How INSPIREd is NERC?

Jerry Giles[1], Robin Waters[2], Rob Walker[3]
[1]NERC, jrag@bgs.ac.uk ; [2]RSW Geomatics, robin.waters@rswgeomatics.com ;
[3]Rob Walker Consultancy, robwalker@cix.co.uk

ABSTRACT

The Natural Environment Research Council (www.nerc.ac.uk) is the UK's main agency for funding and managing research, training and knowledge exchange in the environmental sciences.

In 2007 NERC commissioned a consultancy to prepare an INSPIRE baseline and Road Map to enable it to be compliant with the EU INSPIRE Directive well ahead of the deadlines listed in the Directive.

This study provided:
- A baseline of INSPIRE readiness across NERC with respect to INSPIRE requirements for metadata, discovery, view and download services;
- A description of what NERC will need to provide to fully comply with the INSPIRE Directive;
- A description of the technology options that are currently envisaged to implement the INSPIRE Directive;
- A Road Map to show what NERC must do to meet the INSPIRE Directive;
- An estimate of resources required to implement the INSPIRE Directive.

This paper outlines the findings of this study.

Keywords: INSPIRE, NERC, Natural Environment, spatial data
.
1. INTRODUCTION

The Natural Environment Research Council (NERC) is the UK's main agency for funding and managing research, training and knowledge exchange in the environmental sciences. Its role is: 'to gather and apply knowledge, create understanding and predict the behaviour of the natural environment and its resources, and communicate all aspects of our work'.

NERC has four of its own Research Centres:
- British Antarctic Survey (BAS),
- British Geological Survey (BGS),
- Centre for Ecology & Hydrology (CEH),
- Proudman Oceanographic Laboratory (POL).

There are also fifteen collaborative centres.

The data that is gathered from the research is regarded as the 'most precious asset'. There are six NERC Data Centres which hold and manage this data and provide access to the data holdings:

- Atmospheric science - British Atmospheric Data Centre (BADC) at the Rutherford Appleton Laboratory (RAL),
- Earth sciences - National Geoscience Data Centre (NGDC) at BGS,
- Earth observation – NERC Earth Observation Data Centre (NEODC) at RAL,
- Marine Science - British Oceanographic Data Centre (BODC) at POL,
- Polar Science - Polar Data Centre (PDC) at BAS,
- Terrestrial & Freshwater Science, Hydrology, Bioinformatics – Environmental Information Centre (EIC) at CEH.

In 2007, NERC, through its Data Management Advisory Committee (DMAG) commissioned a consultancy to prepare an INSPIRE baseline and Road Map to enable it to be compliant with the EU INSPIRE Directive well ahead of the deadlines listed in the Directive.

Specifically the brief called for these components:
- A baseline of INSPIRE readiness across NERC with respect to INSPIRE requirements for metadata, discovery, view and download services;
- A description of what NERC will need to provide to fully comply with the INSPIRE Directive;
- A description of the technology options that are currently envisaged to implement the INSPIRE Directive;
- A Road Map to show what NERC must do to meet the INSPIRE Directive;
- An estimate of resources required to implement the INSPIRE Directive.

2. INSPIRE AND NERC

The INSPIRE Directive identifies 34 spatial data 'themes' grouped into three 'Annexes' in order of priority. NERC are actually responsible for input to datasets that fall into 20 of these themes, as shown in Annex A. In addition, almost all of the themes are relevant to NERC as a user of spatial data.

Defra (Department for the Environment, Food and Rural Affairs) is the UK government lead department and provides the UK's representative on the INSPIRE Committee. They are proposing a UK national geo-portal which could be run by, or under contract to, government. A government sponsored metadata portal, Gigateway® has been operated by the Association for Geographic Information (AGI), through which there has been access to several 'nodes' including BADC, NEODC and BGS. There is a marine data equivalent to Gigateway® being developed by MEDIN (Marine Environmental Data Information Network) which includes Defra, BODC and the NGDC.

NERC data centres already have their 'portals' (sometimes several) but not necessarily operating to the standards required for INSPIRE. The minimum requirement for a data centre is that it provides compliant metadata (or a service that maps its own metadata to the standard) and that the datasets themselves can be accessed through a service that provides a minimum functionality.

NERC is active in all modes of interaction with the INSPIRE process from the Drafting Teams working on the Implementing Rules to the UK Transposition process. Documents and examples of datasets and services have been submitted as reference material to Drafting Teams.

