

EU Framework Program for Research and Innovation (WATER-4a-2014 - H2020)



Project Nr: 641821

Applying European market leadership to river basin networks and spreading of innovation on water ICT models, tools and data.

Deliverable D1.3 ***Final Data Management Plan***

Version 1.1

Due date of deliverable: 31/08/2016
Actual submission date: 02/09/2016

Document control page

Title	D1.3 Final Data Management Plan		
Creator	EP_CREAF		
Editor	EP_CREAF, ES_adelphi		
Description	The purpose of the Final Data Management Plan is to provide a definitive analysis of the elements of the data management policy that will be used with regard to all the datasets generated in the project. It reflects the decisions within the consortium about the data produced.		
Publisher	WaterInnEU Consortium		
Contributors	WaterInnEU Partners		
Type	Text		
Format	MS-Word		
Language	EN-GB		
Creation date	28/07/2016		
Version number	1.1		
Version date	01/08/2016		
Last modified by	EP_CREAF		
Rights	Copyright © 2016, WaterInnEU Consortium		
Dissemination level		CO (confidential, only for members of the consortium)	
	X	PU (public)	
		PP (restricted to other programme participants)	
		RE (restricted to a group specified by the consortium)	
	When restricted, access granted to:		
Nature	X	R (report)	
		P (prototype)	
		D (demonstrator)	
		O (other)	
Review status		Draft	<i>Where applicable:</i>
		WP leader accepted	Accepted by the PTB
		PMB quality controlled	Accepted by the PTB as public document
	X	Coordinator accepted	
Action requested		to be revised by all WaterInnEU partners	
		for approval of the WP leader	
		for approval of the PMB	
		for approval of the Project Coordinator	
		for approval of the PTB	
Requested deadline			

Revision history

Version	Date	Modified by	Comments
1.0	06-08-2016	EP_CREAF	Created the basic content of the deliverable
1.0	23-08-2016	ES_adelphi	Introduced some modifications
1.1	01-09-2016	EP_CREAF	Created the final version of the deliverable

Copyright © 2015, WaterInnEU Consortium

The WaterInnEU Consortium grants third parties the right to use and distribute all or parts of this document, provided that the WaterInnEU project and the document are properly referenced.

THIS DOCUMENT IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT OWNER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS DOCUMENT, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Table of Contents

0. Main updates in the Data Management Plan.....	1
1. Inventory of data sources, companies and solutions for the marketplace	1
1. Data summary	1
2. FAIR Data.....	2
2.1. Making data findable, including provisions for metadata	2
2.2. Making data openly accessible	2
2.3. Making data interoperable.....	2
2.4. Increase data re-use (through clarifying licenses)	3
3. Allocation of resources	3
4. Data security.....	3
5. Ethical aspects	3
6. Other.....	3
2. Surveys responses.....	4
1. Data summary	4
2. FAIR Data.....	4
2.1. Making data findable, including provisions for metadata	4
2.2. Making data openly accessible	5
2.3. Making data interoperable.....	5
2.4. Increase data re-use (through clarifying licenses)	6
3. Allocation of resources	6
4. Data security.....	6
5. Ethical aspects	6
6. Other.....	6
3. Product Specification Sheets (PSS)	7
1. Data summary	7
2. FAIR Data.....	7
2.1. Making data findable, including provisions for metadata	7
2.2. Making data openly accessible	8
2.3. Making data interoperable.....	8
2.4. Increase data re-use (through clarifying licenses)	8
3. Allocation of resources	8
4. Data security.....	9
5. Ethical aspects	9
6. Other.....	9
4. Marketplace platform.....	9
1. Data summary	9
2. FAIR Data.....	10

2.1.	<i>Making data findable, including provisions for metadata</i>	10
2.2.	<i>Making data openly accessible</i>	10
2.3.	<i>Making data interoperable</i>	10
2.4.	<i>Increase data re-use (through clarifying licenses)</i>	10
3.	<i>Allocation of resources</i>	10
4.	<i>Data security</i>	11
5.	<i>Ethical aspects</i>	11
6.	<i>Other</i>	11
5.	E-learning materials	11
1.	<i>Data summary</i>	11
2.	<i>FAIR Data</i>	11
2.1.	<i>Making data findable, including provisions for metadata</i>	11
2.2.	<i>Making data openly accessible</i>	12
2.3.	<i>Making data interoperable</i>	12
2.4.	<i>Increase data re-use (through clarifying licenses)</i>	12
3.	<i>Allocation of resources</i>	13
4.	<i>Data security</i>	13
5.	<i>Ethical aspects</i>	13
6.	<i>Other</i>	13
6.	RiBaSE Interoperability experiment data	13
1.	<i>Data summary</i>	13
2.	<i>FAIR Data</i>	14
2.1.	<i>Making data findable, including provisions for metadata</i>	14
2.2.	<i>Making data openly accessible</i>	15
2.3.	<i>Making data interoperable</i>	15
2.4.	<i>Increase data re-use (through clarifying licenses)</i>	15
3.	<i>Allocation of resources</i>	15
4.	<i>Data security</i>	15
5.	<i>Ethical aspects</i>	16
6.	<i>Other</i>	16
7.	Tools for the business strategies services	16
1.	<i>Data summary</i>	16
2.	<i>FAIR Data</i>	16
2.1.	<i>Making data findable, including provisions for metadata</i>	16
2.2.	<i>Making data openly accessible</i>	17
2.3.	<i>Making data interoperable</i>	17
2.4.	<i>Increase data re-use (through clarifying licenses)</i>	17
3.	<i>Allocation of resources</i>	17
4.	<i>Data security</i>	17

5. <i>Ethical aspects</i>	17
6. <i>Other</i>	18
<i>[Bibliography]</i>	19
<i>[Acronyms]</i>	20

0. Main updates in the Data Management Plan

This is the expected final version of the WaterInnEU Data Management Plan. The Draft Data Management Plan was delivered on 09/06/2015, at the beginning of the project activities and so it tried to outline the datasets that would be produced within the project and the storing, accessing and documentation policies to be applied to them.

