

Clear Info Completed Actions Sign Off Sheet



Project Reference	LIFE10 ENV/UK/000175 CLEAR Info	
Action number	2.2	
Description	Identify and apply INSPIRE standards to integrated data.	
Version	Final	
Contents	Action 2.2 Report Annex 1 Consultant's report "CLEAR Info: Publishing in INSPIRE data formats"	
Written By	Emily Stone	Date 11/09/14
Project Manager Sign-off & Date	Suzanne Laidlaw	Date 11/09/14
Project Executive Sign off & Date	Gillian Pratt	Date 15/09/14

CLEAR Info Project Report

September 2014

Action 2.2 Identify and apply INSPIRE standards to integrated data

1.0 Introduction

This activity was scheduled to take place between 1 January 2013 and 31 December 2013. The project team was focussed on other activities, and was unable to assign a dedicated team member to take responsibility for Action 2.2, so relied upon support from other areas of the Environment Agency. Following some delays in procurement, a scoping study was in draft by October 2013.

Analysis revealed that the scale of work required to analyse and match all the CLEAR Info data to the INSPIRE standards was beyond what was possible within the project time and resources. It was decided that the project would deliver a single data set, harmonised to INSPIRE standards as a means to demonstrate the approach. The remaining work, to apply the INSPIRE standards to the data within the data integration system, to produce a report, and to share our findings on the opportunities and challenges this creates, will be undertaken as part of Action 5.1, to explore opportunities to improve access and exchange of data.

This report explains how the INSPIRE standards that apply to the CLEAR Info integrated data were identified. This includes an overview of the scoping study which was commissioned to enable this.

1.1 Overview of INSPIRE

The INSPIRE Directive aims to ensure Member States publish their spatial data in a compatible way. It relates spatial data to an environmental theme and requires common implementing rules for 34 environmental data themes. Compatible spatial data on environmental themes is intended to aid the use of environmental data at a European Community scale, and inform trans-boundary policy making. To do this, the Commission is adopting Implementation Rules as binding regulations.

In England, the Department for the environment, food and rural affairs (Defra) are the lead agency for the implementation of the INSPIRE directive, and make the decisions on how this will be undertaken. The required data sets are currently published on websites such as the

Environment Agency Data Download site. They are not required to be transformed to the INSPIRE standard until 2020.

The CLEAR Info project proposes to harmonise the CLEAR Info collated data set so that it aligns with the standard. Many companies operate internationally so the project aims to publish data in an interoperable format which can be shared across European regulatory boundaries to build a better understanding of corporate compliance with EU legislation.

1.2 Study to identify standards that apply to water quality data

Supported by data specialists from the Environment Agency Evidence department, the project team commissioned a study to identify the INSPIRE standards that apply to one of the data sets which feeds into the CLEAR Info data integration system. The decision to initially commission a study of a single data set was to understand the resources and process required for identifying the applicable INSPIRE standards.

The study focused on data in the Environment Agency's Water Information Management System (WIMS). This data comprises site details, discharge consents, and water quality sample results. This information falls under Annex III of the INSPIRE Directive.

The WIMS data set was chosen as the focus for this study to maximise the opportunity to link with other innovative work the Environment Agency was carrying out about how environmental data is published, This was using WIMS as a testbed for exploring new ways to create a public register, and exploring to what degree linked data could be incorporated into the publication of WIMS.

The study examined how the water discharge consent registrations could be mapped to the data specifications under the INSPIRE 'theme',

→ Utility and Government Services.

The water quality monitoring data was mapped to the theme:

→ Environmental Monitoring Facilities

The study identified data fields in these themes that could be mapped to data in the WIMS data set. It was also identified that the consent registration data could be aligned with the other themes:

→ Environmental Management Facilities, and

→ Production and Industrial Facilities.

A copy of the report "CLEAR Info: Publishing in INSPIRE data formats" is contained in Annex 1.

2.0 Findings and conclusions

The Project Team was able to draw the following conclusions from the discussions and exploration involved in undertaking the study, and from the results.

2.1 INSPIRE themes each cover different activities

The WIMS data set contains water discharge consents for businesses undertaking a range of different activities, but the INSPIRE data themes each apply to only a selected type of activity. The result is that CLEAR Info must develop data sets and select only the WIMS data that relates to sites undertaking the specified activities to present under each theme. These are summarised in the following table.

INSPIRE Theme	Activities
Utility and Government Services (Subclass: Environmental Management Facilities)	Waste industry sites Waste water/ sewage treatment
Production and industrial facilities	Industrial sites permitted under European IPPC
Agriculture and Aquaculture	Fish Farms Agriculture/ Farms
Environmental Monitoring Facilities	Observations and measurement of aspects of the environment

The WIMS data set contains a field for the main activity taking place in the site, and this can be used to filter the data to only those in the appropriate groups. This will exclude a proportion of the data the Environment Agency holds.

The data that is required under the INSPIRE standards for each of the themes is different. This means that a user looking for information relating to water discharge consents will need to look across all three themes, and will find the information presented in a different way in each one.

The monitoring of water quality will also be presented under a separate theme, and will therefore disconnect the information about the permit, and the measured impacts of a discharge point, that are currently held together in the CLEAR Info Data Collation tool.

2.2 Mapping the data to INSPIRE is complex

The INSPIRE data specifications did not exactly match the data held in WIMS, and the mapping process required judgements to be made about what WIMS data was the best fit for the INSPIRE specifications. The consultants had to spend a lot of time with the WIMS data experts to ensure they fully understood what each data field contained, in order to be able to propose a set of matches.

