

Specific contract No 07.0307/2012/638595/D2 implementing Framework
Service Contract ENV.D.2/FRA/2012/0013

URBAN WASTE WATER TREATMENT DIRECTIVE

TOWARD A NEW INFORMATION SYSTEM

SIIF

(Structured Implementation and Information Framework) IT ASPECTS

Final Version
30 January 2014



Date: 30 January 2014
Authors: Katharina Lenz, Edith Hödl-Kreuzbauer (both UBA)
Benoît Fribourg-Blanc, Sylvain Grellet (both OIEau)
Contract Manager: Edith Hödl-Kreuzbauer (UBA)

The contents of this document are subject to copyright and all rights are reserved. No part of this document may be reproduced, stored in a retrieval system or transmitted, in any form or by any means electronic, mechanical, photocopying, recording or otherwise, without the prior written consent of the copyright owner.

This document has been produced by Umweltbundesamt GmbH and OIEau.





Content

I.	INTRODUCTION	1
II.	THE BACKBONE OF THE INFORMATION SYSTEM	2
II.1.	A DISTRIBUTED INFORMATION SYSTEM	2
II.2.	COMMUNICATION CONTENT: A COMMON DATA MODEL	3
II.3.	COMMON COMMUNICATION FORMAT/FLOWS	7
II.4.	SIIF AS A SUPPORT TO E-REPORTING	8
II.5.	SIIF AS A SUPPORT TO REPORTING VIA REPORTNET	10
II.6.	DATA MANAGEMENT RULES	10
II.7.	DOCUMENTING RESOURCES OF THE SYSTEM	12
II.8.	PERSISTENCY WITHIN THE SYSTEM	13
II.9.	QUALITY OF SERVICE	13
II.10.	SPECIFIC SIIF EU-LEVEL NODE IT CONTENT	14
II.11.	SIIF AND MS INSPIRE COMPLIANCY ON UWWT	15
II.12.	ENSURING BACKWARD COMPATIBILITY	16
II.13.	FORWARD LOOKING PERSPECTIVE.....	16
III.	NODES FRONT END	17
III.1.	WEBSITE ACCESSIBILITY	17
III.2.	CONTENT VISUALISATION	18
III.3.	ACCESS TO INFORMATION	18
III.4.	ACCESS TO DOCUMENTATION.....	19
III.5.	HELPDESK.....	19
III.6.	DATA FLOWS REPRESENTATION.....	20
III.7.	LINKS BETWEEN EU AND MS NODES	20
III.8.	UWWTD LEGISLATION OVERVIEW.....	20
III.9.	SPECIFIC SIIF EU-LEVEL NODE INTERFACE	20
IV.	SIIF DEPLOYMENT	21
IV.1.	CORE BRICKS	21
IV.2.	PROMOTING AN IT WATER COMMUNITY.....	22
IV.3.	CENTRALLY MAINTAINED TOOLSET VS FULLY DISTRIBUTED	23
IV.4.	A TRANSITION PERIOD	24
V.	UWWT SIIF AND OTHER SIIFS COHERENCE	24
V.1.	CROSS-CUTTING INFORMATION SYSTEMS ISSUES EXAMPLE	24
V.2.	CROSS-CUTTING IT GOVERNANCE	25
V.2.1.	<i>The approach</i>	25
V.2.2.	<i>Proposal for new roles repartition</i>	27
V.2.3.	<i>Who will benefit from this approach?</i>	28
V.3.	UWWTD SIIF CONNEXIONS WITH WISE AND SEIS	28







I. Introduction

The core of the concept paper document presents an overall vision of the future UWWTD SIIF. Being a concept paper document, technical discussions are kept to the minimum, in order to ensure the message on the vision is as clear as possible. In the meantime, setting up an information system with both its machine to machine and human interfaces also implies entering into technical specificities. These are described in more details in this document.

The various IT items presented in the next chapters are based on IT discussions held between the consultant, DG-ENV and the EEA. A large part of these IT discussions are relevant for the entire SIIF system, whatever the thematic content and this is why a separate document was developed. However they mainly aim at addressing the specific case of the UWWTD SIIF EU node, understood as being the European component of the UWWTD SIIF. Thematic content discussion such as addition or deletion of parameters or reorganisation of them are mainly addressed in the “Background document Thematic Aspects” of the concept paper. It is solely addressed here when it may have implication on the IT system.

The current background document aims at identifying the key elements of the system and the questions and aspects that need to be addressed to fulfil the SIIF requirements. It does not aim at providing already pre-packed full solution and it is anticipated that the roadmap and workplan to be developed will address a possible way forward. In particular, whether the EU node is distributed in countries or centralised has not been decided.

The technical specificities addressed in the following constitute a basis for discussion and refinement toward a progressive and pragmatic implementation of the system.

This document is organised as follows:

- Chapter II focuses on the invisible part of the system to ensure proper machine-to-machine dialogue. Thus defines the Backbone of the System,
- Chapter III describes how UWWT information should be exposed to the public. It defines the functionalities SIIF Nodes interface should provide,
- Chapter III starts the describing how the entire system could be deployed exploring different paths,
- Eventually, chapter IV provides an input on cross-cutting information system issues. Aspect that should not be overlooked as having an overview on those often helps better delineate the system that is being created.
- Annexe I is focusing on the main use cases the future UWWTD SIIF EU node should aim at covering with the main steps this entails. It comprises an identification of the different targeted end users of the UWWTD SIIF, a much broader audience than simply the non expert citizen.





- Annexe II is a tentative draft INSPIRE compliant model based on UWWTD data model and focussed on the components of the UWWTD reporting that have spatial representation: UWWTPs, Discharge point, agglomerations and receiving areas.

II. The backbone of the Information System

II.1. A distributed information system

As mentioned in SEIS communication, “information should be managed as close as possible to its source” (the 7 SEIS principles are available in annex III of the concept paper) which refer to the concept of IT networks based on individual and interconnected data sources. Data dissemination point should be maintained close to the producer to ensure that systems run on the most up-to-date set of information respectively the management of the data and information can be done by the experts more familiar with the content. Furthermore decentralised information systems depending on local data nodes became in the recent years the common architecture for information networks.

In the context of UWWT Directive data information is currently collected and reported on a local level. Further information on this is provided by the sections above dealing with UWWT reporting towards the EU. The distributed systems and the hierarchical structures (data points connected as data node) are implemented at national level which makes it reasonable to stay with the already established network structure approach.

The future UWWT SIIF will then be a distributed system made up of a network of SIIF nodes on different levels (from local to national to European level). Its infrastructure will be based on partners' infrastructure providing access to their respective information nodes.

This approach is similar to other data infrastructure initiatives by the EU which defines only the specifications and requests. Services are seen as resources of the information system. And nodes are then loosely coupled via those services. The only software component constraint on each node is then to be able to communicate using services.

Each MS will deploy its national node and set up the communication with the central node located at the EC level according to the webservices described below. Technical interoperability between each node will be ensured by common open protocols. On that aspect the Internet and its protocols are now considered to be a basis to build the information system on.

The future EU UWWTD SIIF has to respect subsidiarity and be non intrusive. To this extent only the interface and the exchange between partners systems matters (MS et EU-level nodes of the UWWTD SIIF). The SIIF does not aim at defining the way the information should be collected and stored nor how this information should be displayed at MS level. It will however be necessary to





establish clear links between the EU node of the UWWTD SIIF and the national node to allow each MS to display more information than reported in the official UWWTD reporting on each of the geographical objects covered by the reporting: agglomerations and UWWTP as a first priority but also discharge points, sensitive areas. To allow this, and keep it robust, MS will have to implement a single link for each individual object and using a simple and standardised method for all objects, with website root and the EU identification code as a basis.

However, MS may also envisage apply the proposed approach to UWWTD SIIF nodes available at sub-MS level. This could in principle be applied at any of the aggregation units of lower level: region, county, Länder, RBD, sub-unit.

