

Referencing INSPIRE in environmental reporting

A blueprint on the use of INSPIRE data models in reporting

Version 1.0, November 2018



Acknowledgements

This document has been prepared by the European Environment Agency (EEA) assisted by the external contractor Epsilon Italia.

Contents

Acknowledgements	2
Contents	3
Acronyms.....	4
Executive Summary	5
1. Introduction.....	7
2. The “Linked Approach”	12
2.1. Methodology	12
2.2. The use case of the 2018 CDDA reporting exercise	14
2.3. Lessons learned	19
3. INSPIRE extensions approach	22
3.1. Methodology	22
3.2. The use case of the EU Registry on Industrial Sites.....	23
3.3. Lessons learned	27
4. Blueprint for the re-use of INSPIRE data models in environmental reporting data flows.....	29
4.1. Technical and organisational cooperation	29
4.2. Key questions to help guide future INSPIRE-based reporting modelling.....	30
5. Conclusions.....	34

Acronyms

CDDA	Common Database on Nationally Designated Areas
CDR	Central Data Repository
EC	European Commission
EEA	European Environment Agency
EIONET	European Environment Information and Observation Network
E-PRTR	European Pollutant Release and Transfer Register
ETC/BD	European Topic Centre on Biological Diversity
EU	European Union
EU MS	European Union Member States
GML	Geography Markup Language
IED	Industrial Emissions Directive
INSPIRE	Infrastructure for spatial information in Europe
INSPIRE MIG	INSPIRE Maintenance and Implementation Group
INSPIRE PF	INSPIRE Production and Industrial Facilities
INSPIRE PS	INSPIRE Protected Sites
LCP	Large Combustion Plants
ROD	Reporting Obligations Database
QA/QC	Quality assurance / quality control
UNEP-WCMC	United Nations Environment World Conservation Monitoring Centre
UNESCO	United Nations Educational, Scientific and Cultural Organization
WDPA	World Database of Protected Areas
WFS	Web Feature Service
WI	Waste incineration and co-incineration plants
XML	Extensible Markup Language

Executive Summary

The European Environment Agency (EEA) document *“Referencing INSPIRE in environmental reporting”* aims in the first place at describing the practical approaches for reusing INSPIRE data models in environmental reporting in the context of two operational data flows run by the EEA (Nationally designated areas –CDDA- 2018 reporting exercise and the European Union (EU) Registry on Industrial Sites). The document outlines a well the main questions which, based on these experiences, should inform the selection of the most suitable technical approach, taking into consideration that the most appropriate method can only be chosen on a case-by-case basis.

The report starts with a summary of the main policy drivers behind these activities, mainly the streamlining process triggered with the fitness check of environmental reporting that was launched by the European Commission (EC) back in 2015. A subsequent Action Plan published in June 2017 points out to the INSPIRE Directive as a tool with a strong potential for streamlining the reporting process, improving efficiency, effectiveness and coherence in the context of the management of the spatial data included in the reporting obligations. A pragmatic and stepwise approach to transition from the current reporting process to a more effective and efficient one using INSPIRE was described in the EC and EEA concept note *“The future of eReporting and the link to INSPIRE”* also released in spring 2017. The present report is therefore a follow up to that concept note, as it describes the actual application of the data typology and technical approaches spelled out in that note using real life examples.

The report describes in detail the two approaches used to relate INSPIRE data and environmental reporting (“business”) data in the data flows of CDDA and the EU Registry on Industrial Sites. The first case builds on the so-called “Linked Approach”, which has its foundations in the first scenario described in the INSPIRE Generic Conceptual Model addressing existing and well-established data flows. The “Linked Approach” reuses “as is” the already existing INSPIRE data, while the data required by a specific obligation but not required by INSPIRE data models is to be included in separated (but linked) thematic datasets. The linking mechanism between the two datasets is established both at object and dataset level. Several lessons learned can be drawn from the application of this methodology in the 2018 CDDA reporting exercise, in particular the importance of a consistent implementation of the linking mechanism at both technical and organisational levels.

The EU Registry on Industrial Sites data model, the second example covered by this report, was created by applying the general rules for extending an INSPIRE data specification (in this case, the Production and Industrial Facilities –PF- data model) as specified by the INSPIRE Generic Conceptual Model. While the document does not enter in details on how to proceed technically with the different aspects of a generic data model extension, it does describe the specific steps taken during the EU Registry modelling process as well as how the link with other thematic reporting obligations is envisaged. A few lessons learned have also been summarised in the text, among them the appropriateness of using this approach in cases such this one where a new reporting data model is very close, yet not identical, to an INSPIRE one.

Building on lessons learned, this report then elaborates on the main questions which may drive the selection of the best-fitting methodological solution for a reporting workflow based on INSPIRE datasets and provides, where possible, recommendations for mitigating potential issues that may arise (the “blueprint”). A key aspect highlighted in this blueprint is the need for cooperation between INSPIRE and policy working groups for reporting in order to ensure that the model agreed meets the demands from both communities while caters for an efficient, streamlined reporting process. A thorough comparative analysis of the requirements stemming from the particular reporting obligation and relevant INSPIRE data specifications is of course the first step for any INSPIRE based e-reporting development. While not a fixed recipe, the report further proposes a few questions that can help guide future technical developments:

- How similar are the requirements stemming from the thematic reporting obligation and relevant INSPIRE data specifications?
- Are other reporting obligations referencing the same INSPIRE spatial data? And how similar are these reporting obligations?
- How many INSPIRE data specifications are relevant to the reporting data flow?
- How complex could an extended INSPIRE data model become?
- Are there elements that can act as a link between INSPIRE and thematic data, e.g. a thematic identifier? How stable are the thematic reporting requirements.

The document concludes with a final reflection on the expected benefit of streamlining reporting processes reusing INSPIRE. This should ultimately result in less administrative burden in the Member States, as well as in an increase in coherence and consistency, since multiple provisions of the same data will be avoided. Furthermore, by making the geospatial datasets from these reporting obligations INSPIRE compliant, Member States would be as well in a good position to fulfil the relevant INSPIRE obligations. It is however important to note that the INSPIRE obligations go beyond the harmonisation of data models, which is the main focus in this document.

An important caveat concerning this report is referred to not only in the concluding notes but throughout the whole text. The examples provided in this report have not yet embraced the full potential of the INSPIRE web service infrastructure for e-Reporting since they have primarily focused on the reuse of INSPIRE data models. The need to address technical and legal challenges before moving to an operational reporting process based on web service harvesting is one of the main reasons, together with the lack of an homogeneous implementation of INSPIRE across EU Member States. Nevertheless, the steady progress in the implementation of INSPIRE and the recently launched project aiming at modernising the EEA reporting platform (Reportnet 3.0), shall facilitate a transition in the long term to a future reporting system which can harvest INSPIRE web services to collect geospatial datasets relevant for thematic reporting obligations.

1. Introduction

A **reporting obligation** is a legal provision requiring that Member States submit data, information or reports to the European Commission (EC) or the European Environment Agency (EEA), which are needed to check if a particular regulation is effectively meeting its objectives. This is an essential element of the European Union (EU) policy cycle, providing essential facts and information for an informed decision-making.

Developed by the EEA, and operational since 2002, **Reportnet**¹ is the IT infrastructure currently supporting environmental data and information flows linked to reporting obligations. It provides a set of common tools and a technical infrastructure for defining, collecting, transferring, editing, analysing and streamlining environmental data. It provides a framework of data standards, applications and interoperability mechanisms that enable information exchange and sharing within and between information systems. Since it was designed for an earlier generation of EU environment policy, it is now under great strain due to the volume of data now being reported.

In May 2015, in the context of the Better Regulation package², the EC launched a **fitness check on environmental reporting**³. The ultimate aim of this initiative was to develop a more modern, efficient and effective regulatory monitoring and reporting which would better inform the European public about the achievements of the EU law on the ground while simplifying the reporting burden for national administrations and businesses. While the fitness check found that current environmental reporting was largely fit-for-purpose, an Action Plan⁴ was published in June 2017 addressing five areas with potential for further streamlining of the reporting obligations and processes. Among these areas, one concerns the actual reporting process, as more can be done in order to make full use of the ongoing advances in information technology tools, including the standardisation of practices and procedures. Actions underpinning this area are the **modernisation of Reportnet** (Reportnet 3.0), so it can become the central and streamlined EU tool for reporting, as well as the development and testing of tools for **data harvesting**, facilitating the EU level to have better and more flexible access to Member States' data while minimising administrative burden.

The Action Plan also points out to the **INSPIRE Directive**⁵, laying down the general rules establishing the infrastructure for spatial information in the European Community, as a tool with a strong potential for streamlining the reporting process, improving efficiency, effectiveness and coherence of spatial data management. INSPIRE does not require to collect new data but applies to already existing or any newly created data and aims at improving the exchange and re-use of environmental data across Europe.

¹ <https://www.eionet.europa.eu/reportnet>

² "Better Regulation for better results – An EU agenda" (COM(2015) 215)

³ http://ec.europa.eu/environment/legal/reporting/fc_overview_en.htm

⁴ http://ec.europa.eu/environment/legal/reporting/pdf/action_plan_env_issues.pdf

⁵ <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:108:0001:0014:EN:PDF>

One of the original use cases when designing INSPIRE (although not the only one) was actually to facilitate reporting, and that is why many of the themes listed in the Annexes of the INSPIRE Directive refer to relevant geospatial data stemming from EU environment legislation. However, while over the years the use of INSPIRE data specifications has been increasingly discussed in many reporting areas, the alignment of the reporting data flows with INSPIRE has only been achieved to a large or lesser degree depending on the area (e.g. air quality, protected areas, industrial emissions, marine protection, water management). Reasons behind this may be, among others, the differences in implementation timescales of INSPIRE and the thematic obligations, the perceived complexity of the INSPIRE data models and the lack of resources and capacity in Member States that has resulted in a patchy implementation of the INSPIRE Directive across Europe.

