# Deliverable D5.5 Report on the remaining four DMP processes

| | | | |
|---|---|---|---|
| **Project Title (Grant agreement no.):** | ELIXIR-CONVERGE: Connect and align ELIXIR Nodes to deliver sustainable FAIR life-science data management services (871075) | | |
| **Project Acronym (EC Call):** | ELIXIR-CONVERGE (H2020-INFRADEV-2018-2020) | | |
| **WP No & Title:** | WP5 Demonstrator Projects | | |
| **WP leader(s):** | Anne-Françoise Adam-Blondon | | |
| **Deliverable Lead Beneficiary:** | 10 - BSC | | |
| **Contractual delivery date:** | 30/06/2023 | **Actual delivery date:** | 20/06/2023 |
| **Delayed:** | No | | |
| **Partner(s) contributing to this deliverable:** | | | |
| **Authors:** Laura Portell-Silva (BSC), Vanita Haurheeram (INRAE), Anne-Françoise Adam-Blondon (INRAE)<br>**Contributors:** Salvador Capella-Gutierrez (BSC), Marina Popleteeva (UNILU), Federico Bianchini (UiO), Espen Åberg (UiB/UiT), Rob Hooft (DTL), Marek Schuánek (CTU), Jan Slifka (CTU), Adam Hospital (IRB Barcelona), Nils P Willassen (UiB/UiT), Janet Piñero (UPF), Ferran Sanz (UPF), Juan Manuel Ramirez-Anguita (UPF), Manuel Pastor (UPF), Miguel Angel Mayer (UPF), Ernesto Picardi (CNR), Pinar Alper (UNILU), Vilém Ded (UNILU), Nene Djenaba Barry (UNILU), Paulette Lieby (IFB), Teresa d'Altri (CRG), Daniel Faria (INESC-ID), Erwan Le Floch (INRAE), Philippe Rocca-Serra (EMBL-EBI), Bert Droesbeke (VIB), Korbinian Bösl (UiB), Marko Vidak (UL)<br>**Acknowledgments (not grant participants):** Cyril Pommier (EMPHASIS, ELIXIR Plant Science community, INRAE), Sebastian Beier (ELIXIR Plant Sciences community, IPK), Matthias Lange (ELIXIR Plant Sciences community, IPK), Daniel Arend (ELIXIR Plant Sciences community, IPK), Nadja Zlender (IBMI); Isabelle Alic (EMPHASIS) | | | |
| **Reviewers:** | ELIXIR-CONVERGE Management Board (MB) members. | | |

## Log of changes

| DATE | Mvm | Who | Description |
|---|---|---|---|
| 17/05/2023 | 0v1 | Laura Portell-Silva (BSC) | Initial version |
| 20/06/2023 | 0v2 | A-F Adam-Blondon (INRAE)<br>Laura Portell-Silva (BSC) | Sent to PMU after incorporating internal WP feedback |
| 05/07/2023 | 0v3 | Nikki Coutts (ELIXIR Hub) | Circulated to the MB for final review before submission |
| 13/07/2023 | 0v4 | A-F Adam-Blondon (INRAE)<br>Laura Portell-Silva (BSC) | MB comments addressed |
| 13/07/2023 | 1v0 | Nikki Coutts (ELIXIR Hub) | Final version to be uploaded into EC Portal |

# Table of contents

This project has received funding from the European Union’s Horizon 2020 Research and Innovation programme under grant agreement No **871075**.

# 1. Executive Summary

The main objective of this deliverable is to provide resources that can assist life sciences researchers and data stewards in creating reference Data Management Plans (DMPs) for their research projects. The resources provided in this deliverable are intended to promote good research data management practices across the EU research landscape.

To achieve this objective, a range of activities were undertaken, with a particular focus on the needs of the different domains covered by the demonstrator use-cases. In order to create these resources, the RDMkit pages were extended to include domain-specific information, which can be used as a reference when developing DMPs for different research projects. These pages fall under the "Your Domain" category and provide specific information on the data management needs and considerations for each domain. They also highlight challenges that are specific to each domain, such as data types, species, or areas, and offer solutions and considerations to overcome these challenges. In this deliverable, the RDMkit page for the Toxicology data demonstrator use-case was completed and added to the existing RDMkit pages for the other demonstrator use-cases. Additionally, a new Tool Assembly was added to the RDMkit corresponding to the Plant Sciences demonstrator use-case, covering the entire life cycle of experimental plant phenotyping data.