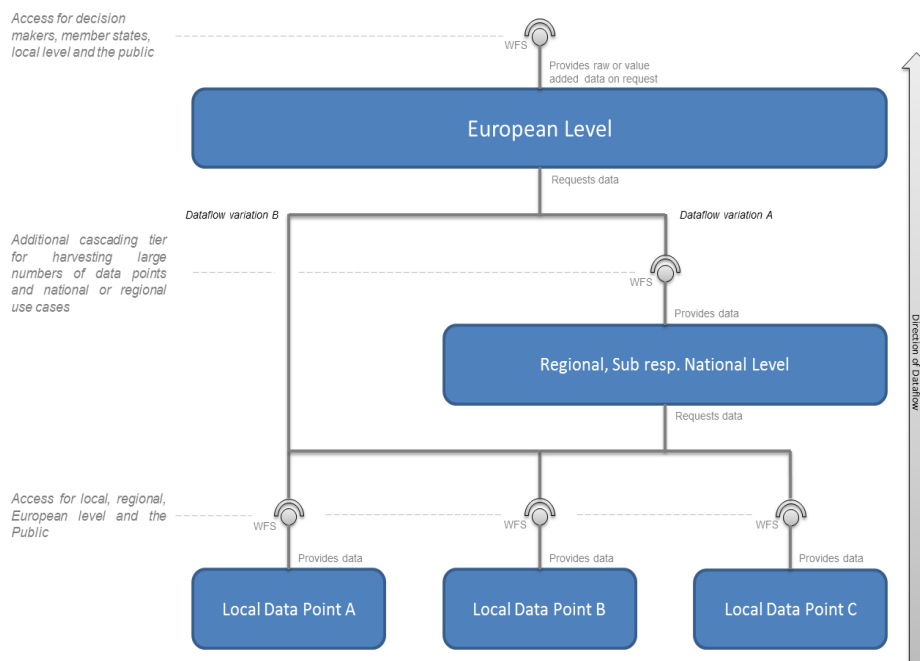


Figure 1 - Distributed architecture and information exchange between nodes

Most important aspect of the SIIF IT issues is then de facto the clarification of the types and formats of information exchanged. These aspects are addressed in the next chapter.

II.2. Communication content: a common data model

The content of the information exchanged is to be based on the information needed for the reporting to the European Commission. In the recent years, many achievements have been done on the technical interoperability. Open standards and their software implementations have improved, which may facilitate data exchange on an almost instant basis compared with today.



Inspire is now providing an excellent incentive for communities to organise themselves around common data models and semantics. The EU Water community has always been active with its water Directives reportings and the setup of WISE.

Common models and reporting sheets already exist but the formats actually used have the following limitations:

- Data models/dictionaries are not always available as plain UML models,
- Semantic is always defined within the perimeter of a single Directive. Thus the same words/expression are used to define different notions in different Directives. Clear examples encountered on the terms Waterbody and Agglomeration (2 terms used in UWWTD reportings), have been identified. For example, for 'Agglomeration' the following meanings have been identified:
 - Environmental Noise Directive (END): 'agglomeration' shall mean part of a territory, delimited by the Member State, having a population in excess of 100 000 persons and a population density such that the Member State considers it to be an urbanised area. Strategic noise mapping: agglomerations with more than 250 000 inhabitants (available in 2007).
 - Air Quality Directive : 'agglomeration' shall mean a zone that is a conurbation with a population in excess of 250 000 inhabitants or, where the population is 250 000 inhabitants or less, with a given population density per km² to be established by the Member States;
 - Urban Waste Water Treatment Directive: 'agglomeration' means an area where the population and/or economic activities are sufficiently concentrated for urban waste water to be collected and conducted to an urban waste water treatment plant or to a final discharge point;
- The modelling technique is not implementing open standards defined in ISO 19100 series and OGC,
- Data models/dictionaries are not using Inspire defined concepts and application schemas.

All the above mentioned points result in:

- Improper understanding of the information asked for in reportings,
- Complexifying the deployment of OGC WFS services which need a proper ISO19100 data model,
- Difficulties when trying to interface with another community.

Which in turn undermines the quality of the information reported and, of course, its potential reuse within the water community or by another community.



European Interoperability Framework¹ recommendation n°2 clearly identifies the use of Open Standards as a principle to be considered for any eGovernment services to be set up at a European level.

Open Standards are indeed crucial for the information sustainability compared to vendor specific formats that can disappear over time.

As proved by the setup of the first reportings under Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and Cleaner Air For Europe (CAFE Directive), basing a reporting format on the Inspire data specifications (extending them when necessary) and on the OGC services is now feasible.

The future reporting data model will then be based on Inspire concepts and will be UML & ISO 19100 compliant. Each element will have to be clearly defined (an element won't have different definitions) and typed.

A first draft of an Inspire compliant data model has been developed during the exercise of writing this concept paper. The model is available in the document entitled "Draft Inspire compliant data model - (T4-T5)". This should be extended and checked by INSPIRE and thematic domain experts: it can be an EEA ETC/ICM work for 2014.

Main outcomes of this exercise are the following:

- Having an UWWTD Inspire compliant data model feasible,
- Reusing Inspire classes does not add extra constraint in terms of data collection
- Doing the model raises domain question: definitions of key notions are missing in the 91/271/EEC itself (ex : Urban Waster Water Treatment Plant, Sensitive Area, ...)
- Cross-cutting terminology issues with other EU directives are quickly encountered. To avoid confusion, all classes defined for UWWTD have their name prefixed with 'UWWTD_'.

The two figures below highlight the major steps to be taken into account when aligning a data specification to INSPIRE's. They are taken from draft 0.3 of the Inspire compliant data model which was done to test the feasibility of basing UWWTD reporting on Inspire data models. Not all fields currently required in the reporting have been recreated.