Along the lines of the Action Plan abovementioned, the joint EC and EEA concept note “**The future of eReporting and the link to INSPIRE⁶**”, released in May 2017, outlines the vision on the future use of INSPIRE in environmental reporting. The document describes a pragmatic and stepwise approach to transition from the current reporting process to a more effective and efficient one using the INSPIRE approach as a key element to achieve this.

In accordance with the INSPIRE Directive, EU Member States (EU MS) are required to:

- **Identify** the relevant spatial datasets,
- Ensure that the identified datasets are documented using certain standards (**metadata**),
- Make the identified data accessible and available through a number of different **web based services** (discovery, viewing, download), and
- Transform the identified datasets into predefined **data models** for the purpose of increased interoperability.

The definition of spatial datasets addressed by the INSPIRE Directive covers a wide spectrum of environmental (and other) data, from geographic reference points (e.g. location of monitoring station) to the environmental data being collected (e.g. concentration of a specific pollutant in the environment). At the same time, **most or all information reported under EU environmental legislation has a geospatial component, overlapping therefore with the INSPIRE scope.**

The concept note distinguishes between the following three groups of data in any reporting obligation (figure 1):

- The **geospatial reference data** (data type 1) which enable to prepare maps or data viewers and can build on INSPIRE data models.
- The **environmental or other business data** (data type 2) such as attributes of the geospatial reference data which are not necessarily covered by the INSPIRE data models.
- The **textual and contextual data** (data type 3) such as permits, exemptions, methodologies, waste management plans, etc.

⁶<https://ies-svn.jrc.ec.europa.eu/attachments/download/2098/eRep%20and%20INSPIRE%20concept%20-%20rev%20-%20May%202017.pdf>

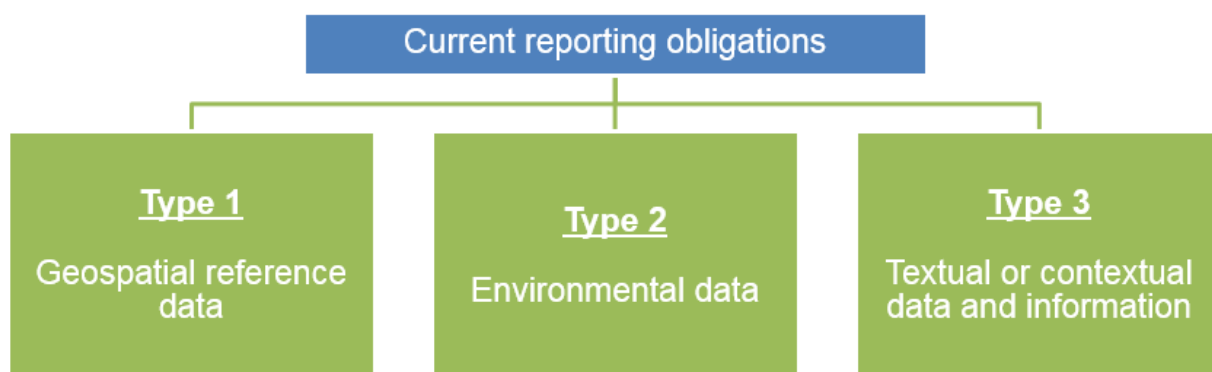


Figure 1: Main data types expected in a generic reporting data flow

In a generic reporting exercise covering at least data type 1 and 2, the competent authorities need usually to rely on the input from a wide range of data holders, involving *inter alia* a heavy processing to turn the spatial data into the agreed format and content. There is a substantial scope to optimise such data transfers and processing by harmonising data structures at the source. Furthermore, spatial data held by a public authority are often needed for more than one reporting activity since the geographical scope of many legal acts overlap. If available through the INSPIRE infrastructure, the relevant geospatial datasets could eventually be harvested online by the corresponding reporting authorities whenever a new report is due, optimising the data flows from different organisations for EU level reporting purposes.

In order to design successful and efficient environmental reporting, these three types of data need to be identified and marked in each reporting obligation. This data typology allows to clearly identify for each reporting obligation which element could eventually be expected from the national INSPIRE infrastructures and which elements needs to be provided in addition to INSPIRE.

The link between the data from type 1 (geospatial reference data typically covered by INSPIRE) and type 2 and 3 (business data and other contextual data) can actually be established in different ways. The **INSPIRE Generic Conceptual Model**⁷ devotes a whole section to the relationships between INSPIRE data and data related to environmental reporting⁸. Given the variety of reporting obligations and data flows, it formulates the principle that where there are existing and well-established data flows for reporting data, INSPIRE data models should be limited to providing spatial objects and attributes that allow joining the reporting data to the spatial objects – via an external object identifier or a thematic identifier. For newly defined data flows instead, INSPIRE data models may also include attributes covered by the thematic legislation, so that in the future INSPIRE compliant data will be able to support the reporting process. Consistency between reporting and INSPIRE obligations should be sought to avoid duplication of Member States efforts.

The following sections of this document therefore elaborate on the technical approaches applied to establish this relationship between spatial and business data in the context of two operational

⁷ <https://inspire.ec.europa.eu/documents/inspire-generic-conceptual-model>

⁸ Section 6.2, pp 33-34.

reporting workflows managed by the EEA: the database of **Nationally Designated Areas (CDDA)** and the **EU Registry on Industrial Sites**.

Nationally Designated Areas (CDDA)⁹ is an annual voluntary data flow for which the EEA member and cooperating countries have committed to report on areas designated under national legislation for the purpose of nature protection, including sites such as national parks and nature reserves. It is the official source of protected area information from European countries to the World Database of Protected Areas (WDPA).

The **EU Registry on Industrial Sites**¹⁰ is a data flow that collects information on the facilities, installations, and plants, which are obliged to report under the European Pollutant Release and Transfer Register (E-PRTR) Regulation and the Industrial Emissions Directive (IED). This obligation finds its legal basis in the Commission Implementing Decision 2018/1135/EU¹¹.

These two practical examples demonstrate the potential of INSPIRE in the context of environmental reporting, facilitating the re-usability of national geospatial datasets while allowing for flexibility in the development cycles of reporting and INSPIRE.

Both developments are underpinned by the following basic assumptions:

- EU MS shall provide data in the scope of INSPIRE in a harmonized way via network services and metadata by 2020 at the latest. This includes the data supporting the environmental reporting, which countries are asked to implement as priority following the fitness check of environmental monitoring and reporting¹².
- EU MS shall avoid multiple provision of the same geospatial data (e.g. to INSPIRE and in parallel as part of the environmental reporting obligations).
- To better support environmental policies and decision making, a more direct, frequent and timely access to the environmental data provided by the EU MS is needed.
- While acknowledging the specific nature of each environmental reporting process, the environmental reporting process should be designed in a pragmatic yet generic way, so it could eventually be re-used without introducing high demands on provider and consumer side.

Each reporting data flow described in this report has followed a different implementation approach to link geospatial reference (type 1) with business data (type 2). In the case of CDDA, the methodology is based on the actual split between the data required by the INSPIRE legal framework (Protected Sites spatial data theme) and those data outside of it, keeping the latter in external reporting files linked to the INSPIRE resources ("**Linked Approach**"). In the case of the EU Registry on Industrial Sites, the reporting model extends a single INSPIRE theme (Production and Industrial

⁹ <http://rod.eionet.europa.eu/obligations/32>

¹⁰ <http://rod.eionet.europa.eu/obligations/721>

¹¹ <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A32018D1135>

¹² <https://ies-svn.jrc.ec.europa.eu/projects/2016-5/>

Facilities spatial data theme) including the specific requirements stemming from reporting (**“Extension method”**).

It is important to note that the practical implementation of these two data flows has not yet at this point in time embraced the full potential of INSPIRE infrastructure, namely shifting from a central data submission to a new reporting paradigm based on **active dissemination** and INSPIRE services **harvesting**. While there is an increasing available of network services, and in particular download services, the uneven implementation of INSPIRE across Europe still prevents this paradigm shift from happening in the short or medium term. Furthermore, challenges in operational settings of technical and legal nature need to be addressed first before this new approach is taken up in an operational stage. The activities related to the list of priority datasets for environmental reporting and data harvesting, as well as the modernising of Reportnet are currently contributing to overcome these issues.

It is therefore in this context that the present document aims at:

- **describing the practical approaches for reusing INSPIRE in environmental reporting**, including an analysis of main advantages and disadvantages, strengths and risks of these methodologies based on lessons learned from real life examples, i.e. CDDA reporting 2018 and the EU Registry on Industrial Sites.
- **outlining the main questions driving the selection of the most suitable approach** (i.e. a blueprint), building on the basic assumption that the most appropriate method can be selected only on a case-by-case basis, starting from a comparative analysis of the e-reporting obligations and relevant INSPIRE data models.

2. The “Linked Approach”

This section focuses on the description of the “Linked Approach” as a technical methodology to re-use INSPIRE data in environmental reporting. It starts with an outline of the technical solution from a generic point of view. It follows with the description of how the methodology has been implemented in the case of the CDDA 2018 reporting exercise, concluding with a summary of the lessons learned, challenges and opportunities of the solution adopted.

2.1. Methodology

In the previous section, it has already been mentioned that the data required by a particular reporting obligation can be categorised in three different types: type 1 (geospatial reference data), type 2 (environmental or other business data) and type 3 (textual and contextual data). As geospatial reference data, type 1 data may typically fall under INSPIRE scope, creating an overlap, which, if not addressed, could result on duplication of Member States efforts.

As indicated above, the relationship between INSPIRE data and environmental reporting had already been foreseen during the development of INSPIRE technical specifications, and the INSPIRE Generic Conceptual Model presents a series of scenarios aiming at streamlining the use of INSPIRE in the context of environmental reporting. The so-called “**Linked Approach**” has its foundations in the first scenario described in the Generic Conceptual Model, namely that where there is already an existing and well-established data flow for reporting data. In this case, it is recommended that INSPIRE data models are limited to providing spatial objects and attributes that allow joining the reporting data to the spatial objects, such as external object identifiers or thematic identifiers.