3. INSPIRE Baseline Readiness
3.1 Applicability
The study included a detailed set of comments on the actual wording of the Directive and how it might be interpreted. This can be summarised as:

1. NERC is a 'public authority' and therefore within scope of INSPIRE.
2. NERC is responsible for datasets that fall into at least 22 of the 34 INSPIRE themes and actually uses all of the others (except 2) for reference frameworks, querying or reporting.
3. BODC, EIC, NGDC & NEODC will need to produce some compliant metadata by May 2010.
4. Data sharing appears to allow the business models used by some UK public sector bodies without interference (e.g. BGS).
5. Some NERC Data Centres are already providing basic versions of most of the services required by INSPIRE.
6. There are various 'get-out' clauses that may allow NERC to opt out of some requirements for some datasets. However it is not clear how these will be enforced or appealed. NERC may decide, for reasons of consistency, that it is in its own interests to comply with INSPIRE compliant rules, even if they are not mandatory for all NERC datasets.
7. It is for the UK (and its public authorities) to decide how to provide portals to access metadata and/or the datasets themselves.
8. Minimum performance criteria are specified but not defined – how onerous these might be will not be known until the relevant IR are finalised.
9. References to structures and mechanisms in Member States are presumed to refer to the UK's Location Strategy and/or to the devolved government GI Strategies.
10. The Directive emphasises the need to use existing international standards wherever possible.

3.2 Web access
All the NERC Data Centres have web sites that enable datasets to be found and accessed. The picture across NERC was found to be very mixed with some datasets in some Data Centres that can easily be discovered, viewed, downloaded and even with some basic transformations also available.

3.3 Metadata holdings
A survey of metadata holdings in the Data Centres was carried out. This looked into the number of spatial datasets for which metadata is available, the number for which it currently is not, formats that the metadata is in, and issues associated with the creation of metadata. A range of portals are used, including the NERC Data Discovery Service (DDS) and each establishment's own portal.

It is estimated that a similar number of datasets exist but do not currently have metadata available through the DDS. CEH in particular is estimated to have several hundred such datasets.

Access conditions apply to the actual datasets in many cases, with restrictions and/or charging on commercial use, and extra licensing conditions where third party copyright data is incorporated (e.g. Ordnance Survey).

Metadata is held in a range of formats including ISO 19115, DIF and various local formats. It is not a major task to migrate metadata from these formats into the INSPIRE metadata format, although the need to capture some additional metadata could be, in some cases, substantial. Where metadata is not currently available, it was estimated that approximately half a person day per dataset would be needed.

A potential issue concerns the use of controlled vocabularies for the selection of keywords to describe the data resource. These vocabularies either exist in most data centres, or are under development. This requires a considerable resource (approximately one full-time person) for development and subsequent maintenance.

4. GAP ANALYSIS
4.1 General
NERC Data Centres generally score well on discovery and download services but less well with viewing and transformations. This reflects both the international outlook of some Data Centres and the technical and service level limitations on automatic and interactive services for spatial data

4.2 Discovery services

Services for the provision of standard metadata are well understood and being made available. Whilst many Data Centres record metadata and make it available externally, the metadata formats used are different from those specified by INSPIRE. The NERC metadata elements will be capable of mapping to the INSPIRE set, but some additional metadata elements will also be required.

4.3 Viewing services

Viewing services require more sophisticated and resource intensive systems that will often need background information (e.g. topography) in order for the user experience to be meaningful. They may also be application specific and therefore better provided by user oriented portals rather than those provided by the data suppliers. For example, subsets of geological data provided for use by planners may be very different from those provided for engineers or for the extractive industries. BGS has recognised this with several different services provided to different markets.

4.4 Download services

These, at least for a whole dataset level, are relatively easy to implement provided that the data volumes are not too large and that licensing is not complex. More complex extraction of part datasets, and requirements to impose sophisticated Digital Rights Management, overlap with the network services and transformations.

4.5 Network services

These include facilities to extract partial datasets – by geography or by other criteria; to transform datasets from their original reference frame into others; and to combine different datasets to synthesise new combinations for viewing and analysis. Environmental Information Centre datasets are typically organised within projects that have relatively high external recognition, often with their own well functioning, but quite separate web sites. The EIC is tackling this issue and should at the same time be able to make them INSPIRE compliant. However, this will require the co-operation of various non-NERC organisations, some of them voluntary, with very limited resources.

5. TECHNOLOGY OPTIONS

The requirements for discovery and view services can be satisfied by standard specifications; whereas download, transformation and invoke services requirements are not yet fully understood and are either not yet or only partially supported by current specifications and thus need further development.