In addition, the general Knowledge Models (KMs) of the Data Stewardship Wizard (DSW) were adapted to address the specific DMP questions needed for each demonstrator use-case. The DSW is a collaborative tool that enables data stewards and researchers to efficiently create DMPs for their research projects and it is designed with a hierarchical KM that guides users through the creation of DMPs. Since the relevant information for DMPs can vary across different domains, these KMs can be modified to contain the information relevant for each demonstrator use-case. For this deliverable, special focus was put on two of the demonstrator use-cases, namely Toxicology and Epitranscriptomics data.  Additionally, related to the Human Data use-case, separate efforts are underway to enhance the sensitive data section of the KM system to ensure the proper management of such data. The improvements and new question suggestions that were found during these sessions were incorporated into the DSW KM by the DSW team.

Furthermore, DMP templates were created in DSW for the demonstrator use-case using two standard approaches: creating a KM or a project template (PT). When creating a PT, a set of answers is saved and can be used to generate a partially pre-filled questionnaire for a new project. In the ideal case scenario, the two methods can be used together to provide domain-specific recommendations by answering questions that better reflect a scientific domain, such as metadata standards.

In conclusion, this deliverable provides several valuable resources for life sciences researchers and data stewards, including extended RDMkit pages, customised DSW KMs, domain-specific DMP templates, and a new KM for creating DPIAs. These resources are designed to encourage good research data management practices across the EU research landscape, ensuring that valuable research data is effectively managed before, during, and after a project.

# 2. Contribution toward project objectives

With this deliverable, the project has reached or the deliverable has contributed to the following objectives/key results:

| **Objective no. / Key Result no. Description** | **Contributed to:** |
|---|---|
| **Objective 1:** Develop a sustainable and scalable operating model for transnational life-science data management support by leveraging national capabilities **(WP1, WP5**) | |
| **Key Result 1.1:** Established European expert network of data stewards that connect national data centres and similar infrastructures and drive the development of interoperable solutions following international best practice, including national interpretations of the General Data Protection Regulation (GDPR) | **No** |
| **Key Result 1. 2:** Development of joint guidelines and common toolkit that are adopted into funder recommendations, with support available nationally and in local languages | **Yes** |
| **Key Result 1.3:** The catalogue of successful national business models incorporated into national strategies | **No** |
| **Key Result 1.4:** The developed “sustainable and scalable operating model for transnational life-science data management support” is adopted into national ELIXIR Node | **No** |
| **Objective 2:** Strengthen Europe’s data management capacity through a comprehensive training programme delivered throughout the European Research Area **(WP2, WP6)** | |
| **Key Result 2.1:** A comprehensive ELIXIR Training and Capacity building programme in Data Management, directed at both data managers and ELIXIR users, and connected to the national training programmes in Data Management in the ELIXIR Nodes and prospective ELIXIR Member countries. | **No** |
| **Key Result 2.2:** Development of a collective group of trainers that support scalable deployment of Data Management training across ELIXIR Nodes. | **No** |
| **Key Result 2.3:** A substantial cohort of data managers, Node coordinators and researchers with specific data management skills, business planning and knowledge of transnational operations across the ELIXIR Nodes | **No** |
| **Objective 3:** Align national data management standards and services through a sustainable, scalable and cost-effective data management toolkit **(WP2, WP3, WP5**) | |
| **Key Result 3.1:** Assemble a full-stack harmonised common toolkit comprising | **Yes** |

| | |
|---|---|
| all aspects of data management: from data capture, annotation, and sharing; to integration with analysis platforms and making the data publicly available according to international standards. | |
| **Key Result 3.2:** Provide exemplar toolkit configurations for prioritised demonstrators to serve as templates for future use. | **Yes** |
| **Key Result 3.3:** Establish national capacity in using as well as updating, extending and sustaining the toolkit across the ERA. | **Yes** |
| **Key Result 3.4:** Enable 'FAIR at source' practice for data generation, and analytical process pipeline implementation by flexible deployment of the toolkit in national operations | **No** |
| **Objective 4:** Align national investments to drive local impact and global influence of ELIXIR **(WP4,WP6**) | |
| **Key Result 4.1:** Development of a Node Impact Assessment Toolkit based on RI-PATHS methodology. | **No** |
| **Key Result 4.2:** Adoption of Impact assessment in ELIXIR Nodes, supported by Node coordinators network and feedback on applicability from dialogues with national funders. | **No** |
| **Key Result 4.3:** Creation of national public-private partnerships and industry outreach where open life-science data and services stimulate local bioeconomy | **No** |
| **Key Result 4.4:** Growth in reach, impact and engagement of stakeholder communication assessed by established ELIXIR Communications metrics | **No** |
| **Key Result 4.5:** Initiating and advancing discussions on Membership (EU and international) or strategic partnerships (international countries) following ELIXIR-CONVERGE workshops. | **No** |

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No **871075**.