Therefore the “Linked Approach” does not extend INSPIRE core data models in order to include new pieces of information or new constraints stemming from environmental legislation. Conversely, it aims at re-using “as is” the already existing INSPIRE data in order to get spatial reference and, potentially, other pieces of information relevant to environmental reporting purposes. All data required by a specific reporting obligation and not available in the related INSPIRE datasets will be included in thematic datasets for the specific reporting.

From a data model point of view this implies that, while the INSPIRE data model is not modified, the data model for the specific reporting takes into consideration only those data elements that are not present in the corresponding INSPIRE data model (and therefore are not available in INSPIRE data). **Of course the thematic data model shall contain all the elements needed to retrieve the necessary INSPIRE data, i.e. the link between the geospatial data (type 1) and the thematic reporting (business) data (type 2).**

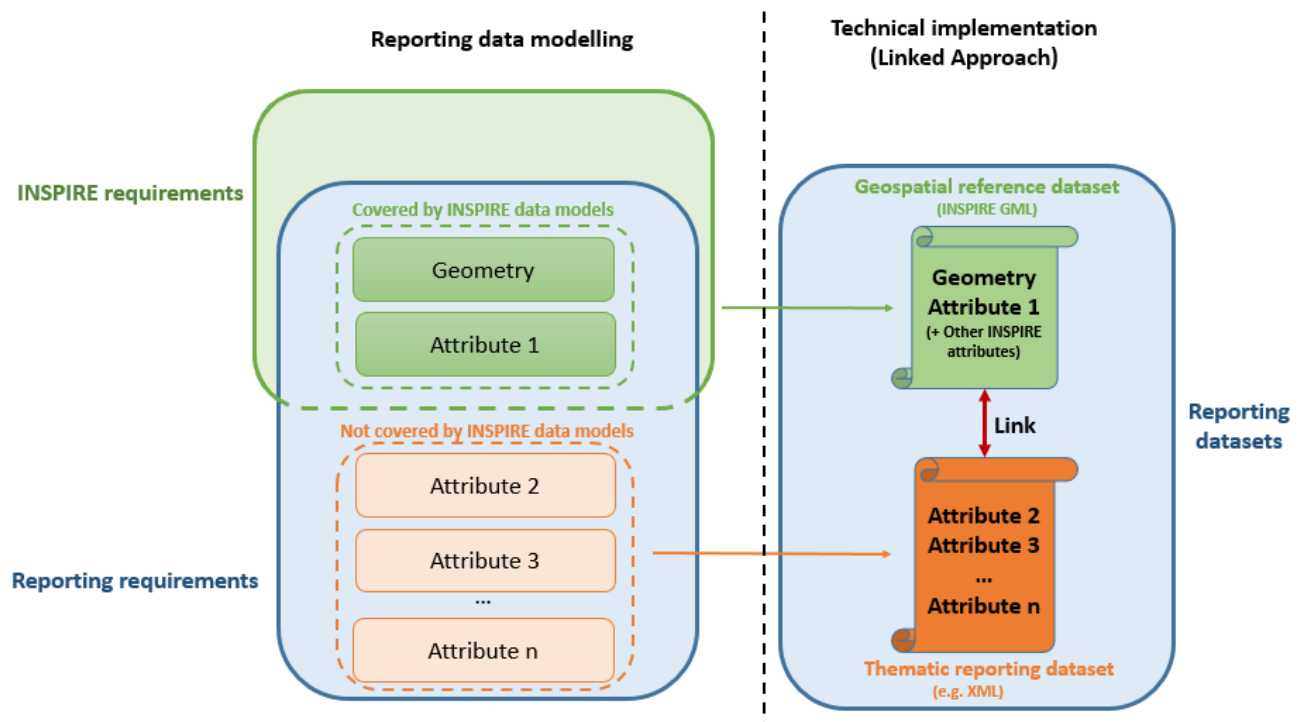


Figure 2: The “Linked Approach” in reporting datasets

As indicated in the figure 2 above, a certain reporting obligation (in blue) overlaps with (some) INSPIRE requirements (in light green). Those common elements (in dark green), among them typically the geometry of the objects, shall then be modelled and provided according to INSPIRE (i.e. described by metadata, harmonised according the INSPIRE data models and available through INSPIRE web services). Elements not covered by INSPIRE (in orange) shall then be provided separately, usually in a plain XML file, avoiding duplication of information in both datasets but ensuring the linking between the two datasets.

The linking mechanism can be established on two levels:

- the **dataset level**, in order to ensure that there is a link between the thematic reporting dataset and the geospatial dataset which provides the spatial reference information. This link can be implemented by pointing out to the corresponding **download service endpoint**¹³ (typically a Web Feature Service –WFS–), or, alternatively, by using the **file name of the GML file**¹⁴.
- the **object level**, in order to ensure that every single record in the thematic dataset points to the corresponding record in the relevant geospatial dataset (which has already been

¹³ A service end point is an URL used for directly calling an operation provided by the service. In turn, a service access point is an URL that can be used to retrieve a detailed description of a service (e.g. GetCapabilities), which typically includes a list of end points to allow its execution. It is therefore necessary that the link referred to here points to the end point and not to the access point.

¹⁴ Storing the GML file name is a specific measure implemented during the CDDA 2018 reporting exercise as the reporting system was not yet ready to harvest INSPIRE download services in an operational setting (see 2.2).

referenced at dataset level). This link is established by referencing an **external object identifier** (e.g.: inspireid) or **thematic identifier** of a particular geospatial object in the corresponding record of the thematic reporting (business) dataset, facilitating the link between the two types of information for a specific object.

Figure 3 below shows this concept for the particular case of the CDDA reporting.

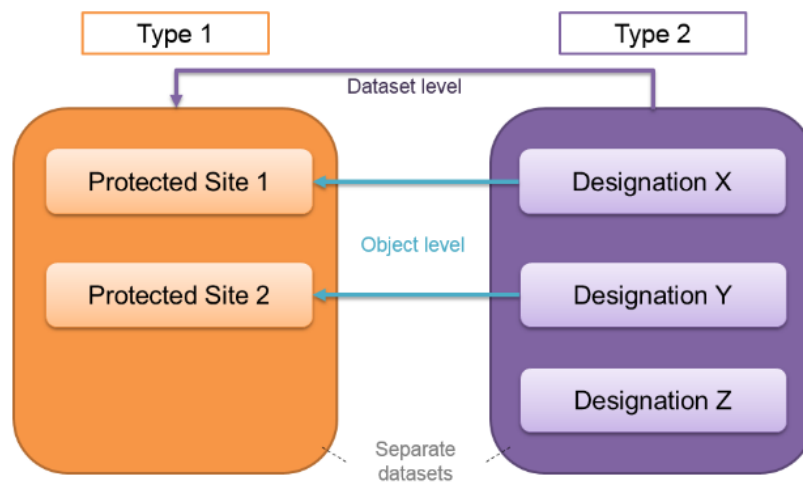


Figure 3: Two linking levels in the CDDA case

While type 1 data shall follow the INSPIRE data model, encoded according to INSPIRE rules and, ideally, provided through web services, type 2 data may be just implemented in a simple format such as plain XML. The only condition, again, is to ensure that each record in the type 2 data file contains a reference to the corresponding INSPIRE object described in the type 1 dataset.

2.2. The use case of the 2018 CDDA reporting exercise

A clear example of a “Linked Approach” implementation is the data model developed for the provision of the **Nationally Designated Areas inventory** in its 2018 reporting exercise. This inventory, formerly known as the Common Database on Designated Areas (CDDA), holds information about designated areas and the national legislative instruments, which directly or indirectly create protected areas in the European Environment Information and Observation Network (Eionet) partner countries.

The CDDA data is delivered by each country as thematic information and spatial information. The CDDA is an agreed annual Eionet core data flow maintained by the EEA with support from the European Topic Centre on Biological Diversity (ETC/BD). The dataset is used by the EEA as well as the United Nations Environment World Conservation Monitoring Centre (UNEP-WCMC) for their main assessments, products and services.

The reporting obligation for the CDDA is defined in the Reportnet's Reporting Obligations Database (ROD)¹⁵. It is a voluntary data flow. Eionet partner countries (EEA member and cooperating countries, 39 in total) have committed to report on areas designated under national legislation for the purpose of nature protection including sites such as national parks and nature reserves (other types of sites may also be reported). Being an annual exercise, the CDDA reporting date is fixed to 15 March every year.

Of the 39 countries reporting to CDDA, 28 EU MS are legally mandated from the INSPIRE Directive to provide some of the CDDA reporting elements following the INSPIRE specifications, in particular the spatial data theme **Protected Sites** (INSPIRE PS)¹⁶. Furthermore, CDDA reporting asks for sites (e.g. sites with statutes such as management schemes) which may have no formal designation in the national regulation and therefore may not be covered by the INSPIRE PS theme. This is presented in a simplified way in the figure 4 below:

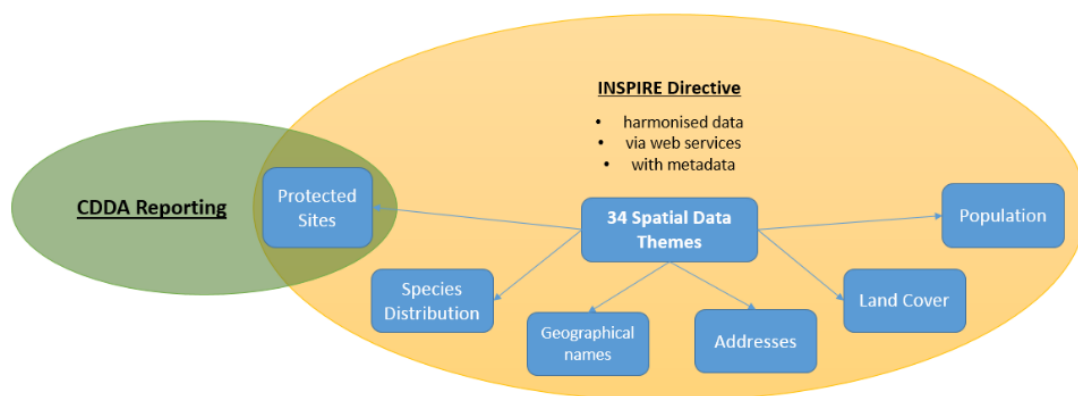


Figure 4: Overlapping between CDDA reporting and INSPIRE obligation

The INSPIRE PS theme is covered by the Annex I of the INSPIRE Directive. Therefore, by the time of the 2018 CDDA reporting deadline, the EU MS were already obliged to provide the corresponding datasets fully harmonised according to the INSPIRE data specifications as the deadline was in November 2017.