During the lifetime of the project, these datasets have been built up and their characteristics have become clearer than at the beginning of the project. For this reason, a revision of the contents of the initial Data Management Plan is required. The main changes in the datasets or in the policies applied to them are collected in this Final Data Management Plan (DMP) and are summarized below:

- The format of the plan has been adapted to the new version of the European Commission.
- The Dataset 2. Feedback from the end users is now Dataset 2. Survey and interview responses.
- The Dataset 3. Compendium of available standards has disappeared as it is not a dataset but a deliverable.
- The Dataset 4. Tools for the business strategies services is now Dataset 7. Tools for the business strategies services.
- The Dataset 5. Technical and support documentation about the use of the marketplace has disappeared as it is no longer considered a dataset.
- The Dataset 6. Marketplace platform internal code is now Dataset 4. Marketplace platform.
- New Dataset 3. Product Specification Sheets
- New Dataset 5. E-learning materials
- New Dataset 6. RiBaSE Interoperability experiment data

1. Inventory of data sources, companies and solutions for the marketplace

1. Data summary

In the initial steps of the project, a database with an inventory of data sources, companies and solutions to support the development of the project has been created based on the information found in specialised websites.

The data sources and solutions come from European research projects funded by the EU Commission and so have been collected from official portals, primarily those from the EC and from the project websites themselves. The listed companies are either identified from the earlier identified European projects, or by the different WaterInnEU partners in their home markets. The database itself contains the list of portals and sources where the information has been obtained.

This database has been collected in a spreadsheet format, using an Excel file. It is not a big file (250 KB), as the information recorded in it is in text format and the water related EU funded projects and companies are limited. The information included in the dataset is:

- Inventory of data sources and solutions: project identification, countries of the partners, location, aggregation level, type of data, end date, keywords, approach and comments.
- Inventory of companies: code, class, name, country, website, keywords.

The inventory dataset is meant as a starting point for the WaterInnEU Marketplace and is serving as input for selecting products, solutions and companies that are being integrated in the prototype of the Marketplace.

2. FAIR Data

2.1. Making data findable, including provisions for metadata

The metadata standard used to describe the dataset will be the Dublin Core Schema, as it is a flexible and common used standard and is also the one adopted by the European OpenAIRE repository (https://guidelines.openaire.eu/en/latest/literature/use_of_oai_pmh.html). It will in some way include references to the WaterInnEU project (through the grant identification number 641821), to the deliverables D2.1 and D2.2 of the project (also deposited in the repository and linked to the database), and to the EC and H2020 program.

The database will be identified with the name *Dataset 2.1. Inventory of data sources, companies and solutions* where the numbers refer to the origin WP (2) and the ordinal number of datasets within the WP (1). This naming convention is the same used to identify the deliverables in the project.

Some keywords possibly used for the description of the dataset when being deposited are *water tools, water solutions, EU projects, water companies*.

As this dataset is now completed and has been used as a starting inventory for the creation of the marketplace, no updates to it are foreseen during the project or after its ending. For this reason, there's no need for versioning control as no new versions are foreseen.

Currently the related deliverables of the project are being deposited in an institutional repository (UAB Digital Repository of Documents - <http://ddd.uab.cat/?ln=en>), which is linked to the OpenAIRE archive (<http://tinyurl.com/gvfryeb>), but for the moment it doesn't accept datasets. If this issue is not solved once the project is finished, all the datasets produced within the project will be deposited in the [OpenAIRE's Zenodo catch-all repository](#). The institutional repository provides a unique URL to access the documents with the format <https://ddd.uab.cat/record/123456>. If finally the dataset is deposited in the Zenodo repository, a DOI will be provided for the dataset.

2.2. Making data openly accessible

All the data in the database will be made openly available as there is no private information stored in it. The identification of the projects and companies and the information about them have been collected from public and open websites and don't include any restricted data.

Some of the information included in the database (that related to the selected products and companies) will be made available through the marketplace created within the project as it will form the basis for the products and companies description. The complete database is available from the project website (http://www.waterinneu.org/deliverables/Database_v7.xlsx) and will also be available from the institutional repository (<http://ddd.uab.cat/?ln=en>) when it is ready to store data or from the Zenodo repository in case the former is not available.

There is not included documentation about the software needed to access the data, but it consists in an Excel file, which is a common used format and widely distributed software. We trust in the repository itself to look after the documentation about this and the necessary updates of the file to keep it accessible from new versions of the software.

2.3. Making data interoperable

The use of the Dublin Core standard will ensure the interoperability of the data. The data itself is in text format recorded in a spreadsheet file. Where possible, standard codes has been followed, for example in the case of country identification, registered following the ISO 3166-1-alpha-2 codes. Other types of data have been registered following internal codifications, clearly specified within the file.

2.4. Increase data re-use (through clarifying licenses)

The deliverables associated to the dataset are licensed through an *All rights reserved* license as they are working papers not intended to be reused. Nevertheless the database should be shared as a possible reusable dataset. For this reason, when deposited to the repository, an Attribution-NonCommercial license (by-nc) will be requested. The data is currently available for re-use from the project website (http://www.waterinneu.org/deliverables/Database_v7.xlsx) and will also be findable and reusable through the final depositing repository (the institutional one or Zenodo) and from OpenAire, the latest by the end of the project.