# 3. Introduction

Effective data management plans (DMPs) are essential for successful research data management, outlining the management of research data throughout the project lifecycle. ELIXIR-CONVERGE WP5 aims to evaluate ELIXIR's and its national nodes' capacity to support users' projects in implementing DMPs on a European scale. This objective is being achieved through the development of six diverse demonstrator use-cases:

1. Harmonised FAIR plant genotype & phenotype data management toolkit for Europe
2. Reproducible, comparable and FAIR Epitranscriptomics
3. Common Data management plans for the marine metagenomics Community
4. Federated access to human genomics data: GDPR
5. FAIR encoding and access to Toxicology data
6. FAIR organisation of biomolecular simulation information

In Deliverable D5.1[1], the demonstrator use-cases were analysed, and a categorization based on users' needs was proposed to determine the process necessary for DMP implementation. Deliverable D5.2[2] builds on this by providing a detailed description of the primary resources available for researchers to create DMPs for the demonstrator use-cases. It outlines the main activities related to the ELIXIR Research Data Management Kit (RDMkit) and the Data Stewardship Wizard (DSW), which were used to develop the initial description of the first DMPs. Additionally, "Your domain" and "Tool assembly" pages were added to the RDMkit, providing information on data management in the demonstrator use-cases domains. D5.2 started the work to create customised KMs, and specific sessions with DSW experts and representatives of the different demonstrator use-cases were held, where the general KM for DMPs was analysed to detect areas that can be improved for four of the demonstrator use-cases (Plant Sciences, Marine Metagenomics, Human Data and Biomolecular Simulation).

## 3.1 Scope of Deliverable

This deliverable builds upon the progress made in D5.2 and provides a more comprehensive overview of the work conducted on the different demonstrator use-cases. The RDMkit and DSW continue to be central to the activities outlined in this deliverable, as they are two primary outcomes of ELIXIR-CONVERGE and valuable resources for creating customised DMPs for the various demonstrator use-cases.

- For the Toxicology data use-case, a dedicated "Your domain" page was added to the RDMkit, providing information specific to data management within this domain.
- For the Plant sciences use-cases, a dedicated “Tool assembly” page was added to the RDMkit, providing data management guidelines that cover the whole life cycle of experimental plant phenotyping data.

[1] https://doi.org/10.5281/zenodo.4674490
[2] https://doi.org/10.5281/zenodo.5139903

- Dedicated sessions were held between demonstrator use-cases representatives and the DSW team to adapt the Knowledge Models (KMs) to the specific DMP questions needed for each demonstrator use-case.
- To further streamline the process of DMP creation, project templates were integrated into the DSW for selected demonstrator use-cases. These templates pre-fill answers to common questions, simplifying the DMP creation process for users in these domains who wish to use the DSW.
- Related to demonstrator use-case #4, (Federated access to human genomics data: GDPR) the DSW now offers a convenient feature for creating a Data Protection Impact Assessment (DPIA) directly within the platform[3]. This has been made possible through the addition of a new KM and a corresponding document output template. Users are now able to create a DPIA project within DSW, input relevant information, and generate a fully-formed DPIA document.

## 3.2 Relationship with other WPs

- ELIXIR-CONVERGE WP1 Data Management experts contributed in the best practices and guidelines in the 'Your domain' pages in the RDMkit for some of the demonstrator use-cases added.
- ELIXIR-CONVERGE WP2 members contributed in the creation of dedicated training materials and capacity building actions.
- ELIXIR-CONVERGE WP3 RDMkit members contributed to the assembly of the 'Your domain' and 'Tool assembly' pages in the RDMkit.
- ELIXIR-CONVERGE WP4 members contributed in providing reference information to the external members to the project, e.g. Industry.
- ELIXIR-CONVERGE WP7 Federated EGA members contributed to the activities related to the human data use-case.

## 3.3 Methodology

The approach for this deliverable was established in collaboration with representatives from the various demonstrator use-cases, primarily through regular monthly meetings and additional discussions with pertinent ELIXIR communities (such as Plant Science), as well as DSW experts and ELIXIR-CONVERGE WP3 partners. The work done for this deliverable was done through different approaches:

1. In order to ensure that the KMs within DSW were tailored to the specific needs of each demonstrator use-case, dedicated sessions with representatives from each demonstrator use-case and the DSW team were conducted. These sessions were designed to facilitate an open exchange of information, allowing for a better understanding of the unique requirements of each demonstrator use-case.

[3] https://doi.org/10.37044/osf.io/cuvqw

2. During the RDMkit contentathons organised by WP3, the opportunity was taken to complete the "Your domain" and “Tool assembly” pages and to complement and expand upon other pages related to the demonstrator use-cases.
3. A dedicated project[4] in the BioHackathon Europe 2021 was used to add a new KM in DSW, allowing the creation of a DPIA directly in DSW.

# 4. Description of work accomplished

The main objective of this deliverable is to provide resources that can assist life sciences researchers and data stewards in creating reference DMPs for their research projects. To achieve this, a range of activities were undertaken, with a particular focus on the needs of the different domains covered by the demonstrator use-cases.