1. First a mapping between domain terminology and INSPIRE data specifications must be done.

¹ <http://ec.europa.eu/idabc/en/document/3473/5585.html>



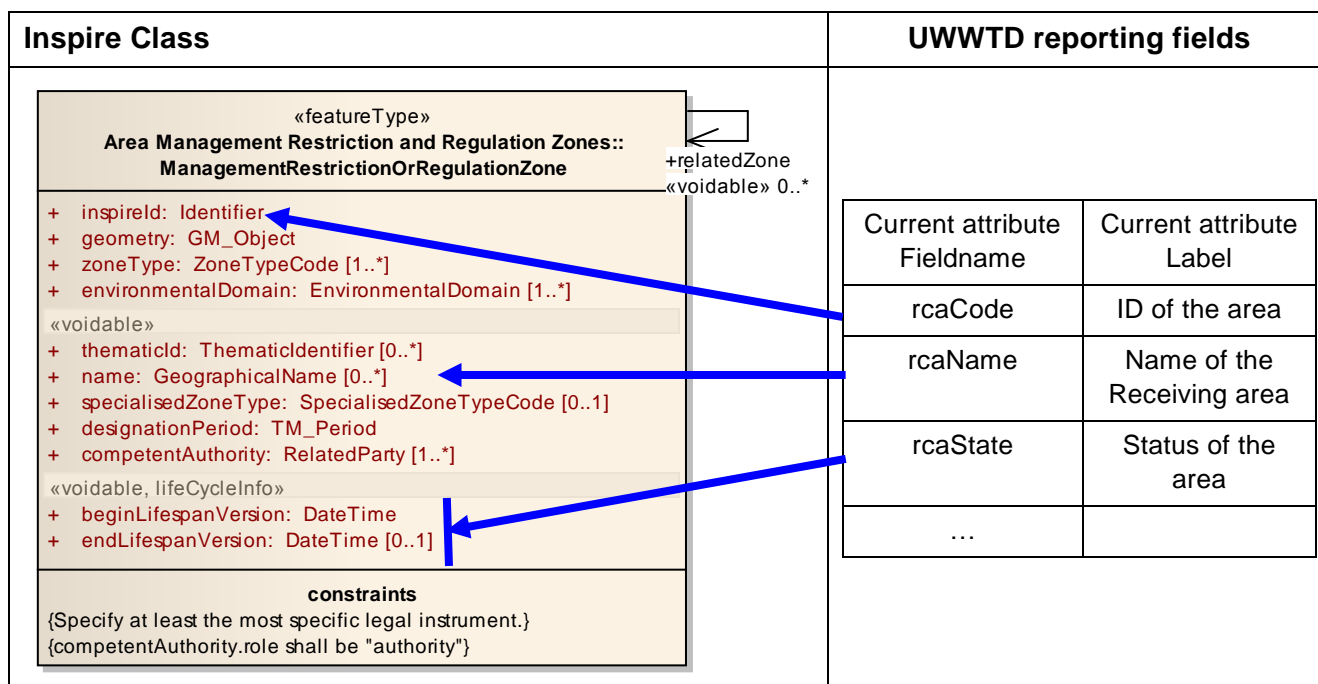


Figure 2 - Mapping between Inspire and UWWTD terminology example

2. Then, extending INSPIRE data specifications to fit extra domain need

Inspire Class	UWWTD receiving area
---------------	----------------------



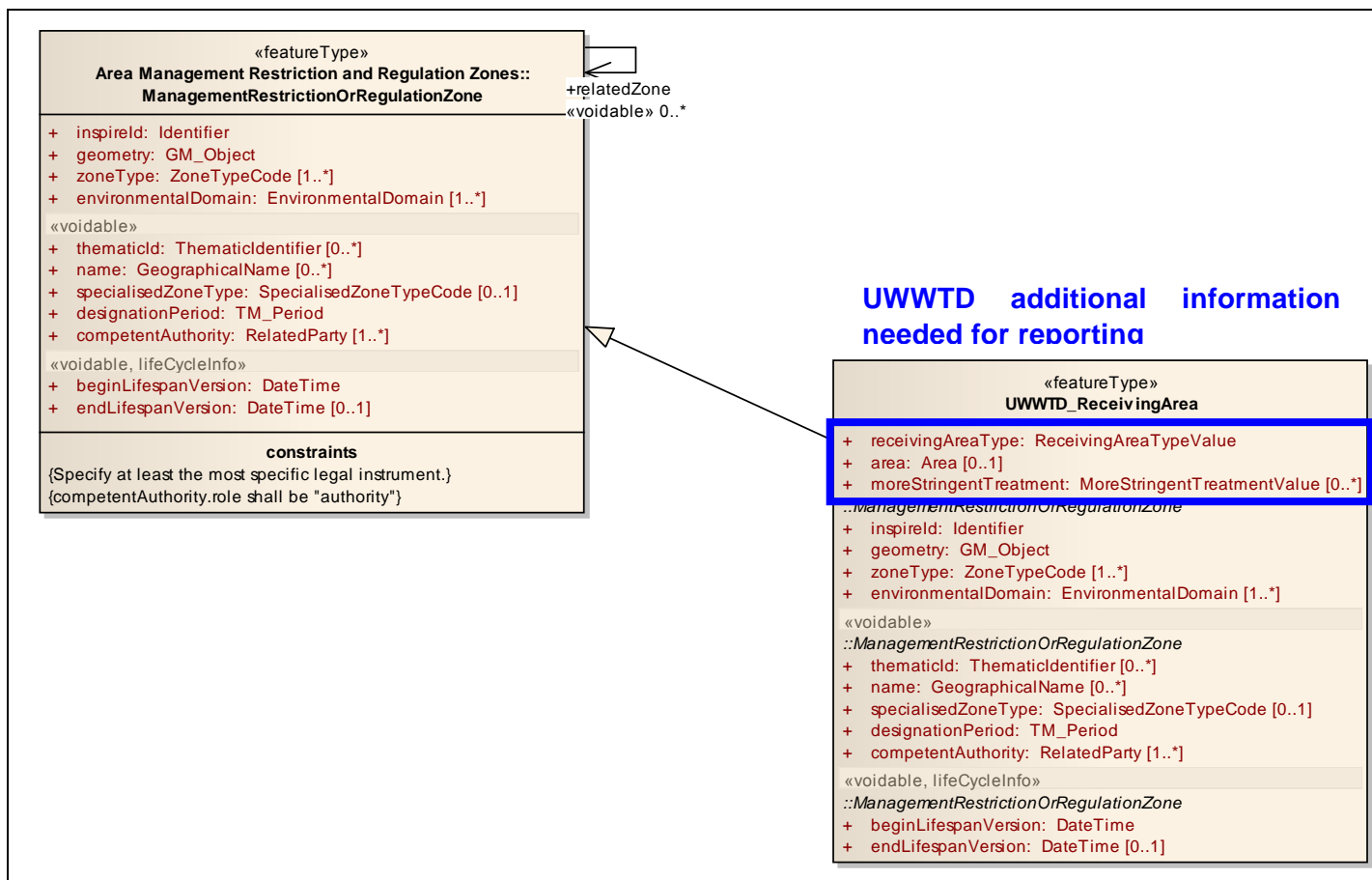


Figure 3 - Extension of an Inspire class to fit UWWTD needs example

II.3. Common communication format/flows

Implementing a distributed system leads to machine-to-machine dialogues in which communication flows and formats need to be standardised.

So far, reporting under UWWTD Art. 15(4) has been done using XML structured files, which could be generated by an UWWTD reporting tool (DEM-tool) provided by the EEA via the Topic Centre ETC/ICM (MS Access database). This XML pathway should be now supplemented with new alternative pathway exchanged using webservice.

The purpose of the UWWT SIIF being mainly to supplement such XML structured data sets between MS and the COM, Open Geospatial Consortium Web Feature Service (OGC WFS) is a technical solution to address this need. Other option, such as automated “file harvesting” (as opposed to calling a webservice) via Reportnet should also be assessed.



To be more specific, information flows will be structured in a specific XML grammar: Geography Markup Language (GML). The structure to be applied in the GML file (xsd) will be generated based on the semantic defined in the Common Data Model. Strict ISO rules define how to generate an application schema the GML/XML structure from a ISO-19100 compliant data model.

It is important to stress that the pre-existing split between XML and Shapefile will be useless, the vehicle of the information can be the same. Every information type exchanged by the combination of those two formats can be exchanged in GML files (even polygonal, polyline, complex datatype information). There is no need to differentiate between what is spatial and non spatial anymore.

II.4. SIIF as a support to e-reporting

As stated in the previous part, the content exchanged within the SIIF will be based on the information necessary to realise the reporting to the Commission. Thus information flows will be able to fulfil reporting constraints.

Full webservice access to UWWTD SIIF information will be the basis of the e-reporting. To achieve that goal, the reporting from MS to the Commission will be done using OGC Web Feature Services (WFS) compliant to the common data model's application schema.

Being a shared information system, the way this exchange should occur has to be agreed precisely.

By definition a WFS is a pull/synchronous oriented dialogue. This has 2 major aspects:

1. **pull**: EU-level node has to trigger the data exchange by querying the MS node,
2. **synchronous**: once the MS node replies, the communication channel between both nodes must be kept open until the complete transfer is done.

The first aspect can result in many request "for nothing" by the EU-level node as it is unaware whether new information is available or no. Even if the EU-level node can use OGC Filter Encoding specifications to query the MS node on "the most recent dataset available", it will result in unnecessary bandwidth and resources consumption.

A **notification service** should be implemented at each MS node to which the COM should subscribe. Various technical solutions could be chosen: from a simple RSS flow to OGC's.

OGC Sensor Alert Service (SAS) and Web Notification Services (WNS) are best practice documents from 2006. Since then, activity continued on Event Service within OGC resulting in the setup of a Publish / Subscribe Standard Working Group (PubSub SWG²) in 2010. A review of the

² <http://www.opengeospatial.org/projects/groups/pubsubswg>





current state of work on that aspect as of the end of 2011 has been published by the chair of the PubSub SWG: OGC 11-088r1³.

The second aspect (being a synchronous data exchange) can suffer from timeouts especially when country wide datasets are being exchanged. WFS are efficient when used for on-the-fly queries, or to extract part of a database. Maximum dataset tolerated size/transfer time in a standard WFS exchange according to the UWWTD SIIF data model should be identified. Testing complete e-reporting with the bigger countries could help on this. If proven necessary, **asynchronous web-services** exchanges should be investigated to ensure the data transfer will occur properly in every situation.

Synchronous/pull services can be considered more intrusive within the MS architecture by person in charge of the IT security in the structures in charge of reporting at MS level. Whether push/asynchronous exchanges should be investigated has to be discussed between MS and the COM.