The data will remain re-usable after the end of the project by anyone interested in it with no access or time restrictions. As this dataset has been used as a starting inventory for the creation of the marketplace, no updates to it are foreseen during the project or after its ending. The database reflects the inventory of projects and companies at the moment it was made for the WaterInnEU work and it's not expected to be updated to include the new ones. The living database will be the marketplace itself.

3. Allocation of resources

There are no costs associated to the described mechanisms to make the database FAIR and long term preserved. The value for the long term preservation is low, as the database will soon become out of date.

As the coordinator, CREAM has the ultimate responsibility for the data management in the project and so, for the Inventory of data sources, companies and solutions management.

4. Data security

All the responsibilities concerning data recovery and secure storage will go to the repository storing the dataset. Both the institutional (<http://serveis.uab.cat/ddd/>) and the zenodo (<https://zenodo.org/features>) have properly addressed this issue.

5. Ethical aspects

No ethical aspects apply to this dataset.

6. Other

Through the use of the institutional repository, we are also following these procedures for data management:

Institutional level

- [Open Access Itinerary](#) (CREAF-DDD)
- [Open Access Institutional Policy of the Autonomous University of Barcelona](#) (Agreement of the Governing Council on April 25th 2012)

National level

- [Law 14/2011](#) of June 1st, on Science, Technology and Innovation (Article 37 Dissemination in open access)

2. Survey and interview responses

1. Data summary

In the initial steps of the project feedback from identified potential end users of the marketplace has been collected through a survey and interviews. This information is being used to design and adapt the functionalities and services to be offered by the virtual marketplace.

The results obtained from this consultation process (survey and interviews) concern the identification of user needs from actors involved in implementing the EU Water Framework Directive, especially their information needs, preferences and challenges in selecting and applying existing products. The assessment of the information obtained resulted in the formulation of concrete user requirements for functionalities and services from the virtual marketplace. The questions have been addressed mainly to European river basin managers.

The survey has been designed as an online consultation using the Free Open Source Software survey tool “limesurvey”¹. The results have been registered into a database in spreadsheet format. Deliverable *D3.1 Draft requirements report* and deliverable *D3.2 Final requirements report* present the results of this consultation process of potential end users. The survey questionnaires collected information from potential end users on the following aspects: general information/user profile, tools (knowledge, barriers, needs), information platforms (knowledge, barriers, needs), user requirements (services). The treatment of personal information in the surveys is described in the ethics document of WaterInnEU.

The information collected has come directly from potential end users, through the online survey in the first round of consultations and from interviews in a second round. 219 people provided answers to the online survey and 22 people participated in the interviews. The end volume of the file is not very big. There was no available information for this kind of data before the consultation process, as it was about qualitative opinions of the stakeholders identified as potential end users of the virtual marketplace.

2. FAIR Data

2.1. Making data findable, including provisions for metadata

The metadata standard used to describe the dataset will be the Dublin Core Schema, as it is a flexible and common used standard and is also the one adopted by the European OpenAIRE repository (https://guidelines.openaire.eu/en/latest/literature/use_of_oai_pmh.html). It will in some way include references to the WaterInnEU project (through the grant identification number 641821),

¹ <https://www.limesurvey.org/>

to the deliverables D3.1 and D3.2 of the project (also deposited in the repository and linked to the survey results), and to the EC and H2020 program.

The database will be identified with the name *Dataset 3.1. Survey and interview responses* where the numbers refer to the origin WP (3) and the ordinal number of datasets within the WP (1). This naming convention is the same used to identify the deliverables in the project.

Some keywords possibly used for the description of the dataset when being deposited are *water tools, water solutions, user requirements, stakeholders*.

As this dataset is now completed and has been used as a guiding input for the creation of the marketplace, no updates to it are foreseen during the project or after its ending. For this reason, there is no need for versioning control as no new versions are foreseen.

Currently the related deliverables of the project are being deposited in an institutional repository (UAB Digital Repository of Documents - <http://ddd.uab.cat/?ln=en>), which is linked to the OpenAIRE archive (<http://tinyurl.com/gvfryeb>), but for the moment it does not accept datasets. If this issue is not solved once the project is finished, all the datasets produced within the project will be deposited in the [OpenAIRE's Zenodo catch-all repository](#). The institutional repository provides a unique URL to access the documents with the format <https://ddd.uab.cat/record/123456>. If finally the dataset is deposited in the Zenodo repository, a DOI will be provided for the dataset.

2.2. Making data openly accessible

The surveys were anonymous, but some personal information was collected in order to get user profiles and allow for the subsequent categorization of the responses. This personal information included data like the e-mail address and telephone of the respondent, his/her country, river basin and institution, and main field of work. The answers to these data were optional as it was not mandatory to fill them to answer the survey, and respondents were informed that the information collected would remain confidential and would not be transferred to third parties. For this reason, this kind of data will be kept closed and will not be made accessible at all. The information regarding the opinions about tools, platforms and user requirements will be made accessible without being related to any personal data.

The spreadsheet containing the surveys responses will be made available through a data repository after deleting fields containing personal data, either on the institutional repository of the UAB (<http://ddd.uab.cat/?ln=en>) or the Zenodo catch-all repository. Both are linked to the OpenAire archive. This way the dataset will be entirely accessible without restrictions (as confidential data will have been removed).

There is no documentation included about the software needed to access the data, as it consists in an Excel file, which is a commonly used format and widely distributed software. We trust in the repository itself to look after the documentation about this and the necessary updates of the file to keep it accessible from new versions of the software.

2.3. Making data interoperable

There is no suitable standard for this kind of information, as it is not structured information. Users have filled the surveys using their own terms and vocabulary. An effort towards the harmonization of the responses has been done in summarising the results in deliverable *D3.1 Draft requirements report*.