In this sense, the need to avoid double reporting and address these two overlapping obligations, underpinned by a process with the Eionet partner countries aiming at reviewing and removing outdated elements from the existing CDDA reporting, led to the development of a new CDDA data model for the 2018 monitoring exercise. The CDDA reporting model was designed in a way that the common elements with INSPIRE would now come from, and be defined by, the INSPIRE Protected Sites data model (**type 1 data**), and the rest would then be reported separately in a tabular format (**type 2 data**) as it had been done before 2018. Therefore, the complete CDDA data flow consists of one (or more) INSPIRE dataset(s), in GML format, plus a thematic dataset, eventually encoded using plain XML.

¹⁵ <http://rod.eionet.europa.eu/obligations/32>

¹⁶ <https://inspire.ec.europa.eu/Themes/117/2892>

It is also important to acknowledge that **not all elements under the INSPIRE PS data model are relevant for CDDA** and only a few of them were considered in the reporting model¹⁷. However, the ultimate objective is to ensure that EU MS are re-using the data that should already be provided through and harmonised according to INSPIRE. The figure 5 provides the CDDA 2018 reporting conceptual model, the '**Type 1 data**' (orange elements) modelled according to the INSPIRE Protected Sites data model, with the CDDA relevant fields in red, and the '**Type 2 data**' (lilac fields) being modelled separately. The dotted lines represent the "**linking mechanism**" which, as already explained in Sec. 2.1, operates on the 'dataset level' and on the 'record level'.

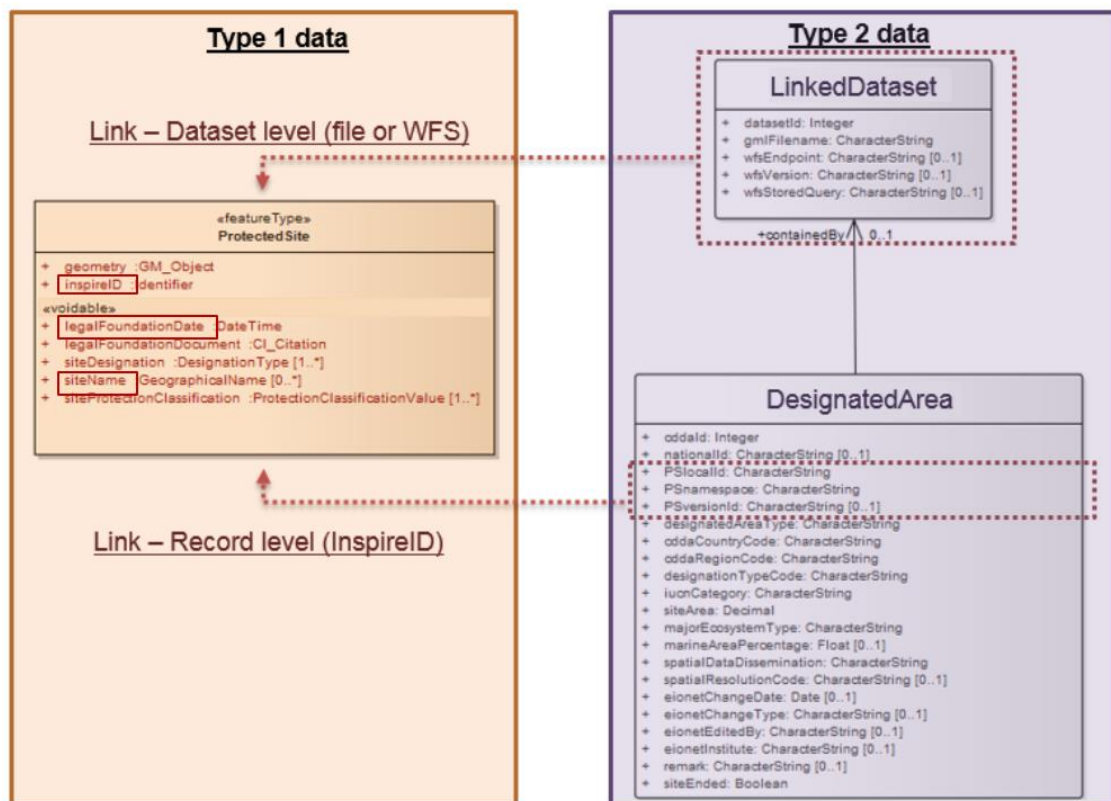


Figure 5: CDDA 2018 reporting conceptual model

The type 1 dataset, modelled according to the INSPIRE Protected Sites theme, is a collection of «**ProtectedSite**» GML features, each of them uniquely identified by their *inspireId*.

The type 2 dataset, instead, is built up by two tables, **Designated Area** and **Linked Dataset**, and encoded in XML. The **Designated Area** table is the main table containing, per each row (which corresponds to a specific designated area) those attributes that are specific for the CDDA reporting and do not belong to the INSPIRE PS data model¹⁸. Apart from these attributes, and in order to

¹⁷ In particular: geometry, inspireId, legalFoundationDate and siteName. The attribute siteProtectionClassification is expected to take the value "natureConservation".

¹⁸ With the exception of the IUCN category. This category is included in the siteDesignation data type of the INSPIRE PS data model, but the reference code list used by INSPIRE is different than the one recommended by UNEP-WCMC and

ensure the link at object level between each feature in the GML file and the corresponding designated area characteristics, the **inspireId** attribute is also included in the Designated Area table (see “Link – Record level” in the figure above). Each designated area in the CDDA dataset therefore incorporates the information on the *inspireId* triple (*localId*, *namespace* and *versionId*) of the relevant INSPIRE PS data model in the *PSlocalId*, *PSnamespace* and *PSversionId* attributes of the table.

The second table, **LinkedDataset**, has been included in the model in order to ensure the connection of both data types on the dataset level. In the 2018 CDDA reporting exercise, this link has been established by providing the name(s) of the GML file(s) which contain the relevant spatial information, since the harvesting of INSPIRE download services was actually not operationally possible. Member States needed to provide instead one or more INSPIRE GML files which, together with an XML file containing the CDDA thematic information, were uploaded to the Central Data Repository (CDR) of Reportnet¹⁹. However, and in order to test the possibility of using download services in the future, national reporters were requested to optionally provide some information about the endpoint of the INSPIRE download service (e.g. WFS) from which the INSPIRE dataset could be harvested.

In the CDDA 2018 reporting exercise, well described in the corresponding reporting guidelines²⁰, particular attention has been paid to the need to ensure **proper coordination at national level** for the provision of the type 1 dataset. This is important as the CDDA reporter is typically not in charge of the provision of INSPIRE PS theme datasets at national level. Coordination with those publishing CDDA datasets according to INSPIRE is essential to avoid duplicating efforts and ensuring coherence in the whole exercise. However, this may not be straightforward and it may be difficult to identify the specific INSPIRE datasets to use among those published by those responsible for INSPIRE implementation in the country. The figure 6 below had been included in the CDDA reporting guidelines explaining the process for creating the GML file that contains the type 1 data.

In the case the CDDA dataset is identified and provided fully harmonized and available through INSPIRE services, it is strongly recommended to reuse it (even if for this reporting exercise the GML data shall be downloaded from the service and uploaded in Reportnet). In the best-case scenario, a specific web service would only serve the CDDA dataset. However, as the INSPIRE Protected Sites data theme covers different types of protected areas - e.g. Natura2000, Ramsar, UNESCO sites - a single INSPIRE PS dataset might contain several typologies of sites. In this case, only a subset is relevant for CDDA reporting and it would then be necessary to filter the dataset before its delivery (which might imply some manual work). Besides this filtering, the INSPIRE dataset should remain unchanged (attributes and geometry).

used in WDPA, which is used traditionally by CDDA. Therefore, this attribute needed to be included again in the type 2 dataset despite being part of the INSPIRE data model.