In order to create these resources, the RDMkit pages were extended to include domain-specific information, which can be used as a reference when developing DMPs for different research projects. In addition, the general KMs of the DSW were modified to address the particular DMP questions necessary for each demonstrator use-case, and project templates were included in the DSW for certain demonstrator use-cases, streamlining the DMP creation procedure for users in these fields who intend to use the DSW.

These resources will be of great value to life sciences researchers and data stewards, providing them with practical guidance on how to effectively manage research data before, during, and after a project. By making reference DMPs available for the different domains covered by the demonstrator use-cases, we hope to encourage good research data management practices across the EU research landscape.

## 4.1 New RDMkit pages for the demonstrator use-cases

The previous deliverable, D5.2, included RDMkit pages for all demonstrator use-cases except for Toxicology data. In this deliverable, the RDMkit page for the Toxicology data demonstrator use-case has been completed and is now available.[5] These pages fall under the “Your Domain” category and provide specific information on the data management needs and considerations for each domain. They also highlight challenges that are specific to each domain, such as data types, species, or areas, and offer solutions and considerations to overcome these challenges.

In addition, a new “Tool Assembly” page was added in the RDMkit corresponding to the Plant Sciences demonstrator use-case. Tool Assemblies serve as examples of how to combine different tools and workflows to support data management across different stages of the research data lifecycle. They also provide guidance on how to distil these assembly patterns into blueprints for combining tools in a more efficient and effective manner. Specifically, the plant phenomics tool

[4] https://doi.org/10.37044/osf.io/cuvqw
[5] https://rdmkit.elixir-europe.org/toxicology_data

This project has received funding from the European Union’s Horizon 2020 Research and Innovation programme under grant agreement No **871075**.

assembly was added to the RDMkit, which covers the entire life cycle of experimental plant phenotyping data.

## 4.2 Adaptation of the DSW KMs for the demonstrator use-cases

To facilitate the generation of DMPs for this deliverable, the DSW was used as a collaborative tool that enables data stewards and researchers to efficiently create DMPs for their research projects. The DSW is designed with a hierarchical KM that guides data stewards through a decision-making process to help them select the appropriate tools, resources, and practices when creating DMPs. Since the relevant information for DMPs can vary across different domains, these KMs can be customised for each demonstrator use-case.

Dedicated sessions were held with DSW experts and representatives from each demonstrator use-case, during which the general KM for DMPs was examined to identify areas that could be improved for each demonstrator use-case and questions to add. Deliverable D5.2 included the results of the sessions of four of the demonstrator use-cases (Plant sciences, Marine metagenomics, Human data, and Biomolecular simulation data). For this deliverable, the sessions focused on analysing the remaining demonstrator use-cases, namely Toxicology and Epitranscriptomics data.

Additionally, related to the Human Data use case, separate efforts are underway to enhance the sensitive data section of the KM system to ensure the proper management of such data. Finally, all the improvements and new questions suggestions that were found during these sessions were incorporated in the DSW KM by the DSW team.

## 4.3 Creation of DMP templates in DSW for the demonstrator use-cases

There are two standard approaches for creating something that could be called a “domain-specific” DMP, namely creating a KM (as explained in the section 4.2) or a project template (PT). Custom KMs are usually generated by “forking” one of the KMs provided by the Data Stewardship Wizard (DSW) registry. When doing this, all the pre-existing features of the KM (e.g. integrations, tags, document templates) can be reused. When modifying the forked KM, questions can be altered or added to address points that are not covered or to add domain-specific descriptions and examples with more specific jargon. When creating a PT, a set of answers (to the questions already present in the KM) is saved and can be used to generate a partially pre-filled questionnaire for a new project. In the ideal case scenario, the two methods can be used together to:

- Create a set of questions that better reflect a scientific domain, also in terms of jargon and, possibly, providing practical examples a researcher would understand.
- Provide domain-specific recommendations by answering the questions above, e.g. by indicating metadata standards, ontologies and a repository.

This project has received funding from the European Union’s Horizon 2020 Research and Innovation programme under grant agreement No **871075**.

The main advantage of using a PT is to provide actual recommendations (domain-specific in this case, but the feature can easily be used also for institutional or national aspects). Additionally, updating a PT once it has been constructed using one of the standard KMs is a straightforward process. The creation of a PT does not require any specific knowledge of the DSW and can be done by any subject matter expert. Moreover, this approach provides access to existing document templates, and DMPs can be exported easily into Horizon 2020 and Science Europe formats.

## 4.4 Data Protection Impact Assessment in DSW

At the BioHackathon 2021, a new KM was introduced to the DSW, enabling the creation of DPIAs directly within the tool. With this new feature, users can easily generate a DPIA document by filling in the necessary information within the DSW. Additionally, users can leverage the many built-in features of the DSW platform, such as version history, comments, and more.