The use of the Dublin Core standard will ensure the interoperability of the data. The data itself is in text format recorded in a spreadsheet file.

2.4. Increase data re-use (through clarifying licenses)

The deliverables associated to the dataset are licensed through an *All rights reserved* license as they are working papers and not intended to be reused. Nevertheless the surveys responses should be shared as a possible reusable dataset. For this reason, when deposited to the repository, an Attribution-NonCommercial license (by-nc) will be requested. The data will be findable and reusable through the final depositing repository (the institutional one or Zenodo) and from OpenAire, the latest by the end of the project.

The data will remain re-usable after the end of the project by anyone interested in it with no access or time restrictions (as the personal information will have been removed). As this dataset has been used as a guiding criteria for the creation of the marketplace, no updates to it are foreseen during the project or after its ending. The responses reflect the opinion of the end users about the existing tools and platforms at the moment the survey was distributed and it is not expected to be updated.

3. Allocation of resources

There are no costs associated to the described mechanisms to make the dataset FAIR and long term preserved. The value for the long term preservation is low, as the responses will soon become out of date.

As the coordinator, CREAM has the ultimate responsibility for the data management in the project and so, for the Surveys responses management.

4. Data security

All the responsibilities concerning data recovery and secure storage will go to the repository storing the dataset. Both the institutional (<http://serveis.uab.cat/ddd/>) and the zenodo (<https://zenodo.org/features>) have properly addressed this issue.

5. Ethical aspects

The ethical aspects related to the personal data collected in this dataset are addressed in the Ethics Requirements version 1.0 document of the project.

6. Other

Through the use of the institutional repository, we are also following these procedures for data management:

Institutional level

- [Open Access Itinerary](#) (CREAF-DDD)
- [Open Access Institutional Policy of the Autonomous University of Barcelona](#) (Agreement of the Governing Council on April 25th 2012)

National level

- [Law 14/2011](#) of June 1st, on Science, Technology and Innovation (Article 37 Dissemination in open access)

3. Product Specification Sheets (PSS)

1. Data summary

To integrate the selected products and solutions in the marketplace and offer a useful description of every one, Product Specification Sheets (PSS) have been prepared for each one. Its purpose is to summarize the main characteristics of the tools and clearly identify the owner, the functionalities and the application and target markets, as well as the stage of commercial development. The PSS are the basis for the presentation of the tools in the marketplace and will be the source of information about them when browsing the platform.

The PSS have been created in a tabular format within a doc file, but they are being translated into a pdf file and also integrated in the online marketplace. The information contained in them comes from the owners of the products and their websites and promotional materials, as well as the websites of the projects that the tools are related to. All the information is in text format, except for the logo of the companies and the identification image of the solution, which are both image files.

Eight PSS have been created that represent approximately 5 MB memory.

2. FAIR Data

2.1. Making data findable, including provisions for metadata

The metadata standard used to describe the dataset will be the Dublin Core Schema, as it is a flexible and common used standard and is also the one adopted by the European OpenAIRE repository (https://guidelines.openaire.eu/en/latest/literature/use_of_oai_pmh.html). It will in some way include references to the WaterInnEU project (through the grant identification number 641821), to the deliverables D6.1 and D6.2 of the project (also deposited in the repository and linked to the database), and to the EC and H2020 program. It will include also references to the tools, the owners, the projects from which the tools come and some keywords describing the purpose of the solutions.

The PSS will be identified with the following naming convention: *NameOfTheTool-PSS_v1.0.pdf*. The suffix “_v1.0” will allow for the control of the versioning. The PSS will all be grouped in a dataset named *Dataset 4.1. Product Specification Sheets* where the numbers refer to the origin WP (4) and the ordinal number of datasets within the WP (1). This naming convention is the same used to identify the deliverables in the project.

Some keywords possibly used for the description of the dataset when being deposited are *water tools*, *water solutions*, and also keywords related to the different applicability of the tools (*socio-economic analysis*, *intercalibration*, etc.).

This dataset is being completed and will be used in the marketplace for the description of the products. When ready, the PSS can be deposited in a repository and later be updated with new versions. As our main repository (UAB Digital Repository of Documents - <http://ddd.uab.cat/?ln=en>) is managing the storing of data for the moment, we will wait until this issue is solved to deposit the data or, in case it is not solved, until the end of the project to deposit the data in an alternative repository. By then no updates are foreseen to the dataset as the project will have finished and the

PSS will have ended their purpose too. In case of new updates to the description of products or of new tools to be added, the marketplace itself will register them and no new PSS will be created.

Currently the related deliverables of the project are being deposited in an institutional repository (UAB Digital Repository of Documents - <http://ddd.uab.cat/?ln=en>), which is linked to the OpenAIRE archive (<http://tinyurl.com/gvfryeb>), but for the moment it doesn't accept datasets. If this issue is not solved once the project is finished, all the datasets produced within the project will be deposited in the [OpenAIRE's Zenodo catch-all repository](#). The institutional repository provides a unique URL to access the documents with the format <https://ddd.uab.cat/record/123456>. If finally the dataset is deposited in the Zenodo repository, a DOI will be provided for the dataset.

2.2. Making data openly accessible

All the data in the PSS will be made openly available as there is no private information stored in it. The identification of the tools and solutions and the descriptions about them have been collected from public and open websites and don't include any restricted data.

The information included in the PSS will form the basis for the products description in the marketplace, so it will be made available through it (<https://wie.dev.52north.org/en/product>). When deposited, they will also be available from the institutional repository (<http://ddd.uab.cat/?ln=en>) when it is ready to store data or from the Zenodo repository in case the former is not available.