Moreover, due to the nature of the information exchanged (legal nature of the act of reporting), it has to be discussed whether various access rights to the system should exist. Not having such credential levels, would ease the building and using the SIIF. Which information field can be sensitive has to be identified. For example, the cost of a project before the attribution of a market can be of sensitive nature in order to avoid introducing a bias the result of the consultation.

Inspire technical architecture commonly identifies a GeoRM (**GeoRightManagement**) level on top of all the web services. This layer enables providing different access rights to the webservices. It will have to be discussed between Member States and the COM whether certain elements to be exposed at MS level for the reporting should only be accessible to the COM/EEA. The webservice endpoint being public, whoever wants to access to information behind the service can simply query it.

The notion of **reporting envelope** will also have to be discussed between COM and MS. This aspect inherited from previous reporting eras including physical exchange of document is to be redefined in a service oriented architecture. As a structuring element of the current reporting process this functionality should be reproduced in the SIIF be it in the data model or the technical exchange between services.

Eventually, the COM and EEA currently supports MS in the reporting process providing them with tools. Those usually take the form of a Microsoft Access database implementing the reporting data dictionary and then able to generate XML compliant to the reporting schema.

³ https://portal.opengeospatial.org/files/?artifact_id=45850





What type of tool can now be developed by the COM and EEA for the MS have to be discussed between Member States and the COM. Various proposals are studied in chapter IV “SIIF deployment”.

The reporting system as we know it today will evolve both in the content and the technologies used. To ensure coherence in the dataserie it will be of the utmost importance to ensure that a proper mapping and development of transfer mechanisms will be done between the previous data model and the new one. This aspect is detailed in chapter II.12 “Ensuring backward compatibility”.

II.5. SIIF as a support to reporting via Reportnet

Previous chapter explores aspects relating to full service based reporting between MS and EU nodes.

Another solution that has been discussed between the consultant, DG-ENV and the EEA is the use of the Reportnet as it is currently deployed.

Reportnet is an archive oriented system that can be fed by harvesting a file located at a MS node. Thus it could point to the complete national GML (XML) file corresponding to a MS reporting.

This could help solve the synchronous/asynchronous service issue depicted in the previous chapter.

On the other hand the way the entire national GML dump of the reporting is generated for an MS will have to be investigated.

II.6. Data management rules

Data duplication should be avoided at all costs (apart from caching when it aims at improving system performances).

Thus reference data banks/sources and corresponding data flow must be identified clearly along with their respective roles within the SIIF and WISE.

Specific emphasis should be put on providing stable external unique identifier when referring to reference datasets within the SIIF. This aspect is crucial to ensure proper link between reportings, really assess trends and for the perennality of the system.

This is highly linked to data’s life cycle (its successive versions). Inspire Generic Conceptual Model has defined a specific type to exchange identifiers. One of his elements refers to the





version of the instance. Life cycle rules have been explored by WISE GIS Guidance document (ex: in which situation a new identifier should be assigned, ...). Such rules should be adapted to UWWTD SIIF - also being a part of WISE.

UWWTD reference data elements (i.e. receiving areas, agglomerations, UWWTPs, discharge points, food- processing industries) with stable external unique identifiers were clearly defined for the current UWWTD reporting. In the context of the UWWTD SIIF their corresponding data flows must be identified clearly along with their respective roles within the SIIF.

WISE GIS Guidance recommendations on life cycle have already been partly implemented in the current UWWTD reporting. Two aspects have to be taken into account in this context for the future UWWTD SIIF:

- The approach of historic data management has only been established partly in the UWWTD Art. 15(4). Reporting and needs to be developed further. For each of the UWWTD reference data elements the status of can be given: 'active' means that the reference data element is relevant and in operation for the reporting period. In case a reference data element marked 'retired' was replaced by another reference data element in the next reporting cycle, the 'old' reference data element has to be reported again, but as 'inactive'. So far, the approach of historic data management had already been established and was used correctly by most of the MS.
- The second important element for historic data management had however not yet been implemented in the data model. In case an 'old' reference data element was replaced by another one, the predecessor of the new element should be additionally provided. The following example presents the situation in more detail:



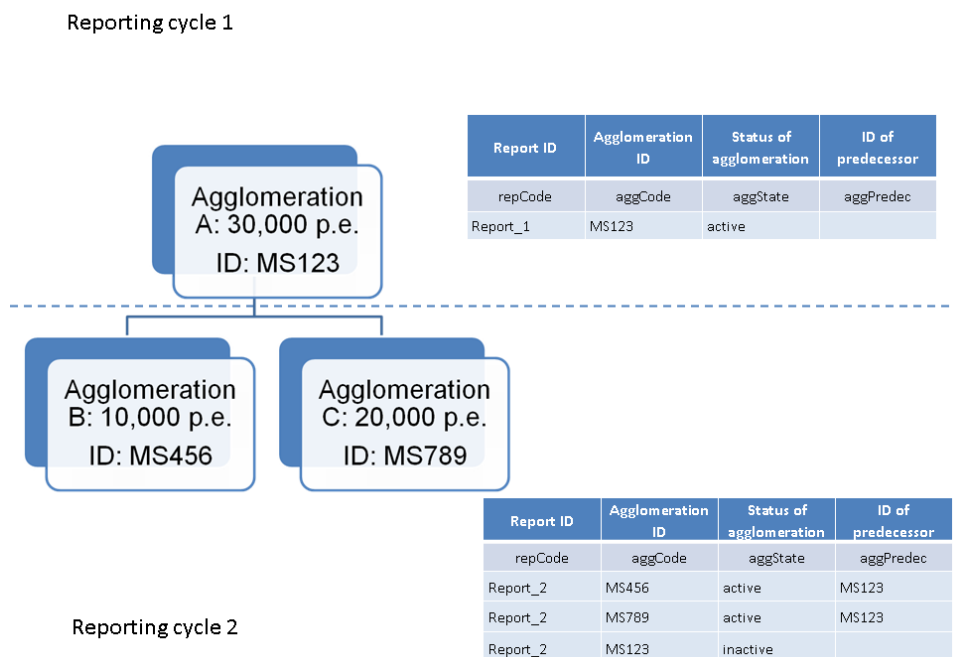


Figure 4 – Predecessor/successor situation example

Further life cycle rules (e.g. in which situations should a new identifier be assigned) need to be discussed between the MS and the EC.

II.7. Documenting resources of the system

Data and services are resources of the future SIIF. In order to ensure proper discovery and use of those, metadata must be available for both.

Each content (geographic, documents) should be documented by metadata. Metadata uploaded in a metadata catalogue that will be queried by the SIIF node interface. The metadata catalogue should also provide an end-point to which the “external world” can connect using open standards:

- For geographic metadata Inspire Implementing Rules on Metadata should be followed and the advised open standard is CSW (minimum version 2.0.2)⁴,
- For document, three ways to exchange them between nodes have been identified :
 - A simple link to the document URL,
 - Using linked data,
 - Exchanging document metadata: in this case, Dublin Core metadata is advised (it will have to be discussed whether there is a need to set a Dublin Core Qualified profile). The advised webservice is OAI-PMH (Open

⁴ <http://www.opengeospatial.org/standards/specifications/catalog>





Archives Initiative Protocol for Metadata Harvesting), as the extension from this towards rdf-a implies a limited effort. This third option could be implemented with both html and xml schemes or solely with XML depending on the needs on the documentation.

- For geographic information: view and download webservices should be set up according to Inspire Implementing Rules, where applicable. Those webservices should also be documented by metadata,
- In those cases where Inspire Implementing Rules have not yet been adopted, but the technical guidance documents have been finalised along with corresponding application schemas, there is an opportunity to test the currently available referential spatial data for Inspire conformity.

II.8. Persistency within the system

Stability of the future UWWTD SIIF overtime will also be covered by the persistency of the various resource locators used. This covers the ability via the SIIF to resolve/dereference:

- Data model(s),
- Stable identifiers assigned to reference datasets,
- Element of the code lists,
- Other resources types: services, documentation, ...