This new KM and DPIA document output template in the DSW offers several benefits for users. First and foremost, it streamlines the process of creating a DPIA by providing a standardised template and guiding users through the process. Furthermore, since the DPIA can be created and managed within the DSW, users can take advantage of the platform's collaboration tools to work with others on the assessment, add comments and notes, and keep track of the version history. Finally, by using the DSW's built-in templates, users can easily export the DPIA into various formats, including Horizon 2020 and Science Europe, for submission to funding agencies and regulatory bodies.

This was done thanks to the integration between DSW and Data Information System (DAISY)[6]. The DSW, which for now is used mainly to help in data management planning, raises awareness for data protection requirements such as the DPIA. However, it is not specialised in DPIA reporting. At the same time DAISY, which allows institutions to keep a register of their projects using sensitive data and stores structured information on the project's GDPR-relevant aspects, lacks the means to combine the project facts into the narrative response needed in a DPIA. As the DSW and DAISY are highly complementary, it was decided to integrate the two to support DPIAs in DSW.

# 5. Results

## 5.1 New RDMkit pages for the demonstrator use-cases

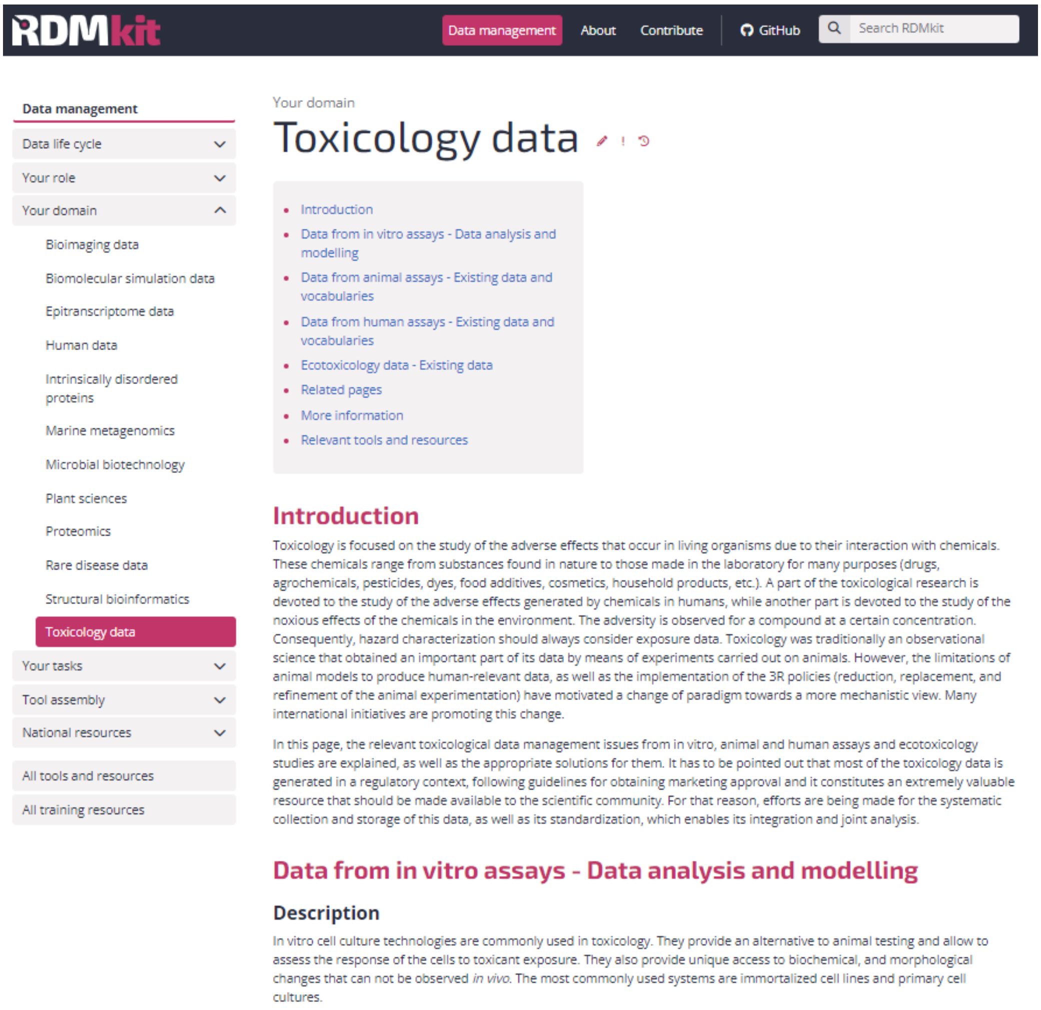
### 5.1.1 Use-case #5: Toxicology data “Your domain” page

This page provides an explanation of toxicological data management issues related to in vitro, animal, human assays, and ecotoxicology studies, along with their corresponding solutions. It is important to note that toxicology data is predominantly generated within a regulatory context, in adherence with

[6] https://doi.org/10.1093/gigascience/giz140

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No **871075**.

guidelines for marketing approval. This data is a valuable resource that should be accessible to the scientific community. Therefore, there are ongoing efforts to collect and store this data systematically, as well as standardise it to facilitate its integration and joint analysis.