6. ROADMAP FOR COMPLIANCE

6.1 Metadata

Metadata for datasets covered by Annexes I and II is required by December 2010 and for datasets covered by Annex III (the majority of NERC datasets) the metadata is required by December 2013. NERC therefore needs to:

Update existing metadata to cover INSPIRE elements not currently held;

Create metadata for datasets where it does not currently exist;

Provide mapping schemas from internal formats into INSPIRE format;

Develop any new metadata repositories to INSPIRE specifications on a common NERC-wide basis.

6.2 Monitoring and reporting

Responsibility for reporting by the UK will lie with Defra and they will in turn determine exactly what information is required from NERC (and other data providers) and when. Indicators will include:

- Existence (of a dataset);
- Existence, accessibility & compliance of metadata;
- Interoperability;
- Existence and performance of view, download, and possibly other services;
- Use of datasets and services.

6.3 Discovery and view services

It is the responsibility of a Member State (MS) to provide Discovery and View Services and to decide which organisations within the MS implement them. It is assumed that NERC will provide at least one Discovery and View portal but there is no INSPIRE mandated reason why there should not be multiple portals offering these services in one or more NERC data centres. The current INSPIRE Roadmap specifies that Discovery and View services should be operational by November 2010.

6.4 Download services

NERC Data Centres already provide download services and these will need to be made compliant. There may be some datasets within NERC that are not currently available for download for various reasons. When metadata for each dataset is published it should become apparent whether any further action needs to be taken for INSPIRE.

6.5 Coordinate transformation service

NERC holds datasets in different co-ordinate systems that have been customary in different disciplines. Transformations are available between most of these but will need to be made available through an easily usable service – perhaps as part of the download service.

6.6 Community access

NERC will need to clarify its own position on licensing its datasets for all users. It is understood that Defra may set up a 'clearing house' in the UK for licence issues to ensure some consistency between public authorities. NERC will also need to clarify the position regarding its use of reference material or derived data (e.g. directly or derived from Ordnance Survey) to European Community institutions and bodies. Defra will need to co-ordinate such issues.

6.7 Data specifications

The Thematic Working Groups (TWGs) for each of the Annex I themes published the first draft specifications for consultations in December 2008 with a view to some, at least, being submitted to the INSPIRE Committee during 2009. NERC's Data Centres are commenting on these specifications individually and through other pan European projects in which they are involved. NERC Data Centres use and often derive information from the reference datasets in Annex I & II.

6.8 Geoportals

There is no specific requirement for national or thematic geo-portals in Member States. However current Defra thinking is that there will be a 'national portal with links to any portals developed by the Devolved Administrations and the EC INSPIRE Geoportal'. NERC's own portals can continue though they may wish to implement the INSPIRE standards.

6.9 Spatial data service network service (Invoke)

There are no details beyond a short section in the Network Services Architecture document. The "Invoke Spatial Data Service" allows definition of both the data inputs and data outputs expected by the service, and definition of a workflow or service chain combining multiple services. It also allows definition of the external web service interface of the workflow or service chain. NERC cannot take any action at this stage.

6.10 UK Implementation

INSPIRE specifically mandates Member States to ensure that datasets and services within scope are available and sharable, not just for Commission purposes, but also within countries. Defra is now co-ordinating INSPIRE with the UK Location Strategy. NERC has provided a representative on the government's Working Group for INSPIRE and is therefore in a good position to be pro-active in ensuring that NERC concerns are taken into account during implementation. NERC should also take into account a likely increase in demand for datasets and services as INSPIRE is implemented. This may manifest itself through other portals (e.g. National and EC Geoportals) as well as direct from any NERC portals.

7. RESOURCES REQUIRED

The INSPIRE Directive essentially requires implementation of best practice for data management. Most of the work that is needed to meet these requirements will have to be carried out anyway, and is likely to form part of the current plans for NERC data management and development of services.

There will however be a requirement for additional metadata, both extending existing metadata to meet the INSPIRE requirements, and the creation of new metadata for existing datasets where none currently exists. This should not be especially onerous. The effort requirement for this is estimated at approximately half a day per dataset. This might be an issue for the EIC, where there are potentially several hundred such datasets.

It is unlikely that any additional resources will be made available from Government funds to meet INSPIRE obligations. It is expected that compliance with INSPIRE will be met with existing planned resources, since there are no new data collection requirements. Any new data collection will be justified in the usual way, and data management will be resourced as at present.