¹⁹ <http://cdr.eionet.europa.eu/>

²⁰ A detailed description of structure, requirements and tools for production of the CDDA 2018 reporting data can be found in the relevant guidelines: <http://cdr.eionet.europa.eu/help/cdda/CDDAv16%202018%20guidelines%20v1.1.pdf>

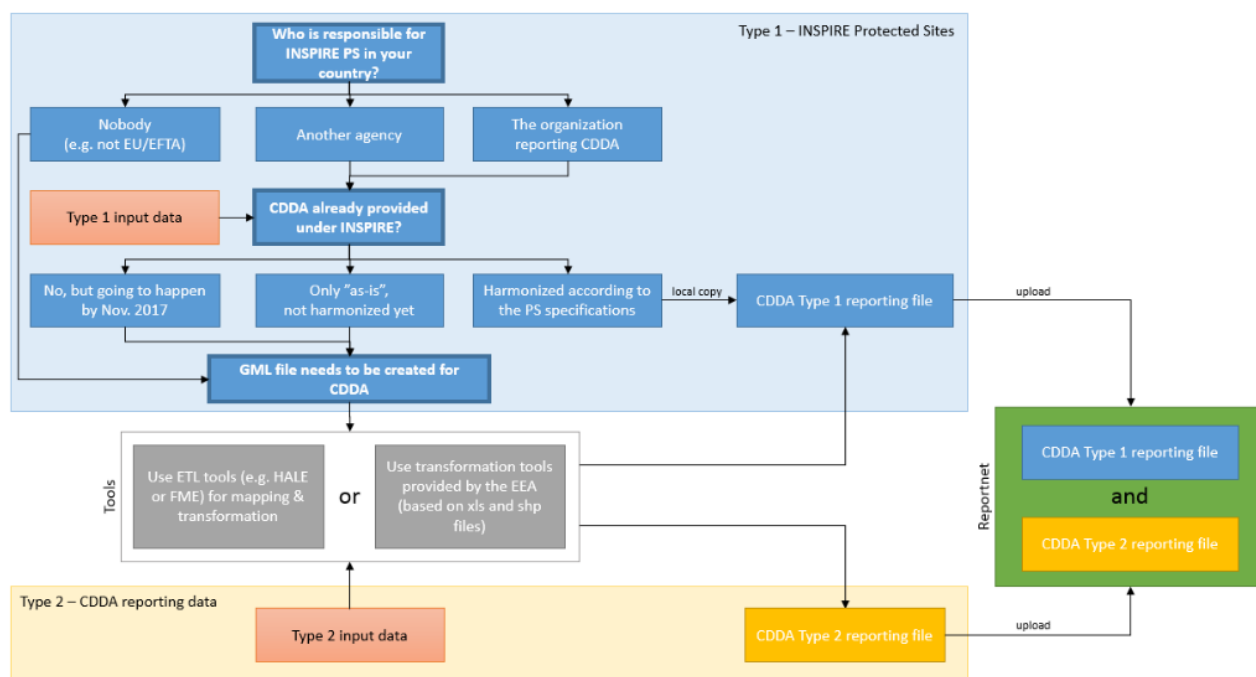


Figure 6: Recommended workflow to create CDDA reporting dataset

In case CDDA is still not provided fully harmonised, it should be transformed to meet the INSPIRE PS data specification requirements. A set of tools has been provided to the national reporters to do so. However, and in order to ensure this exercise aligns with the national INSPIRE strategy, it is again strongly recommended that the CDDA reporter making the transformation acts in coordination with the responsible organisation for INSPIRE in the country. The CDDA guidelines stress that providing INSPIRE datasets in the context of CDDA reporting does not substitute the obligations coming from INSPIRE, meaning that the INSPIRE PS dataset containing the CDDA data still needs to be described by metadata, be accessible via view and download service and available through the INSPIRE Geoportal²¹.

The production of **Type 2 data**, which falls typically only under the remit of the national CDDA reporter, should also consider a proper management of the identifiers that allow connecting this data with its spatial reference. As already explained, the INSPIRE identifier (inspireId) is the element that ensures the link on the 'record level' between the thematic dataset and the relevant INSPIRE dataset and needs to be included in the corresponding elements of the type 2 data file. **The inspireId, when existing, shall not be changed.** However, and similarly to the case of the correct identification of Type 1 data, when the organisations in charge for the reporting are not responsible for the INSPIRE implementation the inspireId values may not be known to CDDA reporters, making difficult the generation of the type 2 dataset. Should the national INSPIRE PS dataset not exist elsewhere, and neither a national policy on how to build INSPIRE identifiers, the CDDA guidelines provides a suggestion which identifies the *localId* with the site code (cddaId in the type 2 dataset) and proposes a specific *namespace* for CDDA.

²¹ <http://inspire-geoportal.ec.europa.eu/>

The CDDA reporting guidelines also cater for the scenario where there are several INSPIRE PS datasets of subnational level provided by one country, i.e. federal states or in countries with overseas territories.

2.3. Lessons learned

The CDDA 2018 reporting exercise represents the first operational example in the EEA of the “Linked Approach” for re-using INSPIRE in environmental reporting processes. As mentioned before, any reflection on the lessons learned of this specific implementation should take into consideration that the current reporting infrastructure (Reportnet) is not ready at this time to leverage the full potential of the “Linked Approach”, namely **to obtain the spatial reference dataset by directly harvesting of INSPIRE download services** (i.e. avoiding uploading the GML file physically). Furthermore, not all countries were ready to provide fully compliant INSPIRE web services during this reporting exercise. The 2018 implementation of the CDDA reporting is therefore the result of a compromise between the need to build on INSPIRE investments at national level avoiding duplication of efforts, and the early stages in the transition from active reporting towards a system which may be based on continuous, active dissemination by countries.

Considering all this, several lessons learned can be extracted from this activity, which are summarised below:

- **On the re-use of INSPIRE core data models:**
 - The “Linked Approach” allows **reusing INSPIRE core data models**, in this case INSPIRE PS, while encoding the specificities of a particular reporting obligation separately in thematic data models. As customisation remains external to the INSPIRE data model, countries can simply reuse “as-is” what has already been produced to fulfill INSPIRE obligations. Vice-versa, the same INSPIRE dataset can also be used in different reporting obligations or even different contexts, maximising its usability and minimising costs.
 - The use of this approach avoids what could eventually become a complex **data model extension** exercise as it allows designing flat structures for the reporting of specific thematic data, encoded in plain XML. Type 2 datasets are therefore exempted from the difficulties related to the consumption and use of GML complex features, typical of the INSPIRE data models, and only contains primitive type elements (namely `CharacterString`, `Integer` and `Boolean`).
- **On the importance of linking mechanism and identifier management:**
 - A major challenge in the technical implementation concerns the **linking mechanism** itself, ensuring the availability of the necessary information to create the links at the ‘dataset level’ (i.e. how to know which INSPIRE dataset contains the ‘official’ relevant information) and at the ‘record level’ (i.e. how to match the corresponding objects in thematic and INSPIRE data). With no proper implementation of the linking

mechanism there is no possibility to connect both data types, rendering the “Linked Approach” completely meaningless. This can only be mitigated by an increased cooperation and transparency among INSPIRE stakeholders.

- An important issue arising from the reuse of the INSPIRE data model in the CDDA reporting is that **the INSPIRE PS data model does not contain any thematic identifier**. The thematic identifier or code is used to identify a real-world phenomenon. Therefore, currently the key linking value at record level is the INSPIRE identifier (e.g. *inspireId* in the INSPIRE PS data model), in principle not associated to any unique identifier of the designated area in the reporting data flow. If a thematic identifier was part of the INSPIRE Protected Sites core data model, it could be used to contain the specific thematic site code directly (e.g. *cddald* in the current CDDA model), thus avoiding management issues related to the encoding of the INSPIRE identifiers (*inspireId*) in the thematic part of Type 2 data.
- **On the change requests to improve INSPIRE alignment and fit for purpose:**
 - Partly inspired by this exercise, a proposal has been submitted to amend the INSPIRE PS data model in order to add a **thematic identifier** element (with the multiplicity 0..*). Such data type ThematicIdentifier is already included in the INSPIRE Generic Conceptual Model²².
 - Also partly due to this exercise, although not directly related to the “Linked Approach” discussion, **additional change requests have been submitted to the INSPIRE community** as a few required elements of the INSPIRE PS data model do not fully align with the CDDA reporting requirement, in particular:
 - the **format** of the *legalFoundationDate* element (year of designation), deemed unnecessarily detailed (i.e. up to the second) for its purpose, being a regular practice by the regulatory authorities to encode only the year, and
 - the **multiplicity** of the elements *legalFoundationDate* and *legalFoundationDocument*, which does not support providing the dates and founding documents corresponding to several designations of the same spatial object. For example, the CDDA and Natura 2000 assign legal references not to the protected site but to the designation itself. The proposal in this case is to couple the legal foundation date and legal foundation document with the designation, and to allow multiple values of those properties.
- **On the coordination between the responsible authorities**
 - The **coordination of INSPIRE implementers** plays a key role for the successful implementation of the “Linked Approach”. Coordination between institutions will avoid further duplication of efforts or the increase in the burden on reporters. Increased cooperation is needed to agree on recommendations on how to build

²² <https://inspire.ec.europa.eu/documents/inspire-generic-conceptual-model>

persistent identifiers for the INSPIRE objects or to give authoritative instructions on whether to label “official” data to re-use in reporting contexts (i.e. using keywords such as those from the list of priority datasets²³);

²³ <https://webgate.ec.europa.eu/fpfis/wikis/display/InspireMIG/Action+2016.5%3A+Priority+list+of+datasets+for+e-Reporting>

3. INSPIRE extensions approach

This section is focused on the specific application of the INSPIRE extension methodology of an INSPIRE data model (Production and Industrial Facilities, PF) to cater for the requirements of a thematic reporting data flow, the EU Registry on Industrial Sites. It introduces briefly the concept of a generic data model extension and continues then with the description of how the methodology was implemented in the case of EU Registry on Industrial Sites data flow. As in the previous case, this section concludes with a summary of the lessons learned from this implementation, including challenges and benefits of the solution adopted.

3.1. Methodology

As its focus is mostly on spatial data harmonisation and publication, INSPIRE “does not aim at modelling all information to execute business processes, scientific simulations, or comply with reporting requirements”²⁴. To reuse INSPIRE data models in specific environmental legislation or thematic domain frameworks there will generally be a need to provide additional thematic information, either linked to the INSPIRE core models, as we have seen in section 2, or included in the (extended) INSPIRE data model, which is the subject of this chapter.