Persistent Uniform Resource Locators (PURL) should be defined and deployed for all those elements. A guideline has been produced under the Interoperability Solutions for European Public Administrations programme (ISA programme)⁵.

Next item that will have to be defined is where the PURL should point to. For example: when an external unique identifier refers to a specific Waster Water Treatment Plant in a given reported version, is the reference node the MS or the EU one? There are no real pros and cons on the IT side. The answer to this question is more on the responsibility aspects. What is the reference version has to be jointly defined with MS. Next question will then be: who has the responsibility to maintain the persistent addresses ?

II.9. Quality of service

UWWTD SIIF being based on Inspire network services rules. Inspire quality of service rules should apply.

⁵ <https://joinup.ec.europa.eu/community/semic/document/10-rules-persistent-uris>





Within Inspire, Annex I from Commission Regulation 1088/2010/EC of 23 November 2010 amending Regulation 976/2009/EC as regards download services and transformation services defines quality of service criteria relating to performance, capacity and availability.

Moreover the SIIF infrastructure must not degrade partners' systems performances and security to respect the non-intrusiveness principle.

II.10. Specific SIIF EU-level node IT content

Even if the system is to be distributed, the EU-level node will have to play a specific role of gatekeeper of the system. It will have to:

- Host the application schema (xsd) of the common agreed data model. WFS set up at MS level should refer to it in order for their GML to be validated,
- This application schema will define shared code lists to be used. Those must also be available in the EU-level node.

Moreover, it is identified in the herewith joint *Annex I: use cases* that common QA/QC tests and procedures to assess compliance to the Directive should be jointly defined between MS and COM. Most those rules and procedures should aim the automatisisation of the process.

In order to ensure a fully shared implementation of those rules and procedures, the UWWTD SIIF EU-level node should provide both validation (QA/QC) and compliance services. Those services should be able to receive as an entry the MS reporting webservice endpoint (or an xml file compliant to the reporting data model) and provide the result of QA/QC and/or compliance assessment as an output.

This implementation at the EU-level node is necessary as it will ensure that QA/QC and compliance assessment procedures have a reference IT deployment. Otherwise, with 27 Member States implementing the same reference documents, there is a high risk some implementation have a different understanding; thus providing different results.

More specifically:

- Based on a QA/QC document commonly agreed between COM, EEA and MS, a data validation endpoint should be made available. It should consist in an XML parser that would both validate Member States' WFS against the data model and also run specific schematron rules to check the content quality
- Following the same rationale, provided compliance assessment can be fully automatised based on Member States' WFS, a compliance assessment endpoint should also be made available at the COM level. This implies MS web services provide access to the underlying data used for UWWTD compliance assessment in the MS. It should consist in a set of





schematron rules to be applied to the reporting data flow and should take into account the options given in the UWWTD and corresponding choices made by MS.

It has to be stressed that the remote compliance assessment described above should be considered a 'pre-compliance' assessment. It will help MS enhance the quality of the reported datasets and speed up the reporting process. However, running an official compliance assessment centrally done at EU level still seems important before publishing the aggregated European datasets

II.11. SIIF and MS Inspire compliancy on UWWT

Among many Inspire implementing Rules, Member States will have to expose under annex III.6 "Utility and governmental services" data theme a layer corresponding to their **EnvironmentalManagementFacilities**.

The definition of an EnvironmentalManagementFacility is the following: "*A physical structure designed, built or installed to serve specific functions in relation to environmental material flows, such as waste or waste water flows, or a delimitable area of land or water used to serve such functions.*"

The accompanying description states: "*In the context of waste management the "specific function" may be a waste recovery or disposal operation. Typically, waste management sites and waste management installations (such as incineration plants, landfills or storages) get distinguished. Multiple waste management installations may be found at the same site. Waste management installations can be a part of other waste management installations. The functions considered for the Environmental Facilities Theme fall mainly under the NACE rev. 2 category E "Water supply; Sewerage; Waste management and remediation activities".*

Water Treatment Plants fall clearly under this data theme. Other concepts like 'Agglomeration' or 'Sensitive Area' do fall under annex III.11 "Area management/restriction/regulation zones and reporting units" data theme.

Those aspects have been introduced in chapter II.2 Communication content: a common data model" and detailed in the document entitled "Draft Inspire compliant data model - (T4-T5)".

From data management and financial point of views it would be nonsense to ask MS to expose somewhere their dataset according to Inspire on the one hand and on the other to realise an e-reporting on another webservice endpoint.

This situation would lead to duplication of work on the exact same datasets and, obviously to incoherence or non-alignment of the datasets exposed.





That's another reason for which the reporting should be based on Inspire data specifications (extended where necessary).

Using MS SIIF nodes to do such Inspire based reportings would allow MS to fulfil two obligations at a time:

- UWWTD reporting,
- INSPIRE compliance.

This will really be an application of the motto "Report once, use many times".

It has been highly discussed during INSPIRE data specifications process that:

- INSPIRE scope is not to specify reportings: For sure, but INSPIRE data specifications do provide the necessary bricks to set up an e-reporting along with an obligation to expose content directly or indirectly related to the UWWTD. As proven by the UWWTD Inspire compliant data model exercise along WFS exchanges testing it is feasible to extend those bricks when necessary. Exposing datasets that do cover Inspire + reporting specific requirement will ensure to fulfil two obligations at a time,
- INSPIRE does not apply to the Commission bodies but to MS: That's also true but INSPIRE is the good incentive to streamline data availability at EU level: between MS, between MS and the COM but also between the COM bodies and from the COM towards MS,
- The WISE products and services provided at EU level, based on MS reported / shared data, should also comply with Inspire implementing Rules, where applicable.

II.12. Ensuring backward compatibility

Transition to the new system will take place over a long period of time. More over UWWTD SIIF will have to be able to deal with datasets exchanged in the current formalism.

There will be a need for developing a transformation service, presumably 2-ways:

- One way from current European dataset in Waterbase into an INSPIRE compliant data model,
- but also the other way from the new Inspire compliant data model back to the current XML reported via Reportnet into Waterbase.

II.13. Forward looking perspective

Other aspects have been identified in the course of writing this concept paper.





Interoperability using ontologies has been identified as a potential future improvement of the UWWTD dataflow into a SIIF approach. This subject is under test within some MS and also at the EU level (at EEA : SENSE 1 & 2 projects).

Exposing information using rdf endpoints at each UWWTD SIIF MS node, realising possible ontology alignment with ontologies available at the EU level and finally running semantic reasoner on top of this was discussed.

This approach would enable more flexibility as the reporting web service wouldn't need to respect fully a data model. On the other hand for each MS, the ontology alignment with the common EU ontology should be done at the EU level.

It could also support a dual approach in which:

- Everything that is relating to reporting obligations should be defined in the INSPIRE, ISO 19000 compliant data model and made available via WFS services
- Ancillary information could be exposed using rdf endpoints in its native (MS) structure.

However, it is considered this approach needs more testing by its respective communities. It is deemed too early to deploy it in production in the short term future UWWTD SIIF.

III. Nodes front end

III.1. Website accessibility

One objective of the UWWTD SIIF initiative is to improve dissemination to local levels. This implies that the information should be accessible to the general public.

Even if a SIIF node could be accessed by various user types (public, wastewater expert, COM,), first level access should be tailored to the general public. SIIF nodes homepage should better be made up of simpler cartographic interfaces along with easy search possibilities (ex: by city, agglomeration...). Such homepages have been already identified in tasks 1 & 2.

The interface could then allow for deeper search, access to more detailed information to suit UWWTD expert needs, and excluding the information declared as confidential by the owner of information for not registered users (a similar principle on confidentiality applied for E-PRTR Regulation could be considered here).

Such a solution is preferred to having homepages forcing the user to attach himself to a profile (ex : "what is your profile ? "General public", "Public association/Media", "UWWTD expert", ...)





and is already the approach adopted at the EEA level. It is not a problem to propose several kind of presentation of information. A simple presentation for public and then possibility to have access to more detailed information.

Links to regulation (including relevant websites), statistics, dashboards, other SIIF nodes (from MS to EU or EU to MS) which are described below should also be easily accessible.