However, increasing use of NERC data could lead to a greater demand for server and other network resources. If NERC believes this is likely then it must bear in mind its own 'public task' – on which INSPIRE should have no direct impact – and consider how to resource extra demands.

8. CONCLUSIONS & RECOMMENDATIONS

8.1 The INSPIRE Directive is applicable to NERC, and NERC Data Centres have responsibility for many spatial datasets within the remit of INSPIRE.

8.2 The INSPIRE Implementing Rules are not yet finalised, but will encompass current best practice for data management.

8.3 INSPIRE requires that metadata be published for UK datasets held by NERC Data Centres. Many of these are currently collecting metadata, but in a different format from that defined in the INSPIRE Implementing Rules. These can be mapped into the INSPIRE format, although some additional metadata elements may be required.

8.4 Basic 'discovery' and 'view' services will be required for datasets within scope. Whilst the detailed mechanisms for these services are not yet known, they are likely to be based upon similar procedures already used in NERC.

8.5 NERC is already standardising data management across its Data Centres and this co-ordination will be vital if INSPIRE is to be implemented efficiently.

8.6 NERC will need to ensure that its Data Policy encompasses INSPIRE requirements and continue to develop common data management procedures across the Data Centres, based upon current best practice. It is recommended that conformance with INSPIRE should be sought across all datasets whether strictly within scope or not, for consistency.

8.7 NERC should continue to harmonise processes for metadata collection across the Data Centres, expanding the DIF standards to meet the INSPIRE requirements. There is already a Metadata Content Subgroup charged with this responsibility. Collection of metadata should be encouraged centrally as well as locally, even to existing standards.

8.8 Common data discovery and view services should continue to be developed across the Data Centres, following current best practice. These should be expanded to follow the INSPIRE Implementing Rules as they develop.

8.9 NERC should clarify licensing of datasets for all users, including regarding its use of reference material or derived data, and particularly for use by EC institutions and bodies.

ANNEX A. NERC RESPONSIBILITY FOR INSPIRE THEME DATASETS

Annex	Theme	Data responsibility in NERC
I	Coordinate Reference Systems	NGDC gravity & Geomagnetism
	Geographical Grid Systems	BODC for Gebco One Minute Grid
	Geographical Names	BODC for Gebco Underwater Names
	Hydrography	EIC
II	Elevation	BODC for Gebco; NEODC have various commissioned datasets
	Land cover	EIC
	Orthoimagery	NEODC have archives
	Geology	NGDC
III	Soil	EIC
	Human Health & Safety	EIC, NGDC Radon etc
	Utility and government services	EIC & NGDC have obligations to EA
	Environmental monitoring facilities	EIC for river flows; BODC; NGDC
	Production and Industrial Facilities	NGDC
	Natural risk zones	NGDC; EIC
	Atmospheric conditions	BADC; EIC
	Meteorological geographical features	BADC; BODC; EIC
	Oceanographic geographical features	BODC
	Bio-geographical regions	EIC
	Habitats and biotopes	EIC
	Species distribution	EIC
	Energy resources	BADC; BODC; NGDC; EIC
	Mineral resources	NGDC

ANNEX B. GLOSSARY

Initials	Full Name	Comment
AGI	Association for Geographic Information	Non profit association currently running metadata service
BADC	British Atmospheric Data Centre	NERC Data Centre (at RAL)
BAS	British Antarctic Survey	NERC Institute
BGS	British Geological Survey	NERC Institute
BODC	British Oceanographic Data Centre	NERC Data Centre (at POL)
CEH	Centre for Ecology and Hydrology	NERC Institute
DDS	(NERC) Data Discovery Service	
Defra	Department for the Environment, Food & Rural Affairs	Lead UK government deptartment for INSPIRE
DMAG	(NERC) Data Management Advisory Group	Commissioned this study
EIC	Environmental Information Centre	NERC Data Centre (at CEH)
Gebco	General Bathymetric Chart of the Oceans	International co-operative programme.
MEDIN	Marine Environmental Data & Information Network	Sponsored by NERC, Defra, et al
NEODC	NERC Earth Observation Data Centre	NERC Data Centre (at RAL)
NERC	Natural Environment Research Council	
NGDC	National GeoScience Data Centre	NERC Data Centre (at BGS)
PDC	Polar (previously Antarctic Environmental) Data Centre	NERC Data Centre (at BAS)
POL	Proudman Oceanographic Laboratory	NERC Institute
RAL	Rutherford Appleton Laboratory	UK Government Laboratory that hosts two NERC Data Centres