There is not included documentation about the software needed to access the data, but it consists in PDF files, which is a common used format and widely distributed software. We trust in the repository itself to look after the documentation about this and the necessary updates of the file to keep it accessible from new versions of the software.

2.3. Making data interoperable

The use of the Dublin Core standard will ensure the interoperability of the data. The data itself is in text format recorded as a PDF file. Controlled vocabulary has been used in all the fields where it was appropriate in order to let for harmonization, but no standards or existing vocabularies have been followed, as the data is too specific.

2.4. Increase data re-use (through clarifying licenses)

The PSS will be shared as a possible reusable dataset, the latest by the end of the project. For this reason, when deposited to the repository, an Attribution-NonCommercial-NoDerivatives license (by-nc-nd) will be requested. The data will also be accessible from the marketplace with the same license since the final version of the platform is being released.

The data will remain re-usable after the end of the project by anyone interested in it with no access or time restrictions. As this dataset has been used as starting material for the marketplace content, no updates to it are foreseen during the project or after its ending. The PSS reflect the status of the tools and solutions at the moment there were made for the WaterInnEU work and they are not expected to be updated to include new information. The living PSS will be the marketplace itself.

3. Allocation of resources

There are no costs associated to the described mechanisms to make the PSS FAIR and long term preserved. The value for the long term preservation is medium, as the sheets can become out of date in a few years.

As the coordinator, CREAM has the ultimate responsibility for the data management in the project and so, for the Inventory of data sources, companies and solutions management.

4. Data security

All the responsibilities concerning data recovery and secure storage will go to the repository storing the dataset. Both the institutional (<http://serveis.uab.cat/ddd/>) and the zenodo (<https://zenodo.org/features>) have properly addressed this issue.

5. Ethical aspects

No ethical aspects apply to this dataset.

6. Other

Through the use of the institutional repository, we are also following these procedures for data management:

Institutional level

- [Open Access Itinerary](#) (CREAF-DDD)
- [Open Access Institutional Policy of the Autonomous University of Barcelona](#) (Agreement of the Governing Council on April 25th 2012)

National level

- [Law 14/2011](#) of June 1st, on Science, Technology and Innovation (Article 37 Dissemination in open access)

4. Marketplace platform

1. Data summary

The marketplace platform is the core of the WaterInnEU project. It will be a web based platform that will connect the outcomes developed in previous EU funded activities (currently collected by the EIP Water marketplace or the WISE-RTD portal) with the already existing data available at European level (INSPIRE directive, Eurostat Open Data Portal, etc.) and also to the SMEs and companies that are able to deploy products and offer services to the users from these tools and data. The platform will also offer an independent marketplace supported by technical and commercial expertise as a service for users (river basin district managers) that will allow them to access products and services best fitting their priorities, capabilities and procurement processes (based on the previous products).

To develop these features, the virtual platform will include some technical components: a companies & SMEs pool, a water data inventory, interoperability experiments, a standardization label, a user feedback component, an e-learning platform and specialist commercialisation support services. Other minor developments will also be included as multimodal search, registration of new products or a spatial component.

Similar initiatives exist like the EIP Water marketplace or the WISE-RTD portal, but the reuse of these kinds of software for online platforms has revealed to be more difficult than expected, so the

code for the marketplace is newly generated within the project. As the platform is currently being developed, the final size of the internal code is not known by now.

2. FAIR Data

2.1. Making data findable, including provisions for metadata

The metadata standard used to describe the dataset will be the Dublin Core Schema, as it is a flexible and common used standard and is also the one adopted by the European OpenAIRE repository (https://guidelines.openaire.eu/en/latest/literature/use_of_oai_pmh.html). It will in some way include references to the WaterInnEU project (through the grant identification number 641821), to the deliverables D6.1, D6.2 and D6.3 of the project (also deposited in the repository and linked to the platform code), and to the EC and H2020 program. Some keywords possibly used for the description of the dataset when being deposited are *online platform*, *marketplace*, *water tools*, *water solutions*, *EU projects*, *water companies*.

The identification and the control for the versioning of the code will be done through the repository itself. The code by now is deposited in the GitHub repository, which allows for the properly control of versions and updates as well as for the identification of the dataset.

2.2. Making data openly accessible

All the data needed to create and maintain the marketplace is being made openly accessible through the GitHub repository (<https://github.com/52North/waterinneu>), along with the corresponding technical documentation, with no restriction access and with indications on the technical requirements needed to properly understand the data.

2.3. Making data interoperable

For the sake of (current and future) maintainability and extensibility, open source solutions are being used. The platform provides open APIs for user authentication (e.g. OAuth) as well as content (e.g. linked open data, or GeoViQua feedback model) to allow integration with other components.

2.4. Increase data re-use (through clarifying licenses)

The development of the platform is being accompanied by proper technical documentation to allow third parties to use parts of the prototype or to mirror the complete system.

The code has been made accessible from the beginning of its development and will remain accessible after the end of the project. When depositing the code to the GitHub repository, a GPL v2.0 license was adopted. The quality of the dataset is guaranteed by the platform functioning.

3. Allocation of resources

There are no costs associated to the described mechanisms to make the marketplace code FAIR. Regarding the long term preservation, in this case it is crucial as the marketplace is the core of the project and needs to continue functioning after the end of the project. For this reason, a parallel preservation strategy needs to be adopted to ensure the sustainability of the platform. So, in the final phase of the project, concepts on how the virtual platform can be sustained beyond the project

duration will be developed with respect to several relevant areas: (i) content (required user groups and levels, moderation needs, guidelines for contributors, linking to other relevant websites), (ii) maintenance and finance (estimation of costs, potential sources to finance a virtual platform), (iii) long term strategy (having demonstrated the features of a marketplace, what use cases are most likely to continue to matter), and (iv) organisational structures (what forms of management are viable). Recommendations for each of these areas will be formulated and published.