III.2. Content visualisation

In order to provide a better access to the information available on a SIIF node, various types of representation should be explored:

- Cartographic :
 - UWW Treatment Plants should be accessible in a first place. A first symbology could consist in having point size proportional to the treatment plant capacity and point color range representing its compliancy to the Directive, as already used on the WISE viewer
 - Other geographic information could be overlaid to this one. First, other information asked for in reportings (ex: sensitive areas, agglomerations,..). Then information coming from other reportings (ex: WFD water bodies),
 - Ideally each layer that can be displayed on the interface should be queryable (especially those directly related to the UWWTD reporting). For example, if the user wants to know more on a sensitive area, it should be able to ask for more information on the object by clicking on it.
- Graphics:
 - Evolutions, trends of information available at TPs, agglomerations should be available using bars or pie charts graphics.
- Ontology
 - Being understood here as covering controlled vocabulary and mapping between models, the recent development of the thematic ontologies and the availability of a norm for UWWTD language allow considering it technically feasible to build a specific ontology to allow for the exploration of the information with a sound technical basis. See also II.13 of the document.

III.3. Access to information

Every information type available at the SIIF Node and directly related to UWWTD should be accessible using the discovery / view / download approach promoted by Inspire:

- Each content (geographic, documents, ...) should be documented by metadata. Every service set up at a SIIF node should also be documented by metadata. Metadata available in a metadata catalogue that will be queried by the SIIF node interface. The metadata catalogue should also provide an end-point to which the “external world” can connect





using open standards. This second aspect has already been presented in chapter II.7 “Documenting resources of the system”.

- Information should also be made available using tabular format downloadable from the SIIF interface.
- CodeLists: code lists used within the UWWTD SIIF should be accessible at the UWWTD SIIF EU-level node. At least in a human readable form, ideally in a format that allows direct links from the XML files used in reportings (XML, XSDs, ontologies). The way codeList repositories should be made available online is a important topic of discussion within Inspire Drafting Teams. Hopefully, technical choice will be made within a short period of time. UWWTD SIIF will implement the solution decided/advised at Inspire level.
- Maps / indicators: maps and indicators available at a SIIF node should also be documented and searchable.

For ancillary information added to the SIIF node it is recommended to provide a link to its source only to avoid duplication of information. Such information could either come from other reportings or Inspire data flows.

Aligned with the rest of this document, it is highly recommended that information layers created at the COM level and accessible via a SIIF node should be Inspire compatible. Which means, wherever possible, based on Inspire concepts and extending them when necessary.

III.4. Access to documentation

The data model describing the content of the information exposed on each SIIF node should be accessible.

It should be available at least as a downloadable document. Ideally an interactive form allowing dynamic search and navigation will help the end-user.

The HTML view of UML models provided under Inspire website could be a starting point for such an interactive tool. For an example (see: <http://inspire-twq.jrc.ec.europa.eu/data-model/draft/r2563/>), the “*interactive HTML view of the complete UML data models. This view includes detailed definitions of spatial object types, data types, enumerations and code lists and UML class diagrams*”

III.5. Helpdesk

Direct access to a helpdesk must be provided by a SIIF node. The user must be able to raise questions on:

- the use of the SIIF node interface,
- the structure of the information available,





- the actual values used in the datasets,
- the definition of terms used in the website in case they are not exactly the same as those defined in the data model.

III.6. Data flows representation

The interface should clarify the roles of the various entities involved in UWWTD reporting: from the lower level acquiring the information at the TP level to structures in charge of checking the information at the COM level via the entity in charge of the reporting at MS level.

III.7. Links between EU and MS nodes

SIIF front end should provide links between EU and MS nodes.

Its cartographic interface should allow displaying both EU layers and national ones.

As for the content when asking for more:

- EU information it should be structured according to the EU data model,
- National information the system should redirect to MS SIIF level (if the initial request is done on the EU node) and the information should be structured according to national data models.

National SIIF nodes should structure their information as follows :

- information exposed via WFS webservices should be, at least structured according to the EU data model,
- their interface should display the content according to local data models.

III.8. UWWTD legislation overview

A clear view of the application of the EU legislation should be provided by each node. For example, an UWWTD SIIF MS node will provide dates of transposition of the EU legislation in the national law along with the various procedures specific to national legislation that have been set up to enforce UWWTD.

Access to the various infringement cases and their final conclusion should also be provided.

III.9. Specific SIIF EU-level node interface

All the IT elements defined in II.10 “Specific SIIF EU-level node IT content” to be accessible on a machine to machine basis should also be provided in a human accessible way.





The user should be able to browse, search, and access:

- The agreed data model,
- The code lists and reference datasets,
- Upload an XML file compliant to the reporting schema and test it either the QA/QC procedure and/or the compliance assessment.

IV. SIIF deployment

The proposed UWWTD SIIF is based on the most up-to-date/robust technologies and practices available at the time of writing this concept paper.

This, on the one hand, ensures proper enhancement of the current reporting system. On the other hand there is a risk that IT levels/competencies and necessary funding are not available in each MS.

Eventually, the SIIF is entirely built on the Internet. Local network limitations could impair the deployment of the SIIF.

Various solutions have been identified to circumvent the above mentioned issues. They are detailed below.

IV.1. Core bricks

Not all the UWWTD SIIF bricks have to be deployed at once. Identifying core-bricks / second level / third ... from an IT perspective will feed the UWWTD SIIF implementation roadmap.

Core bricks identified so far are:

- The Inspire compliant data model,
- Setting up WFS reporting webservice,
- Setting up reporting using ReportNet,
- SIIF Node front end available at the COM level and in volunteer MS,
- Communication between nodes,
- XML transformation service between the current data model and the new Inspire compliant one.

Second level:

- COM validation endpoint : QA/QC service,
- COM compliance assessment endpoint,
- CodeList repositories.

Third level:





- Notification services,
- Asynchronous services,
- Permanent URL (PURL) resolver.

SIIF node frontend functionalities (the interface) can also be prioritised so that a first level of information accessible online is reached quickly. Then progressively new functionalities can be added.

Those aspects will be further developed in the UWWTD SIIF implementation roadmap.

IV.2. Promoting an IT water community

In order to support MS in deploying their SIIF node, COM and EEA should investigate how to best set up an IT Water Community. WISE Steering Group (SG) could be the right forum to initiate this and WISE Technical Group (TG), organising WISE GIS/IT workshops, could be the forum to implement it.

The current approach consisting in providing a tool along with the corresponding helpdesk could be enhanced with a closer cooperation between IT services in both COM and MS.

Moving toward e-reporting will involve more competencies in MS. Tools currently used in reportings can be deployed on a single laptop. On the other hand, tools proposed to be deployed in the future UWWTD SIIF must be part of a real IT service oriented infrastructure the visible part of which will be the SIIF frontend.

In order to foster such a community, the following approach should be studied:

- A first set of tools answering to this concept paper IT bricks should be identified, tested & improved (if necessary) by the COM and voluntary MS.
- This could in turn be transferred to the IT water community via training sessions and workshops. The transfer could be done on a tool by tool basis.

Possibilities provided by virtualisation should also be studied. It is now common to run IT architectures on virtualised servers. Following such an approach, a virtual image of SIIF node server containing all the identified pre-configured tools could be downloadable from the UWWTD SIIF EU node⁶.

- Once the transfer is done, it will be necessary to provide a helpdesk to support MS. Instead of a formal closed helpdesk where MS don't have access to questions raised by the others. Mailing list oriented or ticket services approaches should be advised as they stimulates a cooperative approach from MS toward COM but also between MS.

⁶ Rough illustration: it is like cloning the entire server content on a DVD and read it in another IT infrastructure with the proper reader.





Both the COM and MS will benefit from this win-win oriented approach. The entire IT backbone of the UWWTD SIIF revolving around INSPIRE it will also benefit from and to the IT development done under this umbrella. This will avoid duplication of efforts in this domain as, even if the thematic content is Directive specific, IT approach share commonalities with others.

Eventually, being based on open standards, documentation, and open source software tools, the UWWTD SIIF could be replicated at MS sub-level (all aggregation units of lower level: region, county, länder, RBD, sub-unit).

