ESRI’s Spatial Database Engine used to store spatial data |
| ArcView | ESRI’s GIS Software to perform simple GIS processes |
| Ellipsoid | An irregular sphere |
| Line | A chain of points connected by lines used to represent routes and such as roads and rivers. |
| Point | A single coordinate used to represent a location |
| Polygon | A multi-faceted vector graphic figure used to represent an area and made up of a closed chain of points and connecting lines |

## 6.4 ANNEX 4 - UNEP-WCMC Interactive Map Service

### 6.4.1 Introduction

(1) Data within UNEP-WCMC

Over the last 25 years, UNEP-WCMC has collated data from across the globe to create a comprehensive and reliable global biodiversity information service. This data provides decision-makers in the private sector, governments and international agencies with a source of information on the status of global biodiversity. UNEP-WCMC's data holdings cover a wide range of topics in a large number of formats, from information held in their paper library through to large databases containing thousands of records, such as CITES export permits and global Protected Areas distribution.

(2) GIS within UNEP-WCMC

For the last 15 years UNEP-WCMC has been managing and distributing GIS data. Through close ties with ESRI, UNEP-WCMC was able to develop its GIS data infrastructure and holdings to create a comprehensive Biodiversity Map Library (BML). This BML holds layers on a large number of subjects including Species, Habitats and Protected Areas. This information has been used to support a large number of publications and analyses including the UN List of Protected Areas, the World Atlas of Biodiversity, the World Atlas of Coral Reefs and the World Atlas of Seagrasses (a comprehensive list of the UNEP-WCMC publications is available from their website at http://www.unep-wcmc.org/reception/publications.htm).

### 6.4.2 IMAPS history

(1) Collaboration with the Oil Industry

Since 1995, UNEP-WCMC has been in partnership with the International Petroleum Industry Environmental Conservation Association (IPIECA). IPIECA is a non-profit making organisation whose membership is made up petroleum companies as well as national, regional and international organisations. The aim of the partnership was to design and manage a programme of activities to increase knowledge and develop a greater understanding of international biodiversity issues, relating to the petroleum industry.

Under this partnership and through collaboration with IPIECA's partners, the Interactive Map Service (IMapS) was developed at UNEP-WCMC to provide quick and easy access to accurate and up-to-date biodiversity information for use in emergency planning and response to oil pollution incidents.

IMapS delivers GIS information on the location of species, sensitive sites, infrastructure and background information over the Internet through the use of

ESRI's Internet mapping technologies. This allows you to zoom in and out, turn on and off specific layers, query the underlying data and extract related information.

(2) Support to UNEP-WCMC

Whilst the partnership with IPIECA provided the impetus for the development of IMapS, other projects were also initiated to take advantage of the technology. During the initial stages, IMapS were developed using MapObjects IMS 2.0 to show data on Indian Ocean Marine Turtles as well as the locations of Coral Diseases.

6.4.3 IMAPS Today

(1) Two Levels of Functionality

IMapS has developed over time to provide different resources for different needs. This has meant that two broad levels of functionality and display have been incorporated into the concept.

The first level can be described as the full interactive IMapS that is explained more fully below. This level allows the user to zoom in and out, query the under attribute data and link through to related information.

The second level of functionality is much more simple from the user's point of view. It allows for the incorporation of bespoke maps created on demand and displayed within other database outputs. An example of this is the species distribution maps created through the UNEP-WCMC species database. This simple applet takes country level distribution data from the non-spatial database and uses it to shade a map, highlighting the countries where we have information on a species presence or extinction. Where we have distribution information for the species queried, this can then be added on top.

This process has also been used to display maps on the location of protected areas within a country, as well as mining activities in relation to surrounding environmental sensitivities.

(2) IMapS Platform at UNEP-WCMC

All IMapS are developed using ESRI ArcIMS (4.0.1) that are implemented on Windows Servers with dual processors with high performance multiple hard discs, network cards and power supplies. These servers are all protected with a UPS (Unbreakable Power Supply) and back-up generator. A full list of all currently available IMapS is shown in Table 1.