As the coordinator, CREAM has the ultimate responsibility for the data management in the project and so, for the Marketplace platform management. Despite this, 52N will help in the responsibility of the marketplace management as the developer of the code.

4. Data security

The responsibilities concerning data recovery and secure storage will be shared among the repository storing the dataset (GitHub) and the partners involved in the marketplace management (CREAM and 52N). GitHub will host the final version of the code and it is properly addressing this issue (<https://help.github.com/articles/github-security/>), but copies of it need also to be preserved and maintained by CREAM and 52 N to guarantee the sustainability of the platform.

5. Ethical aspects

The ethical aspects related to the personal data collected in this dataset are addressed in the Ethics Requirements version 1.0 document of the project.

6. Other

-

5. E-learning materials

1. Data summary

One of the added values of the WaterInnEU marketplace are the e-learning materials about the tools offered in the platform. WP7, concerning the e-learning processes, includes the generation of participative e-learning courses and multimedia products, implemented in an e-learning platform. The multimedia products consist of a set of visually appealing presentations, videos and images that will be used to enrich the content of the products offered in the marketplace.

The training materials are being developed within the WP7 and no existing data is being reused. By now the size of the dataset is about 500 MB but it is expected to grow as it is foreseen to include more products.

2. FAIR Data

2.1. Making data findable, including provisions for metadata

The metadata standard used to describe the dataset will be the Dublin Core Schema, as it is a flexible and common used standard and is also the one adopted by the European OpenAIRE

repository (https://guidelines.openaire.eu/en/latest/literature/use_of_oai_pmh.html). It will in some way include references to the WaterInnEU project (through the grant identification number 641821), to the deliverables D7.1 and D7.2 of the project (also deposited in the repository and linked to the database), and to the EC and H2020 program, as well as to the related products and tools.

The materials will be identified with the name *Dataset 7.1. E-learning materials* where the numbers refer to the origin WP (7) and the ordinal number of datasets within the WP (1). This naming convention is the same used to identify the deliverables in the project.

Some keywords possibly used for the description of the dataset when being deposited are *e-learning, tutorials, multimedia, courses, water tools, water solutions*.

The materials will be identified with the following naming convention: *NameOfTheTool- TypeOfMaterial_v1.0*. The suffix “_v1.0” will allow for the control of the versioning.

Currently the related deliverables of the project are being deposited in an institutional repository (UAB Digital Repository of Documents - <http://ddd.uab.cat/?ln=en>), which is linked to the OpenAIRE archive (<http://tinyurl.com/gvfryeb>), but for the moment it doesn't accept datasets. If this issue is not solved once the project is finished, all the datasets produced within the project will be deposited in the [OpenAIRE's Zenodo catch-all repository](#). The institutional repository provides a unique URL to access the documents with the format <https://ddd.uab.cat/record/123456>. If finally the dataset is deposited in the Zenodo repository, a DOI will be provided for the dataset.

2.2. Making data openly accessible

All the data in the training materials will be made openly available as there is no private information stored in it. The materials will be available through the marketplace and also through the project website (http://www.waterinneu.org/deliverables/E_learning_WaterInnEU.zip) and will also be available from the institutional repository (<http://ddd.uab.cat/?ln=en>) when it is ready to store data or from the Zenodo repository in case the former is not available.

There is not included documentation about the software needed to access the data, but all the materials are in common used formats (PDF, MP4, PPT) and widely distributed types of software. We trust in the repository itself to look after the documentation about this and the necessary updates of the files to keep them accessible from new versions of the software.

2.3. Making data interoperable

The use of the Dublin Core standard will ensure the interoperability of the data. In the development of the graphical materials such as presentations, videos and images, the recognised multimedia standards are being adopted, as for example the JPEG standard for the images and the MPEG standards for the videos. In the case of the tutorials or the technical documentation, the standard HTML is being used to let the maximum interoperability of the data.

2.4. Increase data re-use (through clarifying licenses)

The deliverables associated to the dataset are licensed through an *All rights reserved* license as they are working papers not intended to be reused. Nevertheless the training materials should be shared as possible reusable datasets. For this reason, when deposited to the repository, an Attribution-NonCommercial-NonDerivatives license (by-nc-nd) will be requested. The data is

currently available for re-use from the project website (http://www.waterinneu.org/deliverables/E_learning_WaterInnEU.zip) and will also be findable and reusable through the final depositing repository (the institutional one or Zenodo) and from OpenAire, the latest by the end of the project.

The data will remain re-usable after the end of the project by anyone interested in it with no access or time restrictions.

3. Allocation of resources

There are no costs associated to the described mechanisms to make the database FAIR and long term preserved. The value for the long term preservation is medium, as the materials are a good support material for the promotion of the products of the marketplace and a distinctive plus of the WaterInnEU marketplace.

As the coordinator, CREAM has the ultimate responsibility for the data management in the project and so, for the E-learning materials management.

4. Data security

All the responsibilities concerning data recovery and secure storage will go to the repository storing the dataset. Both the institutional (<http://serveis.uab.cat/ddd/>) and the zenodo (<https://zenodo.org/features>) have properly addressed this issue.

5. Ethical aspects

No ethical aspects apply to this dataset.