As far as the software licences are concerned, an OpenSource approach is highly recommended. This will ease the transmission of tools to MS, simplify IT collaboration when it comes down to improving the source code of these tools and support creating a “community effect”. *“Information sharing and processing should be supported through common, free opensource software tools.”* is one of the 7 SEIS principles. Moreover, as stated the European Interoperability Framework: *“Open Source Software (OSS) tends to use and help define open standards and publicly available specifications. OSS products are, by their nature, publicly available specifications, and the availability of their source code promotes open, democratic debate around the specifications, making them both more robust and interoperable. As such, OSS corresponds to the objectives of this Framework and should be assessed and considered favourably alongside proprietary alternatives”⁷.*

Pre-existing OpenSource tools already exist for most of the IT bricks identified. Whether this collaborative approach can be done with pre-existing licences or under the ISA EUPL (EU opensource licence) is to be clarified with the COM legal services. Studies on this topic have been issued by the European Commission under the ISA Programme (Interoperability Solutions for European Public Administrations). The Guideline for Public administrations on Procurement and Open Source Software (updated June 2010)⁸ seems an interesting starting point

IV.3. Centrally maintained toolset VS fully distributed

Provided the other solutions are not sufficient enough to help all MS move to the UWWTD SIIF, a complementary approach would be for the COM/EEA to deploy a “centrally maintained toolset”.

It will provide the exact same functionalities a UWWTD SIIF MS node is expected to, but hosted on the EEA IT infrastructure. The term “functionalities” refers here to both the IT service oriented architecture and the SIIF Node frontend (the interface).

MS could remotely manage their own datasets with specific credentials.

⁷ <http://ec.europa.eu/idabc/en/document/3473/5585.html>

⁸ <https://joinup.ec.europa.eu/elibrary/case/guideline-public-administrations-procurement-and-open-source-software-updated-june-201>





Acting as a UWWTD SIIF MS Node for a given MS, the overall rationale of the UWWTD SIIF would be respected.

This scalable approach will ensure that all MS do have a UWWTD SIIF Node deployed according to the agreed roadmap.

IV.4. A transition period

From the discussion between the consultant, DG-ENV and the EEA it has been concluded that real reportings need to be run in order to clarify which IT solution fits MS needs.

The three solutions identified above are all based on the same Inspire compliant data model:

- e-reporting using WFS (see chapter II.4 “SIIF as a support to e-reporting”),
- using Reportnet tools (see chapter II.5 “SIIF as a support to reporting via Reportnet”),
- using a centrally maintained toolset (see chapter IV.3 “Centrally maintained toolset VS fully distributed”).

Thorough testing with pilot MS will validate how to deploy those three solutions.

However it is foreseen that those three solutions should be proposed to MS in parallel as possible information channel to carry out the reporting. After 1 or 2 reporting cycles, the experience gained at MS level should help pinpoint which of the three solutions is the most pragmatic for MS.

This approach is deemed to be the most realistic in order to take into account the diversity of the situations encountered in MS both on the IT and financial aspects

V. UWWT SIIF and other SIIFs coherence

The purpose of the mother document is to define the future UWWTD SIIF. The exercise won't be complete without a more global vision. Tasks 1 and 2 have identified certain issues that should ideally be addressed at an upper level.

V.1. Cross-cutting information systems issues example

Below are examples of issues located between information systems

- Each reporting theoretically concerns one Directive which defines its own domain terminology without taking into account others. As a result the exact same words used in different directives can refer to highly different realities. WaterBody and Agglomeration examples have already been introduced in this document (see example for Agglomeration given in chapter II.2 “Communication content: a common data model”).





From a data management point this is a crucial issue as those terms often end up in the current reporting sheet. It is then easy for people to “take shortcuts” and deem two things are the same only because they share the same ‘label’. This situation has often been encountered. It undermines data quality and costs a lot of time clarifying what is asked for/exchanged not mentioning that it impairs cross-domain or cross-reporting work.

- Additional reporting of the same dataset in different reportings should be avoided at all cost. For example WFD register of protected areas should be able to refer to sensitive areas already reporting under UWWTD not requiring MS to report the content again. Not doing so there is a risk dataset reported are not exactly the same or in the same version for a given reporting cycle.

Moreover there is a strong overlap in the content between UWWTD reporting, Eurostat Joint Questionnaire and also PRTR. They all have their specific reporting cycle, and can involve different contact points at MS. This can induce incoherence in the data transmitted at the European level.

- IT is transversal per se but there are however many different working groups that deal with the IT aspects of reportings in the water domain :
 - Working group for WFD reporting : mostly content,
 - WISE TG : so far mostly been IT freshwater,
 - Working group on Data, Information, Knowledge Exchanges (DIKE),
 - Working group for MSFD reporting,
 - other MSFD sub-group on WISE, EMODNET and decentralised systems.

V.2. Cross-cutting IT governance

V.2.1. The approach

As illustrated in the examples above, other obstacles hinder proper data exchange because the sum of the current reporting systems lacks a joint IT approach.

DG ENV / EEA / Eurostat have done a lot to streamline reportings be it within working groups or in the deployment of tools (ex: Reportnet).

In order to support that endeavour part of the recommendations of this concept paper could be applied at an upper level. Indeed, it has been clearly stated that the experience gained at the UWWTD SIIF level will be generalised to other SIIFs later on. When setting up those other SIIFs it could then be the relevant time to promote the following roles:

- An EU level interoperability board to ensure the overall IT consistency (data models, SIIFs architecture, ...),
- One normalisation structure per information silo (inland water, air,) responsible for the consistency in their domain,





- One SIIF per Directive if deemed necessary.

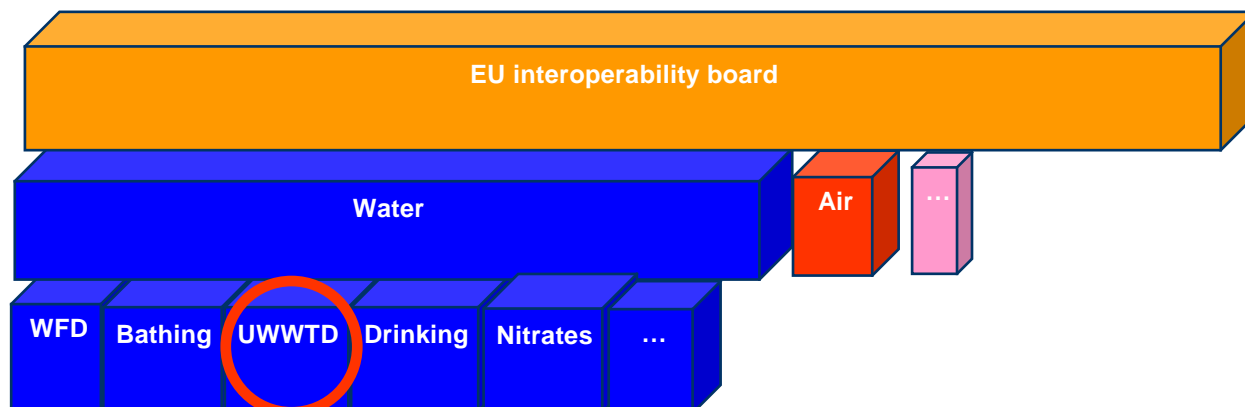


Figure 5 - Sketch of a European cross-cutting IT governance

It has to be mentioned that part of this approach is already on its way as Bathing, Urban Waster Water Treatment and Drinking are now regrouped under the “Water Industry” denomination.

As a consequence of this proposal, each SIIF will still be SEIS pillar in its respective domain but the proposed approach will ensure the “Shared” part of the system is achieved when using the sum of the SIIFs.

Having such board responsible for the overall consistency of the IT system already exists in structures like the OGC, Inspire and also in some countries. For example in France there is one structure in charge of the standardisation in the water domain (the Sandre⁹) and another one on top (the COVADIS¹⁰) created by ministries of Environment & Agriculture. Both work in close cooperation on the IT aspects so that information is not duplicated and to optimise the reuse of already created concepts and reference datasets.