The data layers viewable through the IMapS are held on separate database servers within ArcSDE 8.3 running on Microsoft SQL or Oracle 9i. Data is distributed to the Internet down twin 2 megabit/sec Internet connections. Data is distributed between the IMS and SDE servers on an internal 1gigabit network connection.

(3) Additional functionality

The basic ArcIMS html viewer has been extensively customised in IMapS by the inclusion of additional functionalities requested by the users.

Email:

The addition of an email button has allowed the user to email the URL of the current map to themselves or others. This URL links the recipient of the email back to the specific extent and visible layers the user was using at the time of sending the email.

Multiple Layer Selection:

The user is able to turn on multiple layers with a single click. This has been particularly useful for grouping specific themes together such as related species, protected areas or infrastructure layers.

Location Zoom:

A button has been added to the toolbar that allows the user to zoom to the country or point of interest without having to use the map. This has been of particular use where the user needs quick information on a specific location or country. By using this tool, which allows one to turn on groups of layers, it is possible for the user to view information on the map without the need for an in-depth knowledge of the area of interest.

Location Zoom:

Bespoke development of the original IMapS tool has also allowed the user to select from a range of languages. Although the layer names are still written in English, it is now possible to view the buttons, tools and help page in one of six different languages.

Extended Hyperlinks:

IMapS makes extensive use of the hyperlink function within ArcIMS to combine information held within the GIS layer with other information sources. This has been of particular use with the Protected Areas database. Both the GIS layers and non-spatial database have corresponding unique identification codes that allow the user to click on a particular feature on the map and link through to additional information held within a relational database. Using Macromedia coldfusion, UNEP-WCMC has developed many interfaces to additional attributes and related databases.

(4) Support to Industry

A large part of the momentum behind the development of IMapS and its ongoing evolution has been driven by the partnership between UNEP-WCMC and IPIECA. This development has been guided with the specific needs of the oil industry in mind.

Today there are currently four regional IMapS online which have been created with the support of IPIECA and its members. These IMapS cover the Mediterranean, Black Sea, Caribbean and Caspian. Due to the specific needs of the users for environmental information in exploration, contingency planning and emergency response operations, the layers included and the information available has been tailored to their needs. Main sets of information include:

- Nationally Designated Protected Areas
- World Heritage Sites
- RAMSAR Sites
- UNESCO MAB Biosphere Reserves
- Important Bird Areas
- Endangered Species Distributions
- Coastal Environmental Sensitivity Indexes
- Coral Reefs
- Seagrasses
- Mangroves
- Wetlands
- Pipelines
- Airports
- Roads
- Railways
- Cities

These IMapS are designed to provide information that is similar to that found in local Environmental Sensitivity Maps. Further information on the design and uses of Environmental Sensitivity Maps is available in the IMO/IPIECA Report Series Volume 1, which is available from the IPIECA Website at:

http://www.ipieca.org/publications/oilspill.html

**(5)** Support to UNEP-WCMC

Since their inception, IMapS have been used to support programmes and projects ongoing within UNEP-WCMC. Today IMapS have been created to support the World Atlas of Biodiversity and the World Atlas of Seagrasses. These IMapS allow the user to view an online version of the maps produced in these books, giving the added advantage that the user can zoom in on a specific area of interest and interrogate the underlying data and source information.

For the World Atlas of Biodiversity, the IMapS was created to allow the user to turn on the layers that make up specific maps in the book. Users are also able to combine layers that were not necessarily viewable together previously.

For the World Atlas of Seagrasses, the IMapS allows the user to view all the data that makes up the maps in the book. The dropdown list on the right hand side of the map allows the user to view the distribution of the entire list of individual species. By using the hyperlink on any of the species points or polygons, it is possible to link through to the source information for the data at that location.

(6) Support to Organisations

UNEP-WCMC has developed various IMapS to support the programme of work of other organisations. These organisations often have information management and dissemination needs that require the use of GIS, but that they are unable to supply in-house. This has advantages for the organisation and UNEP-WCMC. The organisation acquires abilities that are not available in-house and UNEP-WCMC gains access to data that would not be available to it ordinarily.