6. Other

Through the use of the institutional repository, we are also following these procedures for data management:

Institutional level

- [Open Access Itinerary](#) (CREAF-DDD)
- [Open Access Institutional Policy of the Autonomous University of Barcelona](#) (Agreement of the Governing Council on April 25th 2012)

National level

- [Law 14/2011](#) of June 1st, on Science, Technology and Innovation (Article 37 Dissemination in open access)

6. RiBaSE Interoperability experiment data

1. Data summary

As part of the standardization and interoperability activities in the project, the River Basin Standards Interoperability Experiment (RiBaSE) is designed to demonstrate how current ICT-based tools and water data can work in combination with geospatial web services in the Scheldt, Maritsa and Severn river basins. The results of the pilot will be documented in an engineering report and may result in change requests to OGC standards as appropriate. The experiment is designed to comply with one of the main goals of the project: to assess the level of standardization and interoperability of the EU funded ICT outcomes related to water.

The data collected (re-used) for the experiment includes: height/flow data from past flood events in the river basins (WaterML 2.0) with their corresponding geographic coordinates of the monitoring points, detailed Digital Elevation Models (raster files), catchments limits and land use maps of the areas (raster files). The data generated in the experiment include documentation (engineering report, change requests) and a video illustrating the results of the pilot.

The height/flow data for the experiment is being obtained from different web services of each river basin that register and offer this kind of data through monitoring stations. The Digital Elevation Models have been obtained from the NASA (ASTER Global Digital Elevation Map) and the Land Use maps and catchments limits from the EEA (European Environment Agency), all of them freely available through their websites.

The collected data for the three river basins is about 5 GB volume, but the overall size of the experiment is expected to be quite bigger as multiple raster files need to be developed.

2. FAIR Data

2.1. Making data findable, including provisions for metadata

The metadata standard used to describe the dataset will be the Dublin Core Schema, as it is a flexible and common used standard and is also the one adopted by the European OpenAIRE repository (https://guidelines.openaire.eu/en/latest/literature/use_of_oai_pmh.html). It will in some way include references to the WaterInnEU project (through the grant identification number 641821), to the deliverables D5.2, D5.3 and D5.4 of the project (also deposited in the repository and linked to the database), and to the EC and H2020 program. It will also refer to the RiBaSE experiment, to the three river basins and to the OGC.

The dataset will be identified with the name *Dataset 5.1. RiBaSE Interoperability Experiment data* where the numbers refer to the origin WP (5) and the ordinal number of datasets within the WP (1). This naming convention is the same used to identify the deliverables in the project.

Some keywords possibly used for the description of the dataset when being deposited are *standardisation, OGC, interoperability experiment, water standards, water data, WML*.

Once the experiment is finished no updates to it are foreseen during the project or after its ending. For this reason, there's no need for versioning control as no new versions are foreseen.

Currently the related deliverables of the project are being deposited in an institutional repository (UAB Digital Repository of Documents - <http://ddd.uab.cat/?ln=en>), which is linked to the OpenAIRE archive (<http://tinyurl.com/gvfryeb>), but for the moment it doesn't accept datasets. If this issue is not solved once the project is finished, all the datasets produced within the project will be deposited in the [OpenAIRE's Zenodo catch-all repository](#). The institutional repository provides a unique URL to access the documents with the format <https://ddd.uab.cat/record/123456>. If finally the dataset is deposited in the Zenodo repository, a DOI will be provided for the dataset.

2.2. Making data openly accessible

All the data in the pilot experiment will be made openly available as there is no private information stored in it. The data has been obtained from public institutions and open websites and don't include any restricted data.

The complete dataset will be available from the project website and probably also from the OGC website (<http://www.ogcnetwork.net/ribase>), as well as from the institutional repository (<http://ddd.uab.cat/?ln=en>) when it is ready to store data or from the Zenodo repository in case the former is not available.

The engineering report and the deliverables about the experiment will detail the types of data and the software needed to replicate the experiment. This software will also be included in the dataset as some parts of it are being developed within the experiment.

2.3. Making data interoperable

The use of the Dublin Core standard will ensure the interoperability of the data. In the case of the content of the dataset, it will be surely interoperable as all the formats and types of data managed in the experiment are standards being interoperable: SOS (OGC SOS 2.0 Hydrology Profile), WPS (OGC WPS 2.0), WaterML 2.0.

2.4. Increase data re-use (through clarifying licenses)

The deliverables associated to the dataset are licensed through an *All rights reserved* license as they are working papers not intended to be reused. Nevertheless the dataset should be shared as a possible reusable dataset. For this reason, when deposited to the repository, an Attribution-NonCommercial license (by-nc) will be requested. The data will be available for re-use from the project website and will also be findable and reusable through the final depositing repository (the institutional one or Zenodo) and from OpenAire, the latest by the end of the project.

The data will remain re-usable after the end of the project by anyone interested in it with no access or time restrictions. Once the experiment is finished no updates to it are foreseen during the project or after its ending.

3. Allocation of resources

There are no costs associated to the described mechanisms to make the dataset FAIR and long term preserved. The value for the long term preservation is medium, as the experiment can be repeated through the available data.

As the coordinator, CREAM has the ultimate responsibility for the data management in the project and so, for the RiBaSE Interoperability Experiment data management.

4. Data security

All the responsibilities concerning data recovery and secure storage will go to the repository storing the dataset. Both the institutional (<http://serveis.uab.cat/ddd/>) and the zenodo (<https://zenodo.org/features>) have properly addressed this issue.

5. Ethical aspects

No ethical aspects apply to this dataset.

6. Other

Through the use of the institutional repository, we are also following these procedures for data management:

Institutional level

- [Open Access Itinerary](#) (CREAF-DDD)
- [Open Access Institutional Policy of the Autonomous University of Barcelona](#) (Agreement of the Governing Council on April 25th 2012)

National level

- [Law 14/2011](#) of June 1st, on Science, Technology and Innovation (Article 37 Dissemination in open access)

7. Tools for the business strategies services

1. Data summary

This dataset will include the materials and procedures developed within WP4 to provide support services to the companies interested in implementing the solutions of the marketplace. In WP4, these companies will be guided by experienced innovation brokers towards successful application and marketing of the identified innovation solutions. This support will involve the development of roadmaps and business strategies for every case, as well as the writing of short procedures and checklists to be followed during these processes.