Inspire Drafting Team on Data Specification addresses already part of these coherence questions when it comes down to Inspire data specifications. The Inspire Generic Conceptual Model (GCM) is already a mutualisation of shared concepts that shows the way interconnected SIIFs should go. For example, the ActivityComplex concept (FeatureType), used by the proposed UWWTD Inspire compliant datamodel is shared between many other Inspire themes. Some of those concepts can be reused in other EU Inspire compliant reportings.

⁹ Service d'Administration Nationale des Données et Référentiels sur l'Eau: <http://www.sandre.eaufrance.fr/?lang=en>

¹⁰ COmmission de VALidation des Données pour l'Information Spatialisée : <http://www.cnig.gouv.fr/Front/index.php?RID=120>





The upcoming Inspire “Maintenance and Implementation Framework” and “Maintenance and Implementation Group” (MIF/MIG¹¹) will have to be associated with the proposed SIIF anyway. Be it applied with a UWWTD SIIF only or an integrated SIIF approach.

Proper responsibilities and communication flows should be drawn between the various information silos (inland water, marine, air,...), the normalisation structures, EU interoperability board, Inspire gatekeepers...

Last but not least, as stated by Recommendation 23 of the European Location Framework White Paper¹²: “*Public administrations should lead or actively participate in standardisation work relevant to their needs*”. Indeed, the work done under Inspire has raised many precise questions on ISO, OGC standards. On the field of water, many standardisation activities are ongoing under the umbrella of the OGC hydrology domain working group that should prove really useful to the European Water Community.

The link created between ISO/OGC and Inspire European Commission contact points should be maintained and enhanced as there is no scale equivalent to such normalisation of information systems as it is occurring now in Europe.

In the Information System domain, Europe has a real potential to be a true driver of standards

V.2.2. Proposal for new roles repartition

Based on the consultant experience, it is deemed important presenting this need for an IT governance from a pure IT point of view.

This idea needs time to mature especially on decision competence issues.

In order to achieve the SEIS objectives and not add extra and unnecessary administrative burden, there is no need to create new organisations. But these IT governance roles should be in the system.

It is highly advised they should be taken on board by already existing structures by extending their mandate.

During the discussion between the consultant, DG-ENV and the EEA the following solution emerged:

- EU level interoperability board: Inspire MIF/MIG,

¹¹ <http://inspire.jrc.ec.europa.eu/index.cfm/pageid/5160>

¹² <http://www.eurogeographics.org/category/cumulus/elf>





- Normalisation structure for the water information silo: WISE SG supported by WISE TG,
- One SIIF per Directive if deemed necessary.

V.2.3. Who will benefit from this approach?

Many levels will benefit from the proposed organisation.

At MS level, it will clarify datasets to be exchanged, in what way, how they are validated / aggregated. It could also help define properly responsible Parties roles. This approach will help make a better use of public money, avoiding data duplication and parallel flows at all costs.

From an EU point of view the exact same benefits can be raised. This will ensure data produced on the basis of MS reports or at EU level will be reused as much as possible.

This approach along with the one described in part III “SIIF deployment - Promoting an IT water community”, will ensure a closer IT cooperation with MS. Both parties will benefit from this, clarifying data exchanges, standards/technologies to be used. It will ensure proper technology transfer in both directions.

More generally all the semantic and technicalities of the data exchange being clarified, organised and documented in a commonly agreed way it will be possible to properly work on the actual content.

All the above mentioned benefits will in turn help:

- Enhance data quality,
- Empowering local levels: as the proposed architecture and organisation will then easily support the deployment of SIIFs in the various domains. It will be then possible to focus on how datasets should be brought to the general public, which indicators can be set up,...
- Eventually improve the state of the environment which is the concrete goal of all the environmental EU legislation. For example it will help making the link between the pressures and the status of waterbodies and hence choose the best appropriate measures to reduce the pollution.

V.3. UWWTD SIIF connexions with WISE and SEIS

Parts developed above identify main functionalities that should be provided by the UWWTD SIIF. As an information system UWWTD SIIF is clearly under WISE umbrella.

Thus, in order to avoid duplication of information systems technical deployment, those should be put next to what is already in place in the EEA IT infrastructure.





This exercise will help:

- clarify the relative position of UWWTD SIIF with WISE,
- thus, de facto explain UWWTD SIIF relation to SEIS (WISE being SEIS water pillar),
- delineate roadmap and workplan for the deployment of the SIIF.

UWWTD SIIF main IT functionalities	Reportnet infrastructure	Reportnet infrastructure in the context of UWWTD
Provide access to Art 15, 16, 17 reporting obligations	Reporting Obligations DataBase	UWWTD Art 15, 16, 17
2006 reporting template + thematic updates + Inspire compatible	Data Dictionary	UWWTD 2006 agreed template
E-reporting using WFS	Data Exchange Module	UWWTD reporting tool/DEM
QA/QC and compliance assessment services	Data repository	Data repository, ETC ICM
Web services available at EU SIIF node	EEA data services	Waterbase UWWTD dedicated content
EU SIIF node front end including cartographic interface	Information products	WISE viewer UWWTD part

The following figure applies the table content to commonly used diagram of Reportnet.

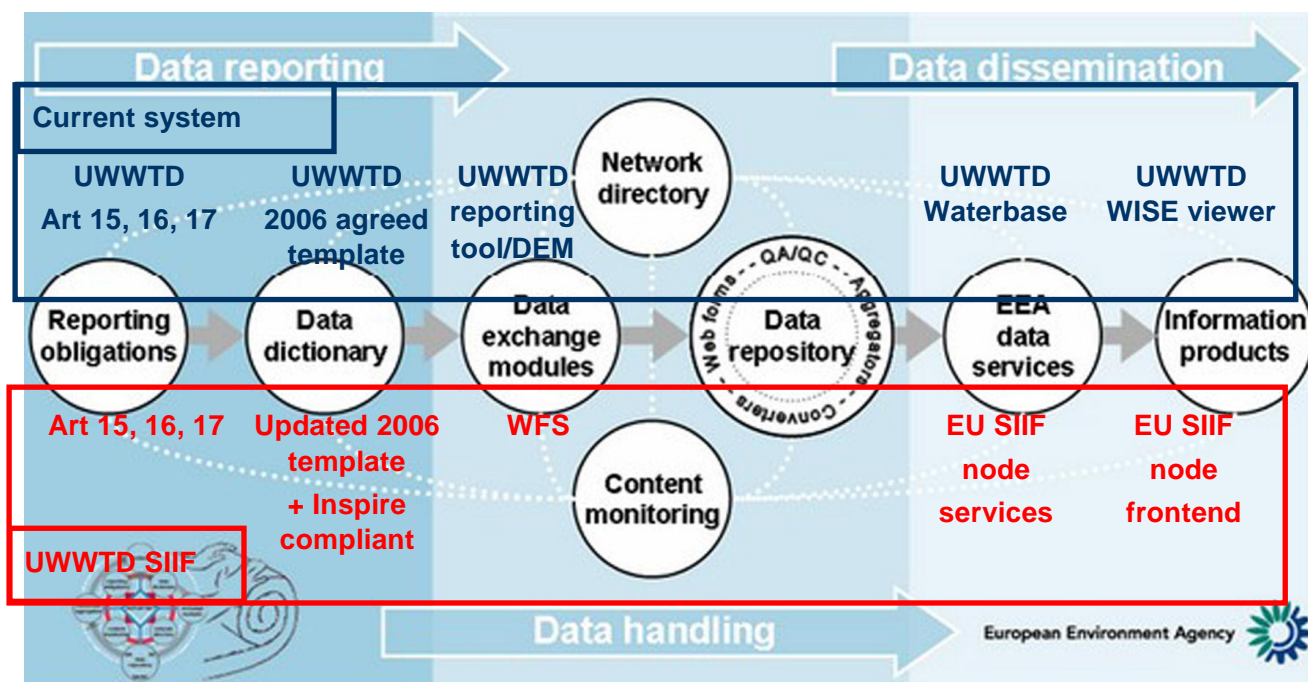


Figure 6 - UWWTD SIIF functionalities applied to Reportnet

Further work on the roadmap / workplan will help clarify connexions between UWWTD SIIF and WISE.