Your domain

Toxicology data

- Introduction
- Data from in vitro assays - Data analysis and modelling
- Data from animal assays - Existing data and vocabularies
- Data from human assays - Existing data and vocabularies
- Ecotoxicology data - Existing data
- Related pages
- More information
- Relevant tools and resources

Introduction

Toxicology is focused on the study of the adverse effects that occur in living organisms due to their interaction with chemicals. These chemicals range from substances found in nature to those made in the laboratory for many purposes (drugs, agrochemicals, pesticides, dyes, food additives, cosmetics, household products, etc.). A part of the toxicological research is devoted to the study of the adverse effects generated by chemicals in humans, while another part is devoted to the study of the noxious effects of the chemicals in the environment. The adversity is observed for a compound at a certain concentration. Consequently, hazard characterization should always consider exposure data. Toxicology was traditionally an observational science that obtained an important part of its data by means of experiments carried out on animals. However, the limitations of animal models to produce human-relevant data, as well as the implementation of the 3R policies (reduction, replacement, and refinement of the animal experimentation) have motivated a change of paradigm towards a more mechanistic view. Many international initiatives are promoting this change.

In this page, the relevant toxicological data management issues from in vitro, animal and human assays and ecotoxicology studies are explained, as well as the appropriate solutions for them. It has to be pointed out that most of the toxicology data is generated in a regulatory context, following guidelines for obtaining marketing approval and it constitutes an extremely valuable resource that should be made available to the scientific community. For that reason, efforts are being made for the systematic collection and storage of this data, as well as its standardization, which enables its integration and joint analysis.

Data from in vitro assays - Data analysis and modelling

Description

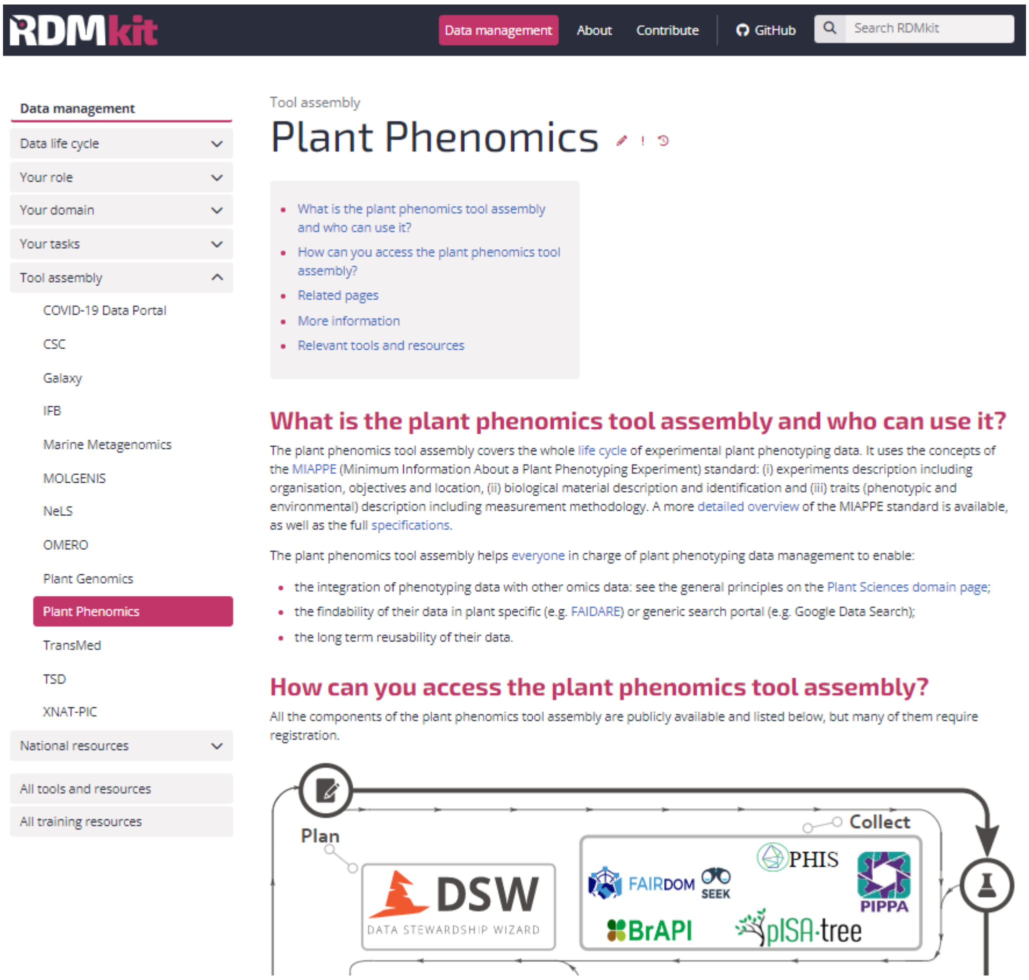
In vitro cell culture technologies are commonly used in toxicology. They provide an alternative to animal testing and allow to assess the response of the cells to toxicant exposure. They also provide unique access to biochemical, and morphological changes that can not be observed *in vivo*. The most commonly used systems are immortalized cell lines and primary cell cultures.