The data created will therefore be specific for each company and will include a set of documents to guide the development and market implementation of the solutions. The data could be useful for other commercial strategies, but only in general terms, as the roadmaps will be designed specifically for each solution.

Adelphi, which is the main responsible partner for this task, has a broad experience in developing marketing strategies, so they will know and surely integrate and reuse some similar data, depending on the solutions to be supported.

2. FAIR Data

2.1. Making data findable, including provisions for metadata

The business strategies are likely to be defined by a SWOT analysis, outlining the weaknesses, the opportunities, the threats and the strengths of the initiative, and by an analysis of the macro-environmental conditions, the competitive edge and the sales strategy, the marketing strategy (pricing and promotion), and the strategic alliances.

The tools will be identified with the following naming convention: *NameOfTheCompany-ToolType_v1.0*. The suffix “_v1.0” will allow for the control of the versioning.

The dataset will be identified with the name *Dataset 4.1. 8. Tools for the business strategies services* where the numbers refer to the origin WP (4) and the ordinal number of datasets within the WP (1). This naming convention is the same used to identify the deliverables in the project.

2.2. Making data openly accessible

This information contains the strategic plans for the companies in order to bring a product or a solution to the market. These support tools and data generated in the development of the companies' market strategies constitute private commercial information concerning only the companies involved and their business, so they will not be shared and will remain in the hands of the companies.

2.3. Making data interoperable

Although there are a lot of procedures to develop business strategies, some of them following similar rules, there is not a standardized way to do them. During the process of development of the business support tools, deep research will be conducted to find the most applied and interoperable method to generate such strategies.

2.4. Increase data re-use (through clarifying licenses)

The deliverables associated to the dataset are licensed through an *All rights reserved* license as they are working papers not intended to be reused. Nevertheless the dataset won't be available for re-use due to its commercial interest.

The data generated in the process of provision of support services to the interested parties will be useful during the process of developing the market strategy and bringing the product to the market. After this period, the data will be transferred to the corresponding companies, which will decide what to do with the data. Given that the information generated will not be of public interest, the preservation of the dataset will fall to the companies themselves.

3. Allocation of resources

There are no costs associated to the described mechanisms to make the dataset FAIR and long term preserved. The value for the long term preservation will depend on each company.

As the coordinator, CREAM has the ultimate responsibility for the data management in the project and so, for the Tools for the business strategies services management.

4. Data security

All the responsibilities concerning data recovery and secure storage will go to the WP4 leader (adelphi) and to the coordinator (CREAF). They should take care of the materials during the project development providing the needed storage facilities and access controls to guarantee the data integrity.

5. Ethical aspects

No ethical aspects apply to this dataset.

6. Other

-

[Bibliography]

Deliverable D1.2 Draft Data Management Plan.

http://www.waterinneu.org/deliverables/641821_Draft%20Data%20Management%20Plan.pdf

Deliverable D2.1 Data availability report.

http://www.waterinneu.org/deliverables/641821_Data%20availability%20report.pdf

Deliverable D2.2 Water companies and solutions report.

http://www.waterinneu.org/deliverables/641821_Water%20companies%20and%20solutions%20report.pdf

Deliverable D3.1 Draft guidance report.

http://www.waterinneu.org/deliverables/641821_Water%20companies%20and%20solutions%20report.pdf

Deliverable D5.2 European water interoperability experiment request for participation.

http://ddd.uab.cat/pub/estudis/2015/145183/641821_D5.2_-_European_water_interoperability_experiment_request_for_participation.pdf

Deliverable 6.3 Marketplace platform.

http://ddd.uab.cat/pub/worpap/2016/148388/641821_D6.3_-_Market_Place_platform.pdf

Deliverable 7.1 Draft tutorials and multimedia products.

http://www.waterinneu.org/deliverables/641821_D7.1%20-%20Draft%20tutorials%20and%20multimedia%20products.pdf

UAB Digital Repository of Documents. <https://ddd.uab.cat/?ln=en>

OpenAIRE's Zenodo catch-all repository. <https://zenodo.org/>

Creative Commons website. <https://creativecommons.org/>

River Basin Standards Interoperability Pilot (RiBASE) website.

<http://www.ogcnetwork.net/ribase>

[Acronyms]

DMP – Data Management Plan

EU – European Union

EC – European Commission

WP – Work Package

UAB – Autonomous University of Barcelona

DOI – Digital Object Identifier

ISO – International Organization for Standardization

FAIR – Findable, Accessible, Interoperable and Re-usable

CREAF – Ecological and Forestry Applications Research Centre

DDD – Digital Repository of Documents

PSS – Product Specification Sheet

OGC – Open Geospatial Consortium

PDF – Portable Document Format

MP4 – Moving Picture Experts Group 4

PPT - Microsoft PowerPoint format

JPEG - Joint Photographic Experts Group

MPEG - Moving Picture Experts Group

HTML - Hyper Text Markup Language

GPL – General Public License

EIP – European Innovation Partnerships

WISE-RTD - Water Information System for Europe

SME – Small and medium-sized enterprise

RiBaSE – River Basin Standards Interoperability Experiment

SWOT - strengths, weaknesses, opportunities, and threats

ICT - Information and communications technology

NASA - National Aeronautics and Space Administration

ASTER - Advanced Spaceborne Thermal Emission and Reflection Radiometer

EEA – European Environment Agency

WML - Wireless Markup Language

SOS - Sensor Observation Service

WPS – Web Processing Service