The INSPIRE Generic Conceptual Model, specifying the common modelling rules for the development of all INSPIRE data specifications, sets out as well general rules for the design of conformant INSPIRE extensions and provides an implementation example in order to facilitate harmonised extension processes efforts across sectors and across Europe. The general rules for extending an INSPIRE data specification in conformity to INSPIRE can be found in the Annex F of that document:

“Extending an INSPIRE data specification would imply at a minimum that:

- the extension does not change anything in the data specification but normatively references it with all its requirements*
- the extension does not add requirements breaking any data specification requirement. “*

The above rules are minimum requirements, a baseline for the development of valid INSPIRE extensions. In order to achieve further harmonisation of extended models, the Annex F also recommends that the extensions are conformant to all modelling rules of the Generic Conceptual Model and to the requirements of the Guidelines for the encoding of spatial data²⁵.

Based on above Generic Conceptual Model requirements, a generic data model extension may:

- add new feature types and data types in the extended application schema.

²⁴ [INSPIRE Generic Conceptual Model - Sec. 6.1](#)

²⁵ https://inspire.ec.europa.eu/documents/Data_Specifications/D2.7_v3.3rc3.pdf

- and new constraints to the elements in the original schema, provided that these do not break any existing requirement (meaning that they can only be more demanding e.g. INSPIRE optional elements can become mandatory in the extended model but vice versa is not allowed).
- add new attributes to existing feature types and data types.
- add new code lists or extend (extensible) INSPIRE code lists, with obligation to make new values and their definition publicly available in a register. Guidelines and best practices for setting up registers for extended INSPIRE code lists can be found in the INSPIRE Best Practices for registers and registries & Technical Guidelines for the INSPIRE register federation²⁶.

This section will not enter into details on how to proceed technically with all these different steps, which should be the subject of a dedicated work. There are already available additional resources to facilitate the design and implementation of INSPIRE data model extensions, such as the WeTransform INSPIRE Data Specification Extensions²⁷, including a step-by-step guidance, a tutorial and some best practices and implementation examples. This section will rather focus on a specific implementation of the generic INSPIRE data model extension methodology in the context of the EU Registry on Industrial Sites.

3.2. The use case of the EU Registry on Industrial Sites

In the wider context of streamlining the reporting of industrial emissions, the EU Registry on industrial sites is intended as a coherent database of industrial entities that release emissions centralising existing mechanisms for the collection of administrative facility data. The EU Registry collects identification and administrative data for the:

- European Pollutant Release and Transfer Register (E-PRTR) Regulation facilities²⁸,
- Installations under the scope of the Industrial Emissions Directive (IED)²⁹,
- Large combustion plants (LCP, chapter III of the IED)
- Waste incineration and co-incineration plants (WI, chapter IV of the IED).

The EU Registry will be the reference dataset to which relevant thematic reporting on LCPs under the IED and facilities under the E-PRTR Regulation will link to (i.e. data on releases and transfers referring to the entities reported to the EU Registry). The planned data flows for reporting and handling administrative and thematic data on industrial emissions are presented in a simplified manner in figure 7. The most relevant aspect here is that the EU Registry contains all relevant permit and geospatial information of these industrial entities, avoiding its duplication in the thematic data flows. The geographic location information of industrial sites held in the EU Registry is the foundation behind its compliance with the INSPIRE Directive. This shall reduce reporting burden and

²⁶ <https://inspire.ec.europa.eu/id/document/tg/registers-and-register-federation>

²⁷ <http://inspire-extensions.wetransform.to/index.html>

²⁸ <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:033:0001:0017:EN:PDF>

²⁹ <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32010L0075>

data management costs, ensuring further investments by reporting countries are not required in the medium term to comply with INSPIRE.

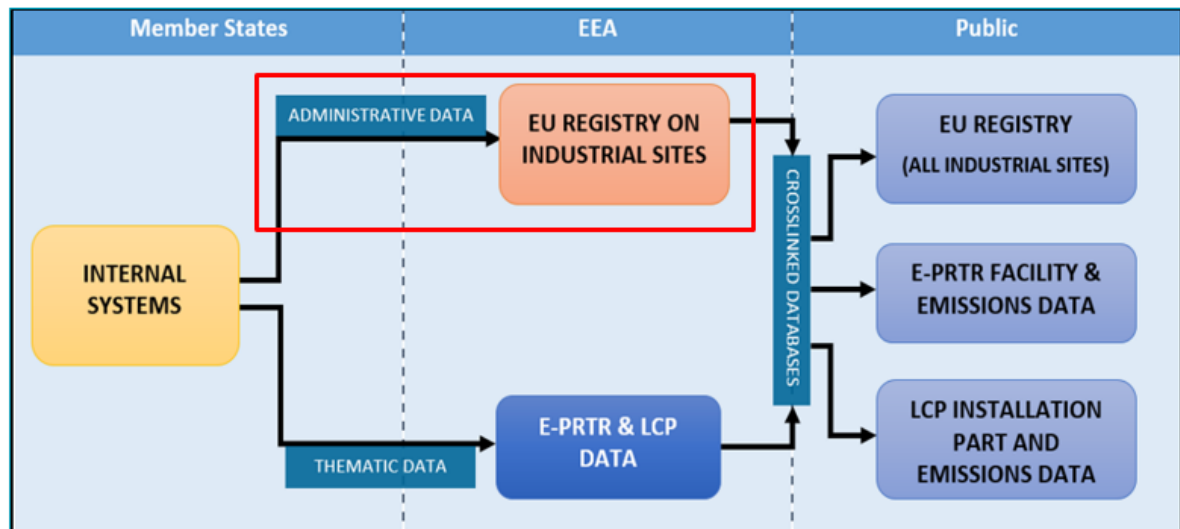


Figure 7: Planned data flows for reporting and handling administrative and thematic data on industrial emissions

While avoiding inconsistencies across the sectorial legislation, the EU Registry data model³⁰ extends the INSPIRE Production and Industrial Facilities (INSPIRE PF) core model³¹ in accordance to the rules set out in the INSPIRE Generic Conceptual Model introduced in 3.1. This means that the EU Registry data model does not change anything in the INSPIRE PF data specification and the new constraints do not break any existing data specification requirement.

During the modelling exercise, it became however clear that the INSPIRE PF data model was very comprehensive and catered for a very high level of detail, with a number of fields and some feature types which were identified as not relevant in the EU Registry context. Therefore, a streamlined view was developed, including only feature types and fields relevant to the EU registry reporting. This streamlined view incorporates all mandatory and non-voidable elements, together with domain reporting requirements and a series of modifications which have no impact on the INSPIRE compliance of the reported data by countries. The rationale for this choice is based on the assumption that Member States do not collect or make available further data on this subject. The modified schemas of the INSPIRE PF model and the INSPIRE Activity Complex (which is imported by the former) are stored locally in the Reportnet Data Dictionary³², facilitating the workflow and the schema validation process. This technical choice is also meant to avoid that possible updates to the INSPIRE schema are directly reflected (with no possibility for the extension developers to check undesired inconsistencies with sectorial legislation) in the EU Registry extended data model.

³⁰

https://cdrtest.eionet.europa.eu/help/ied_registry/documents/Guidance/EU%20Registry_datamodel_CID_Final%207%20August%202018.pdf

³¹ <http://inspire.ec.europa.eu/schemas/pf/4.0/ProductionAndIndustrialFacilities.xsd>

³² <https://dd.eionet.europa.eu/schemaset/euregistryonindustrialsites/view>

The modelling process followed the steps described below:

- Non-required feature types were removed (voidable types such as Production Building and Production Plot) as they provided more specific details than the current reporting requirements.
- Non-required attributes were also removed (optional or voidable)
- Extensions to each remaining feature type containing additional information required by the industrial legislation were then added:
 - New code lists were added and existing INSPIRE code lists were extended, with relevant definitions available in the EU Registry on Industrial Sites vocabulary of the Reportnet Data Dictionary³³.
 - New attributes were added to existing feature types and data types.
 - New data types and constraints were added, following the principle that new constraints applied shall be always more demanding than the ones imposed by the core model. More specifically, the EU Registry data model introduces the following constraints on elements of INSPIRE PF data model:
 - the location of *Installation* and *Installation Parts* must be provided ('pointGeometry' element from optional to mandatory)
 - the value describing the status or condition of the *Facility/ Installation/ Installation Part* must be provided (only one 'status' and void values not allowed)
 - the geometry of the *Facility* is forced to be a point (whilst INSPIRE PF allows any kind of geometry type).

It is important to stress that all implemented changes do not break INSPIRE rules, since they operate on INSPIRE-allowed values (they are restrictive with respect to the relevant domain).

³³ <http://dd.eionet.europa.eu/vocabularies?expand=true&expanded=&folderId=27#folder-27>