**Figure 1.** Toxicology data Domain page in the RDMkit.

### 5.1.2 Use-case #1: Plant sciences “Tool assembly” pages

The plant phenomics tool assembly is a comprehensive tool that covers the entire life cycle of experimental plant phenotyping data, based on the MIAPPE standard. It helps with the management of plant phenotyping data by providing an organised and structured approach to data management. The tool enables the integration of phenotyping data with other omics data, makes the data findable in both plant-specific and generic search portals, and ensures the long-term reusability of the data. The plant phenomics tool assembly is available to everyone in charge of plant phenotyping data management.

This project has received funding from the European Union’s Horizon 2020 Research and Innovation programme under grant agreement No **871075**.



Tool assembly

Plant Phenomics

- What is the plant phenomics tool assembly and who can use it?
- How can you access the plant phenomics tool assembly?
- Related pages
- More information
- Relevant tools and resources

What is the plant phenomics tool assembly and who can use it?

The plant phenomics tool assembly covers the whole life cycle of experimental plant phenotyping data. It uses the concepts of the MIAPPE (Minimum Information About a Plant Phenotyping Experiment) standard: (i) experiments description including organisation, objectives and location, (ii) biological material description and identification and (iii) traits (phenotypic and environmental) description including measurement methodology. A more detailed overview of the MIAPPE standard is available, as well as the full specifications.

The plant phenomics tool assembly helps everyone in charge of plant phenotyping data management to enable:

- the integration of phenotyping data with other omics data: see the general principles on the Plant Sciences domain page;
- the findability of their data in plant specific (e.g. FAIDARE) or generic search portal (e.g. Google Data Search);
- the long term reusability of their data.

How can you access the plant phenomics tool assembly?

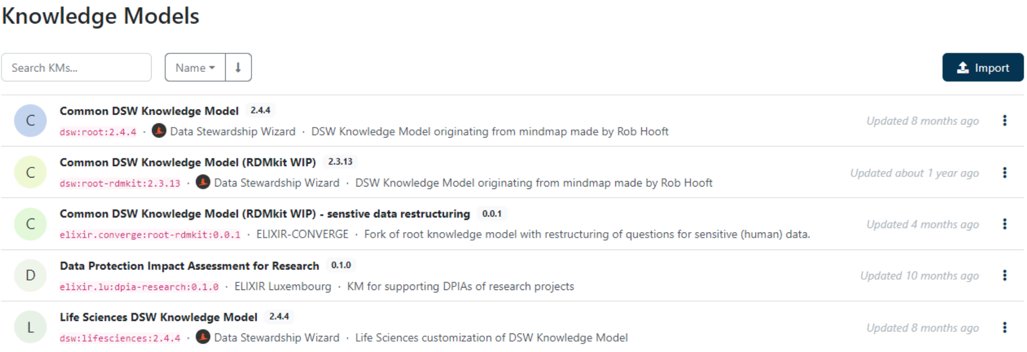
All the components of the plant phenomics tool assembly are publicly available and listed below, but many of them require registration.

**Figure 2.** Plant Phenomic Tool Assembly page in the RDMkit.

## 5.2 Adaptation of the DSW KMs for the demonstrator use-cases

The KMs in DSW contain the knowledge about what should be asked and how to get the necessary information from users to generate a DMP. These KMs are structured like a tree, with Chapters, Questions, Answers, and supplementary resources, and form the basis for the questions presented to users.

In the CONVERGE DSW instance, there are different KMs available for consideration (see figure 3). One of them is a more general one called "*Common DSW Knowledge Model*" and another specific for life sciences called "*Life Sciences DSW Knowledge Model*". Also, there are two KMs that that are Work in Progress (WIP) which are the "*Common DSW Knowledge Model (RDMkit WIP)*" that will include new links to the RDMkit to get further context of some questions and the "*Common DSW Knowledge Model (RDMkit WIP) - sensitive data restructuring*" that will improve the sensenstive data section of the Common DSW Knowledge Model. Finally, there is one KM that is used for the DPIA, described in Section 4.4.