As a specific example of this, UNEP-WCMC has been working in conjunction with the Earthwatch Institute to develop an IMapS that allows the user to identify and query their project locations to return the project reports for those sites.

(7) Support to Conventions

In its normal programme of work UNEP-WCMC provides support to various conventions including CITES (the Convention on International Trade in Endangered Species of Wild Fauna and Flora) and the CMS (Convention on Migratory Species). Under specific projects within this programme of work UNEP-WCMC has developed several IMapS to support these conventions, as follows:

Species Trade:

A prototype IMapS has been developed for CITES to highlight the possibilities of displaying trade information on a specific species. The current prototype uses the example of the Common Iguana (*Iguana iguana*) to identify the exporting and importing countries per year.

Marine Turtles:
Through a project developed by UNEP-WCMC on behalf of CMS and the IOSEA (Indian Ocean – South-east Asian Marine Turtle Memorandum of Understanding), UNEP-WCMC has developed an IMapS as part of an interactive facility to allow the exchange of data, whilst enabling contributors to retain ownership and credit for their work.

This Marine Turtles IMapS is designed to facilitate the integration of public-domain field data; such as distribution, abundance, migration, trends, status and photographs. Information is also provided on index beaches, habitat information such as presence and extent of sea grasses, coral reefs, mangroves, priority areas such as Internationally and Nationally Protected Areas, and physical background parameters.

It is intended that in the future users will be able to upload and manage their own data independently of other users, whilst allowing it to be viewed on a regional scale with comparable data from elsewhere. The facility has been developed initially for the Indian Ocean and it is hoped that in the near future the project can be developed to cover the Pacific, Atlantic and Caribbean.

(8) Exercise Clean Seas

As one of the primary uses for IMapS is emergency response and planning, it has been a natural progression to incorporate IMapS into an emergency response training exercise called Exercise Clean Seas. This exercise is designed to be a realistic, multi-scenario exercise that fills a gap in the training of responders and emergency planners at all levels. A set of standard scenarios is available in the public domain and companies can customise these or create their own to suit their particular training requirements. Exercise Clean Seas complements existing IMO (International Maritime Organisation) and industry oil spill management initiatives and has been adopted by the IMO as their standard training exercise.

The basic concept is to provide the players with a package that is as versatile as possible to allow them to explore the issues and concepts relating to oil spill response and planning. The package contains a map with playing pieces and supporting information including a supporting IMapS of the area. The map shows two sovereign nations and the players receive information on the infrastructure and environmental situation in both countries. From this the user plans and responds to one of a number of different marine pollution scenarios.

(9) Interoperability

The aim of UNEP-WCMC with IMapS is to fulfil the prerequisites required to be OGC compliant, which therefore allow IMapS to be included in other web map viewers.

In addition, two specific layers, Nationally Protected Areas (points) and Coral Reef (lines) have been included in the FGDC (Federal Geographic Data Committee) Clearinghouse. This allows any user of the clearinghouse to search for metadata on these two layers, after which they can include the GIS data into an OGC Map Viewer with data from other sources.

**Table 1: Currently Available IMapS**

| TITLE | URL |
|---|---|
| Mediterranean | http://nene.unep-wcmc.org/imaps/ipieca/mediterranean |
| Black Sea | http://nene.unep-wcmc.org/imaps/ipieca/blacksea |
| Caribbean | http://deben.unep-wcmc.org/imaps/ipieca/caribbean |
| Caspian Sea | http://nene.unep-wcmc.org/imaps/ipieca/caspian |
| World Atlas of Seagrasses | http://stort.unep-wcmc.org/imaps/marine/seagrass |
| World Atlas of Biodiversity | http://stort.unep-wcmc.org/imaps/gb2002/book |
| Earthwatch Projects | http://bure.unep-wcmc.org/imaps/earthwatch |
| CITES Trade Prototype | http://bure.unep-wcmc.org/imaps/CITES/trade |
| Indian Ocean Marine Turtles | http://deben.unep-wcmc.org/imaps/IndTurtles |

| | |
|---|---|
| World Heritage Sites | http://nene.unep-wcmc.org/imaps/W_Heritage |
| Donor Information Sharing | http://nene.unep-wcmc.org/imaps/DFIDProjects |
| Bird Migration Routes and Breeding Areas | http://bure.unep-wcmc.org/imaps/AEWA |
| GRASP - Great Apes Survival Project – In Development | http://stort.unep-wcmc.org/imaps/GRASP |