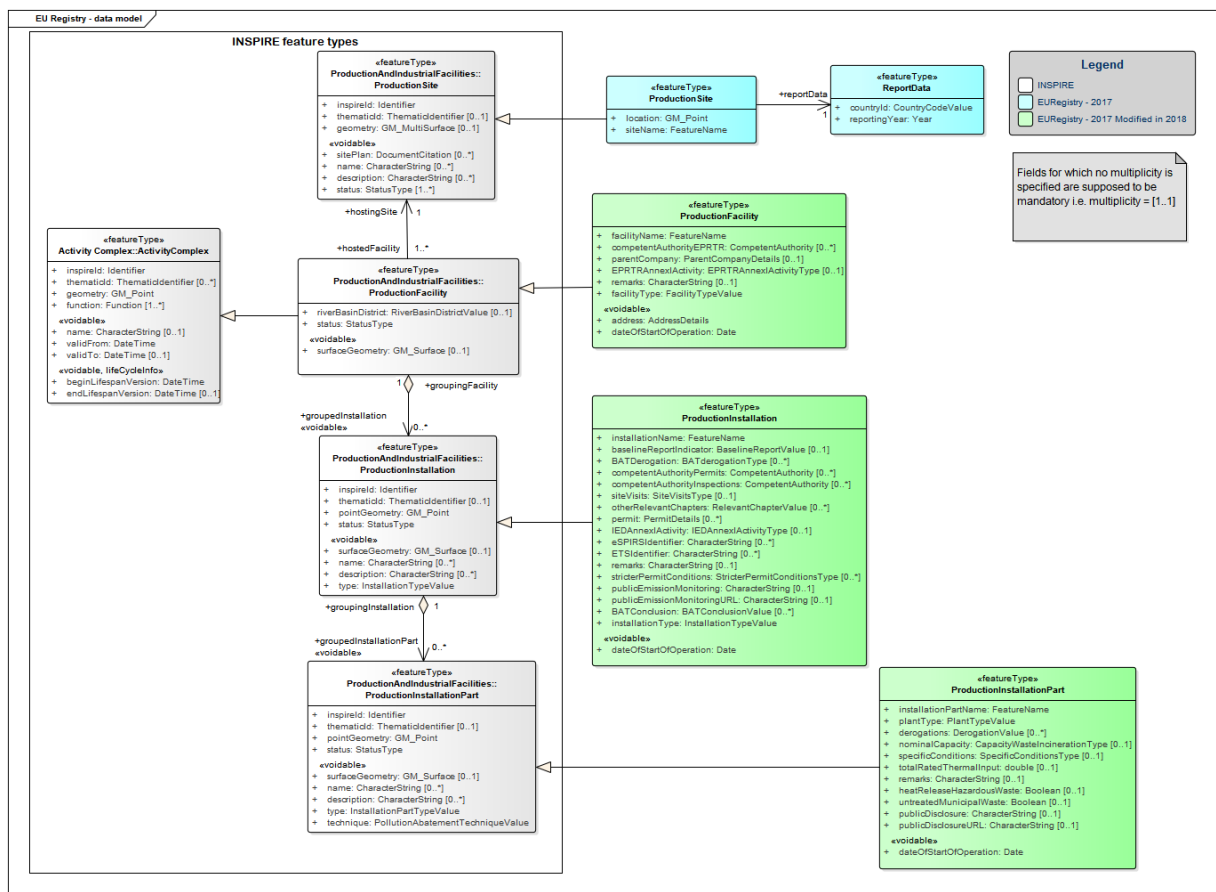


Figure 8: EU Registry data model

The resulting model (see figure 8) is basically structured according to a hierarchy of the feature types “Production Sites”, “Production Facility”, “Production Installation” and “Production Installation Part”, which is defined according to geographic and ownership attributes. While fully compliant with INSPIRE, this is a model easily understood by national thematic staff and industrial reporters.

The thematic data (e.g. emissions) which collection is required by IED and the E-PRTR regulation have been also subject to a streamlining exercise and an integrated data model³⁴ has been designed in order to operate in parallel with the EU Registry data flow. The thematic data model under both legislative instruments only concerns the transmission of emissions and associated activity data for industrial emissions entities that are included in the EU Registry. The link between these two reporting data flows follows a similar conceptual approach as the “Linked Approach” described in Section 2 (although applied to an INSPIRE extended schema) and is based on a consistent use of identifiers between related feature types. In particular, the INSPIRE identifiers (inspireId) of the relevant Production Installation Parts or Production Facilities will be the key attribute that ties these entities (and therefore their geospatial information) to the corresponding emission data, as described in the diagram below (see figure 9). The implementation of this relation eliminates the need to provide geospatial information or identification / administrative data in each thematic data flow.

³⁴ https://cdrtest.eionet.europa.eu/help/eptr_lcp/help/eptr_lcp/Guidance/EPTR-LCP_datamodel_v3.0.pdf

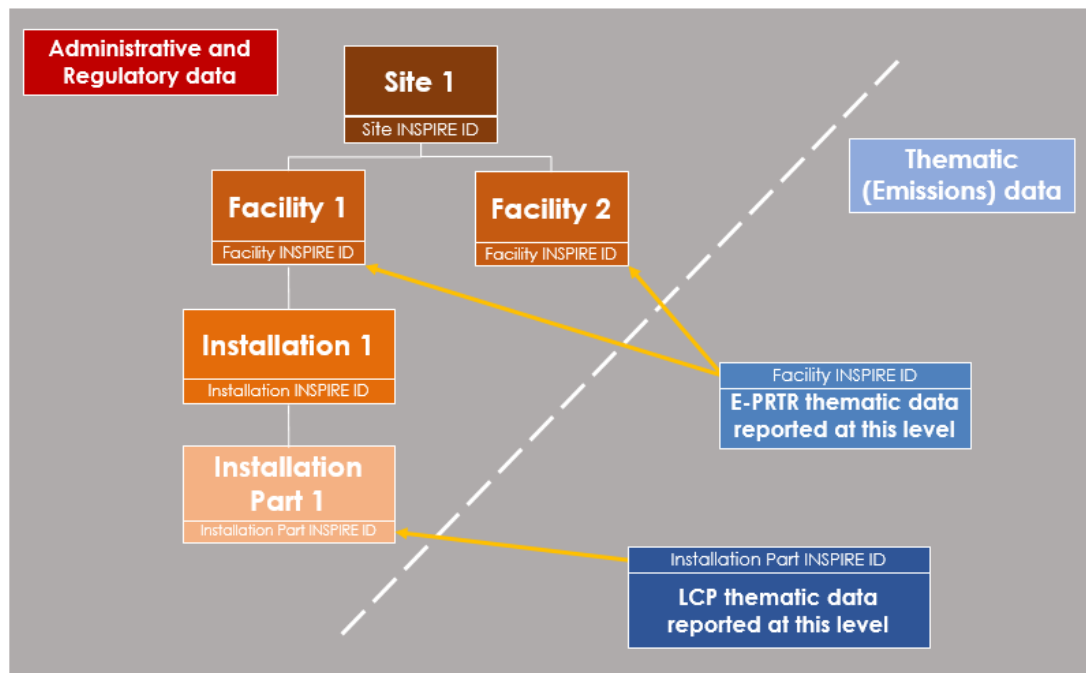


Figure 9: EU Registry entities hierarchy and link to emissions data

QA/QC procedures ensure that, when combined, data submitted via both data flows remain logic and coherent to the requirements of the industrial legislation. The EEA will then harvest all country data submission and aggregate them in a European-wide multi-year data set of industrial sites that can be used for analysis, publication and support of environmental programmes.

3.3. Lessons learned

The INSPIRE core data models are designed to be extensible so as to benefit from the overall INSPIRE infrastructure while ensuring consistency with national as well as domain specific legislation. The EU Registry is arguably a good example of the use of an extended INSPIRE data model to effectively streamline the reporting under a specific thematic legislation.

As in the “Linked Approach” case, any reflection on the lessons learned from this specific implementation should also take into consideration that the current reporting infrastructure is not yet ready to build on a web service based approach so Member States will need to upload their INSPIRE compliant EU registry files to Reportnet, at least for the time being.

The most important lessons learned from this activity are summarised below:

- **On the common register for several reporting data flows:**
 - As part of the streamlining efforts in the reporting of industrial emissions, the EU Registry will be the reference (geospatial) dataset to which thematic reporting under several thematic legislative instrument will link to – through the use of unique

identifiers (InspireId). The main benefit from this approach is that from now on, geospatial information will not be duplicated in each thematic data flow, with the corresponding optimisation in reporting efforts by the Member States.

- **On the extended data model adapted to specific needs:**

- As a basic modelling principle, extending a data model implies that the new (extended) model typically inherits most of the elements from the core model. Therefore an extension approach is particularly convenient when the original data model is quite similar to the new one, i.e. it shares most of the original data information demands. This was the case of the EU Registry, where it was justified the creation of an extended model as the INSPIRE PF model includes data structures for E-PRTR facilities, IED installations, LCP and WI, which emissions and transfers data are to be reported under IED and E-PRTR Regulation. By extending the INSPIRE PF data model, the EU Registry inherits all basic data structures needed to satisfy different reporting obligations (including the one regarding INSPIRE data interoperability) through a single model.
- Although it is not currently envisaged, the system is now flexible enough to accommodate future extensions, enabling the EU registry to become the reference dataset for more (related) reporting obligations.

- **On the encoding:**

- It can also be argued that the major challenge in this activity was to deal with the complexity of extending the INSPIRE data model, as well as with the (more general issue of) management of GML encoded data. Some of the current software solutions are unable to take advantage of the full potential of GML encoding, especially in the case of nested data structures (complex feature types) of the INSPIRE data models and the big size files. However, it shall be pointed out that in the EU Registry case, the streamlined view of the model only contains simple data types, or complex data types which in turn only contain simple data type properties. Geometries are only point-based, which makes unlikely that the size of the GML files becomes unmanageable.

- **On facilitating INSPIRE implementation:**

- Since the deadline for the provision of fully harmonised INSPIRE PF datasets is October 2020, many countries have not yet implemented all relevant obligations with respect to INSPIRE data interoperability at the time of the first EU Registry reporting (mid 2019). Therefore, the EU Registry data model will provide the building blocks for the provision of INSPIRE compliant data, streamlining the efforts of Member States to meet INSPIRE requirements. However, as indicated in the case of CDDA, this does not substitute all the obligations coming from INSPIRE, meaning that the INSPIRE PF datasets still need to be described by metadata, be accessible via view and download service and available through the INSPIRE Geoportal.

4. Blueprint for the re-use of INSPIRE data models in environmental reporting data flows

The CDDA 2018 and EU Registry reporting projects are good examples on how INSPIRE can effectively contribute to streamline environmental reporting data flows. While these two projects can be considered as templates for future streamlining activities, individual reporting specificities may lead to significantly different technical solutions, to be adopted on a case-by-case basis - e.g. a combination of INSPIRE extension and “Linked Approach” methodologies.

Building on lessons learned (described in 2.3 and 3.3), this section elaborates on the main questions which may drive the selection of best-fitting methodological solution for a reporting workflow based on INSPIRE datasets and provides, where possible, recommendations for mitigating potential issues that may arise. This section will not however elaborate on issues associated to a reporting following an INSPIRE web-service based approach, which is to be covered as part of the Reportnet 3.0 project as indicated above as well as in section 5.