This had two sets of implications for CLEAR Info.

First, the mapping of the WIMS data set to INSPIRE standard was a resource intensive process, and the project team did not have the expertise or resourced to repeat this process for all data sets in the Data integration tool. The scope of further work needed to reflect this limitation.

The second implication was wider. Discussions with the data specialists revealed that the process of publishing data for INSPIRE would require further checks to ensure that data held in a number of data sets had been considered before the most appropriate data was matched to fulfil the INSPIRE requirements, and that the Data Custodians (individuals in the Environment Agency who are responsible for safeguarding data and releasing it appropriately) would have to be consulted to ensure they agreed that the most appropriate match has been made. The CLEAR Info project decided not to undertake this assessment of alternative data sets, consultation and decision making because this work will be the responsibility of the Environment Agency as part of its normal activities, in preparing for the Directive publication deadline of 2020.

It was decided that the CLEAR Info project would use the data specifications identified by the consultants to create an example INSPIRE compliant data set, for use in project actions 5.1/5.2, and would recognise that the final solution delivered by the Environment Agency will require further checking and analysis before deciding which data fields best fulfil the requirements of the INSPIRE data specifications.

This will test and demonstrate the ability of the data integration tool to create a company focused data set harmonised to INSPIRE standards.

2.3 INSPIRE is not company focused

CLEAR Info data is focused on linking environmental data to company ownership and collating a picture of performance across a company. It had been imagined that by harmonising a CLEAR Info data set to INSPIRE standards it could be possible to link to similar INSPIRE data sets from across Europe and collate an overview of data for a company operating in a number of European states.

However, the INSPIRE standards seem to be designed for users looking at the spatial location of environmental impacts, rather than at corporate responsibility. The INSPIRE data specifications do not require Company Number, Company name or Parent Company Name to be published for activities taking place on sites. There is a possibility that field for "site name" in the INSPIRE data standards might be populated with a company name in some

cases, but it could be derived from the address, e.g. “Green Road Landfill site”, and so does not provide the reliable corporate identity needed for Company data collation.

If the CLEAR Info project produces an INSPIRE data set using only the fields specified in the INSPIRE standards, this will not enable the linking up of data to create a full picture of a company’s environmental performance across Europe, because INSPIRE does not include the necessary data fields.

The CLEAR Info project has explored the value of company data to:

- regulators
- companies - looking at their own management and assessing their suppliers
- investors and banks
- data literate public

By not serving these user needs the INSPIRE data standards miss the opportunity to maximise the benefits of publishing data in interoperable formats across the European member states.

However, the INSPIRE model can be extended.

2.4 The INSPIRE model can be extended

The INSPIRE technical guidance recognises that national data sets will typically hold data not covered by the INSPIRE data specifications. The Generic Conceptual model applied by the guidance has been designed to be extended, enabling additional data to be included. The CLEAR Info project will produce an extended model to include company ownership information as part of Action 5.1.

A Parent Company cannot be created in the INSPIRE model as a single free standing entity which multiple sites can link to. It would have to be created as a data field that is added as an extension to each site it relates to. The contents of the field could be searched and where they are repeated they could be re-collated as a single company outside of the INSPIRE environment.

In order for the CLEAR Info project or a third party to make the connections between all the sites in the EU owned by a single Parent Company, EU Member States also would need to extend their INSPIRE data in the same way as the CLEAR info project.

INSPIRE therefore provides a building block towards sharing corporate environmental data across EU boundaries because it provides consistent data, but it does not provide the full solution. The barriers and opportunities for sharing corporate environmental data are explored further in Action 5.1 and 5.2.

2.5 The WIMS data can be supplemented

The WIMS data has been mapped to the four most relevant INSPIRE themes, but these themes contain many fields for which WIMS data does not provide a suitable match. Whilst it

is appropriate for the CLEAR Info data set to return a “void” value for most of these unmatched fields, there are a small number where it would be reasonable to derive a result from the available data. For example, where details of a permit are given, INSPIRE requires a description of the permission type. WIMS does not contain a description of the type of permit but as this would provide useful context to the data, the CLEAR Info data integration tool could generate a description, such as ‘discharge to the water environment’. Another example is the requirement in a number of fields for a European industrial activity classification (NACE) code for the economic activity taking place. WIMS does not contain this code, but it would provide useful context to the data.

The CLEAR Info project team will explore the degree to which these additional data fields can be added to the INSPIRE data set through use of the CLEAR Info data integration system under Action 5.1 and 5.2.

3.0 Findings and next steps

The key findings from this action are:

- Transforming data to INSPIRE standards is a resource intensive activity which requires data experts to work hand in hand with INSPIRE specialists.
- INSPIRE does not serve the user needs of regulated businesses or investors interested in a company’s environmental performance.
- Extending the conceptual model may significantly improve the usefulness of the data but, as this is optional, it may not provide a complete data set across Europe. It would be important to share the value of doing this with other regulators in order that they adopt the same approach.

It is proposed that the INSPIRE mapping will be used to:

- Design a process for transforming the CLEAR Info WIMS data to INSPIRE standards.
- Produce an extended Conceptual Model to include company information.
- Produce a data set and registry tables/ code lists, as appropriate.

The harmonised data set and the findings from this report will be used to engage stakeholders from Defra, EC’s Joint Research Centre and other European regulators, and to explore the feasibility of linking data across Europe, for action 5.1 and 5.2