### 6.4.4 Conclusions

Using ArcIMS technology, UNEP-WCMC has produced a wide range of Internet Map Services to distribute GIS data over the Internet. Whilst the individual IMapS have been developed for specific projects and user needs, some of the functionality has applications in other systems. With regard to IABIN, some of the functionality that may be of interest is the ability to include bespoke maps in the output from databases and the ability to link from locations on a map through to additional outputs, drawing on information held elsewhere.

## 6.5 ANNEX 5 - GIS and mapping by the European Environment Agency

### 6.5.1 Introduction

(1) About the European Environment Agency (EEA) and the European Environment Information and Observation Network (EIONET)

1) The European Environment Agency (EEA)

The European Environment Agency's core task is to provide decision-makers with the information needed for making sound and effective policies to protect the environment and support sustainable development.

The Agency ensures this information is available to the general public through its publications and website. The EEA does not make or enforce European Union environment policy or legislation: this is the responsibility of the European Commission and the other EU institutions.

The information provided by the EEA focuses, in particular, on assessing the current and future state of the environment across Europe and the pressures upon it. The Agency's tasks also include disseminating best practice in environmental protection and technologies, and supporting the European Commission in diffusing information on the results of environmental research.

The Agency both gathers and distributes its data and information through the European environment information and observation network (EIONET), which brings together over 300 environmental bodies, agencies, public and private research centres and centres of expertise across Europe. The EEA is responsible for coordinating the EIONET.

The Agency currently has 31 member countries, and is the first EU body to open its doors to the 13 countries in central and eastern Europe and the Mediterranean basin that have applied for membership of the EU. Current member countries are:

The 25 European Union Member States

Iceland, Norway and Liechtenstein, which are members of the European Economic Area, the Slovak Republic and Turkey.

Source: http://org.eea.eu.int/documents/who_we_are

The work of collecting the information/data sets is carried out by EEA and associated Environmental Topic Centres ETCs. The EEA is the body responsible for final preparation and publishing of information.

2) The European Environment Information and Observation Network (EIONET)

EIONET is a collaborative network of the EEA and its Member Countries, connecting National Focal Points in the EU and accession countries, European Topic Centres (ETCs), National Reference Centres, and Main Component Elements. These organisations jointly provide the information that is used for making decisions for improving the state of environment in Europe and making EU policies more effective. EIONET is both a network of organisations and an electronic network (e-EIONET).

Source: http://www.eionet.eu.int/Best_Practice/About_EIONET

### 6.5.2 GIS within the European Environment Agency

(1) The GIS Department

GIS and spatial data interoperability and standards within the EEA fall under the umbrella of EIONET. The GIS team, in existence since 2002 has responsibility for the development of improved routines and systems for handling the flow of geographical information, partly within the EEA and partly between the EEA, the ETCs and the outer world. The GIS team have been instrumental in the development of an EEA spatial data infrastructure covering issues such as metadata, map guidelines and data standards and practices. This is also important in the context of INSPIRE (the Infrastructure for Spatial Information in Europe).

INSPIRE is a recent initiative launched by the European Commission and developed in collaboration with Member States and accession countries. It aims at making available relevant, harmonised and quality geographic information to support formulation, implementation, monitoring and evaluation of Community policies with a territorial dimension or impact.

Source: http://www.ec-gis.org/e-esdi/

(2) Development of the System

In late 2002, the Information Technology and Services (ITS) - GIS operations group of the European Environment Agency published a project proposal for ‘Implementing Spatial Data Infrastructure at the EEA’ which is available from their website at:

http://eionet.eu.int/gis/docs/EEA_SDI_Proposal2002.doc

This proposal outlines the concept, architecture and data management infrastructure that needed to be put in place to fully support the needs of the EEA and wider user groups. The proposal was made in line with the recommendations of the Global Spatial Data Infrastructure (GSDI) Group which has developed a recommended approach for the development of a spatial data infrastructure.