4.1. Technical and organisational cooperation

A first key aspect considered as the baseline for a blueprint of the re-use of INSPIRE in environmental reporting data flows can be quoted from section 6 of the INSPIRE Generic Conceptual Model:

“It is important that INSPIRE and policy working groups for Reporting closely liaise and cooperate during the INSPIRE maintenance & implementation process in order to further streamline reporting guidelines and harmonise them with INSPIRE data models and infrastructure.”

Independently from the technical or methodological approach adopted, a successful reuse of INSPIRE in reporting contexts will have to build on a strong cooperation of all involved actors, both from the thematic and INSPIRE fields. Both use cases described in this document point at the need for a close interaction to make the streamlining exercise effectively succeed:

- The **data model development phase**, and especially when proceeding to an extension of the INSPIRE core data models, can only be successful if cooperation exists between environmental reporters and INSPIRE experts. This will ensure that the final reporting model fits the demands of both communities. Likewise, at organisational level, major benefits of the extension approach - i.e. the possibility of sharing workflows and resources - require proper management and information-sharing procedures for the departments in charge of the different related data flows.
- **Identification of authoritative INSPIRE data relevant to the specific reporting purposes** is another critical aspect that requires a solid cooperation between institutions, which could

benefit from some specific (meta)data labelling (e.g. using approved keywords in the INSPIRE dataset and services metadata such as the keywords of INSPIRE Priority Datasets³⁵).

- Especially in the case of the “Linked Approach”, such as the one used in CDDA, but also in many reporting data flows that can build on INSPIRE, the spatial object identifiers (e.g. inspireId, thematic identifier) play a key role in the retrieval of INSPIRE data relevant to reporting. **It is absolutely necessary that the values of the INSPIRE identifiers are known to the environmental reporters.** Clear rules for the creation and management of such identifiers, agreed at national or, even better, at European level, as well as information-sharing procedures for the departments in charge of the different related data flows would avoid potential reporting issues and/or inconsistencies.

4.2.Key questions to help guide future INSPIRE-based reporting modelling

A thorough comparative analysis of the requirements stemming from the particular reporting obligation and relevant INSPIRE data specifications is clearly the first step for any INSPIRE based e-reporting development. In order to help guide technical developments in future reporting workflows reusing INSPIRE, this analysis should answer, among others, the following questions (see also summary in figure 10 below):



Figure 10: Some key questions to address when planning to reuse INSPIRE in environmental reporting

³⁵ <http://inspire.ec.europa.eu/metadata-codelist/PriorityDataset>

- **How similar are the requirements stemming from the thematic reporting obligation and relevant INSPIRE data specifications?**

The harmonisation principle is already embedded in the environmental reporting obligations requiring the provision of binding and agreed content that allows a certain level of automatisisation to collect the Europe-wide environmental data and information.

Very similar information requirements would typically point towards an INSPIRE extension approach such as the one used in the EU Registry, entailing a relatively small effort to make the INSPIRE data model fitting specific reporting requirements. On the other hand, common requirements limited to the geo-spatial reference (that could cover several INSPIRE themes) and a few other data elements would point at the “Linked Approach” as a more suitable option allowing the re-use of the required INSPIRE data in various and specific thematic data flows. With this, it would be possible to avoid several modelling exercises which would need to cater for data elements not relevant to each specific reporting contexts.

The selection of the best fitting methodology should take into consideration that the general rules for a conformant INSPIRE extension prescribe the compliance with all requirements in a data specification to ensure interoperability. This could however be at odds with specific reporting requirements, which may not foresee the collection of additional data outside the specific obligations rendering the retrieval of these non-reporting data elements difficult and not cost-efficient. Inconsistencies between reporting obligations and INSPIRE requirements should be addressed on a case-by-case basis approach that goes beyond the scope of this document and is dealt with in the context of several actions under the Action Plan to Streamline Environmental Reporting³⁶.

- **Are other reporting obligations referencing the same INSPIRE spatial data? And how similar are these reporting obligations?**

When there are several reporting obligations, very different in nature but referencing the same INSPIRE core data, the use of the “Linked Approach” could be the preferred option, since it allows the “as-is” re-use of the same INSPIRE core data in different reporting contexts. The business data related to a specific reporting would be modelled separately, avoiding the development of several different extensions of the same INSPIRE data model, or of a very complex single one.

An exception to this rule is actually the EU Registry example, where a limited extension of the INSPIRE core data model allows catering for several reporting obligations, which have already many elements in common. In this case, an extension approach has been the most cost-efficient solution,

³⁶ http://ec.europa.eu/environment/legal/reporting/pdf/action_plan_env_issues.pdf

but it should be noted that a streamlining process of the industrial emissions reporting had previously taken place.

- **How many INSPIRE data specifications are relevant to the reporting data flow?**

In case a single reporting obligation references data covered by different INSPIRE themes, the use of the “Linked Approach” is definitely more advantageous, as it allows retrieving the required information from different INSPIRE datasets without creating very complex schemas when extending and integrating different INSPIRE data models.

- **How complex could an extended INSPIRE data model become?**

Extending an INSPIRE data model may become a challenge when relevant reporting information and/or mandatory INSPIRE elements are modelled through very detailed data structures making use of complex data types. Apart from the difficulties related to the extension at the conceptual level, the presence of complex structures with several nesting levels - i.e. data structures embedding other complex data structures – might result in big reporting data sets that typically used tools may find it difficult to retrieve and properly manage.

While alternative encodings for basic INSPIRE data exchange and direct visualisation in standard GI tools are being investigated by the INSPIRE Maintenance and Implementation Group (MIG) experts³⁷, the use of a “Linked Approach” to provide thematic data currently not modelled under INSPIRE core data models may help sort out some of these issues. Thematic data can be simply encoded as flat XML structures, easy to manage, update and exchange, while maintaining the link to the geospatial data modelled according to INSPIRE. But the considerations above should also be taken into account.

Are there elements that can act as a link between INSPIRE and thematic data, e.g. a thematic identifier?

As already indicated, a key aspect in the implementation of a “Linked Approach” concerns the identification of the actual relevant INSPIRE data and, especially, the availability of key values to link the thematic data with the corresponding spatial objects in the identified INSPIRE dataset. These key values may typically, but not always, be thematic identifiers, i.e. unique spatial object identifiers established to meet data exchange requirements of different reporting obligations at international, European or national levels. Thematic identifiers are not necessarily always available in the INSPIRE data models (e.g. INSPIRE PS data theme), in which case a strong recommendation is to embed the thematic identifiers in the INSPIRE identifiers (InspireId). Agreed keywords such as the priority dataset code values can also help identify the relevant reporting data. In a nutshell, the existence of common thematic identifiers in both reporting and INSPIRE data models will facilitate the implementation of a “Linked Approach” technique.

³⁷ <https://ies-svn.jrc.ec.europa.eu/projects/2017-2/wiki>

- **How stable are the thematic reporting requirements?**

A very dynamic, evolving thematic reporting may benefit from a simpler modelling exercise and flat structures rather than a fully-fledged INSPIRE extension process, which may take some time to ensure the final model respects all INSPIRE requirements. This may be therefore a criterion to select a “Linked Approach” when re-using INSPIRE in e-reporting. Stable reporting data models may instead be more suited to build upon extensions of INSPIRE core models, especially if there are many commonalities with the reporting data flows.

On an individual basis there may be many other questions or criteria that may drive the implementation of an INSPIRE-based reporting data model. The questions above shall be only considered as an orientation and not as a fixed recipe.

5. Conclusions

By using real life examples, the previous sections of this document have elaborated on the different technical approaches to embed INSPIRE in the context of environmental reporting, facilitating the streamlining of reporting data flows by re-using national INSPIRE geospatial datasets. Lessons learned from the 2018 CDDA reporting and the EU Registry on Industrial Sites, operational workflows at the time of writing this document, have been the basis for developing a first blueprint that can help guide the modelling of future environmental reporting data models based on INSPIRE. It is important to highlight that one of the key lessons learned out of these projects is actually of a non technical, organisational nature: a strong cooperation between thematic reporters and INSPIRE implementers is vital in order to set up an efficient reporting workflow based on INSPIRE. As it stands now, the similarities between thematic reporting data flows reusing the same INSPIRE datasets or the existence of well established thematic identifiers in the INSPIRE core data models have a clear influence on the choices made during the modelling process.

As mentioned on several occasions in this document, the reuse of INSPIRE in e-reporting should result in less administrative burden at Member State level, since multiple provision of the same data will be avoided. Coherence and consistency should also increase, as the same datasets are shared across different reporting flows. Furthermore, by embedding INSPIRE in these reporting obligations, Member States would be supported to fulfil the INSPIRE obligations while starting to reap the benefits from a decade of investments in INSPIRE implementation. However, it needs to be stressed that providing harmonised datasets in the context of a specific reporting obligation, as covered in this report, does not substitute all the obligations coming from INSPIRE. All INSPIRE datasets used in the context of a reporting obligation and appropriately harmonised need to be described by metadata, be accessible via view and download service and available through the INSPIRE Geoportal.

The first section of this paper already points at the fact that the examples described have not however yet embraced the full potential of INSPIRE for e-Reporting. The streamlined reporting workflows have not yet moved to a web-service based approach. There are several reasons – technical and organisational - behind this choice, being among them the lack of homogeneous implementation of INSPIRE across Europe or the need to address technical and legal challenges before moving to an operational reporting process based on web service harvesting. Nevertheless, the steady progress in the implementation of INSPIRE, activities such as the list of priority datasets for environment reporting and the Reportnet 3.0 project currently under preparation, shall facilitate a transition in the long term to a future reporting system which can harvest INSPIRE web services to collect geospatial datasets relevant for thematic reporting obligations.

Page left intentionally blank