As part of the proposal four main scenarios were identified for needs of the spatial data infrastructure. These are to:

1. Discover Existing Data
2. Reuse a known data service
3. Upload a dataset
4. Publish Metadata

These needs cover the use and requirements of spatial data within the EEA, between the EEA and ETCs and the wider world. All these scenarios are interrelated and by developing the infrastructure to support one of the scenarios steps are taken to develop the others.

1) Discovering Existing Data

This scenario identifies the need of people within the EEA, ETCs and wider world to identify existing datasets, where they are located and what are the access rights and constraints.

2) Re-using an existing data service

This scenario identifies the need of the users to have access to the data identified in the first scenario. By developing online access to identified data it should be possible for the user to analyse and report on a given subject.

3) Uploading datasets

This scenario covers the need of the EEA and ETCs to upload datasets held within their local systems to a central spatial data store for use in the wider EEA community and outside world

4) Publishing Metadata

To support all the other scenarios there is the need of the EEA and ETCs to publish metadata on spatial datasets to allow them to be queried, viewed and stored effectively within the central spatial data store.

5) Development of the infrastructure

To develop the infrastructure to support the above scenarios, the EEA has taken several steps to install the relevant hardware and software environment as well as deciding on the standards and practices that need to be in place before the scenarios can be fulfilled.

The spatial data storage system is based on ArcSDE installed on a server running SQL server. This system manages both the spatial data and the metadata. Various bespoke applications have been added to this system to allow for the management, querying and display of the data and metadata.

At this stage, the actual spatial data is collected from the ETCs on a variety of formats and uploaded into the system by the managers within the GIS team of the EEA data can be supplied in different ways such as via FTP, email and posted on CD.

Once this data has been added to ArcSDE, the metadata needs to be included before it can be queried online. This Metadata is stored in line with ISO standard 19115 for spatial data metadata. The metadata catalogue is held within ArcSDE along with the spatial layers. This has advantages because, as the spatial data is updated, some areas of the metadata are also updated with it, such as bounding boxes and projection information. It is also useful, as there is less chance of the metadata and data becoming inconsistent with each other.

The developers within EEA found that the Metadata entry mechanism within AreGIS ArcCatalog was not ideal for there needs so developed a bespoke metadata entry tool which allows them to easily copy delete and clean metadata records stored within ArcSDE.

After creating the mechanism to add data and metadata to the spatial data infrastructure it was important to develop ways in which it could be accessed, queried and displayed over the web. This was done through bespoke applications, one for the querying of the metadata catalogue and one for the display of the relevant spatial data.

The metadata query tool relies on web pages developed using Active Server Page (ASP) technologies to display queries on the ArcSDE data tables. The pages that are created allow the user to query metadata, and return the record including access to online resources including map services, maps and downloadable data where applicable. As some datasets are restricted to within the EEA and ETCs and cannot be used by the outside world there is also a restricted site that requires the user to log in before querying metadata and downloading the data.

The application used to display the data was developed using ESRI's ArcIMS and MapObjects technology. A bespoke Map Service was developed within the GIS team that allows web developers to query the GIS data either by entering a URL that links using ASP through to the map service that returns a map image, or by an

XML request for more complicated maps. This map service complies with the standards created through the Open GIS Consortium (OGC) for the development of web mapping services.

### 6.5.3 Conclusions

The EEA has made great gains in the development of it GIS capability through the growth of its spatial data infrastructure. By developing in line with the guidelines set down by the GSDI with the use of international standards and practices, it has been able to create a system that is distributed in terms of its data providers and users, whilst maintaining a level of control over the data quality and ownership.

The development of this infrastructure is still underway and it is envisaged that in the future the ETCs will be more full integrated with the central EEA data store. At the moment the uploading of spatial data is undertaken through the use of FTP, email and CD delivery, which has several issues with regard to how current the data is that is held with the EEA.