Common DSW Knowledge Model 2.4.4
dsw:root:2.4.4 · Data Stewardship Wizard · DSW Knowledge Model originating from mindmap made by Rob Hooft
Updated 8 months ago

Common DSW Knowledge Model (RDMkit WIP) 2.3.13
dsw:root-rdmkit:2.3.13 · Data Stewardship Wizard · DSW Knowledge Model originating from mindmap made by Rob Hooft
Updated about 1 year ago

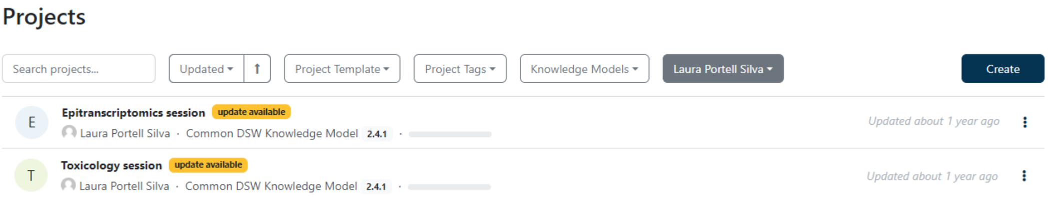
Common DSW Knowledge Model (RDMkit WIP) - senstive data restructuring 0.0.1
elixir.converge:root-rdmkit:0.0.1 · ELIXIR-CONVERGE · Fork of root knowledge model with restructuring of questions for sensitive (human) data.
Updated 4 months ago

Data Protection Impact Assessment for Research 0.1.0
elixir.lu:dpia-research:0.1.0 · ELIXIR Luxembourg · KM for supporting DPIAs of research projects
Updated 10 months ago

Life Sciences DSW Knowledge Model 2.4.4
dsw:lifesciences:2.4.4 · Data Stewardship Wizard · Life Sciences customization of DSW Knowledge Model
Updated 8 months ago

**Figure 3**. Data Stewardship Wizard Knowledge Models.

For this deliverable, the Common DSW Knowledge Model was used as a template to be evaluated, commented and refined by each of the demonstrator use-cases. This involved a gap analysis to identify any missing components and make the KM more domain-specific for each demonstrator use-case. This work was started in the deliverable D5.2 and it is continued here, for the demonstrator use-cases in Epitranscriptomics and Toxicology data. Figure 4 shows some of the DSW projects created to document this process.

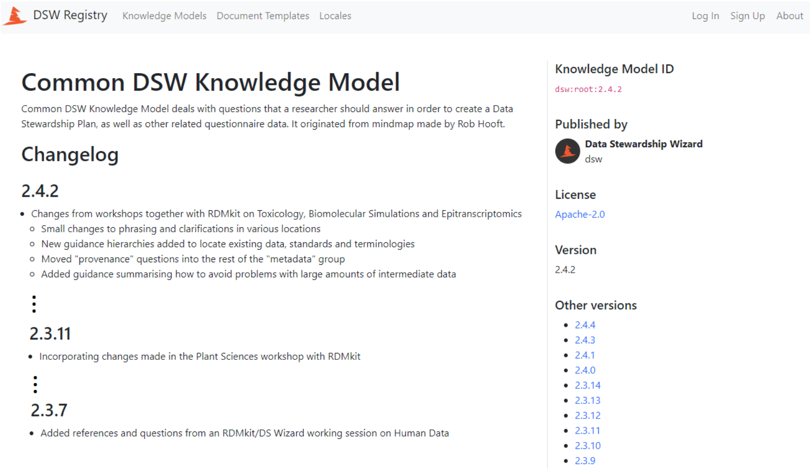


Epitranscriptomics session update available
Laura Portell Silva · Common DSW Knowledge Model 2.4.1 ·
Updated about 1 year ago

Toxicology session update available
Laura Portell Silva · Common DSW Knowledge Model 2.4.1 ·
Updated about 1 year ago

**Figure 4**. Data Stewardship Wizard projects.

As seen in Figure 5, the improvements of the Common DSW Knowledge Model that came out of the sessions with the different demonstrator use-cases have been incorporated in the KM in different versions by the DSW team. For the marine metagenomics use-case, the norwegian instance of the DSW was used and therefore it is not shown in the change log.



Common DSW Knowledge Model

Common DSW Knowledge Model deals with questions that a researcher should answer in order to create a Data Stewardship Plan, as well as other related questionnaire data. It originated from mindmap made by Rob Hooft.

Changelog

2.4.2

- Changes from workshops together with RDMkit on Toxicology, Biomolecular Simulations and Epitranscriptomics
  - Small changes to phrasing and clarifications in various locations
  - New guidance hierarchies added to locate existing data, standards and terminologies
  - Moved "provenance" questions into the rest of the "metadata" group
  - Added guidance summarising how to avoid problems with large amounts of intermediate data

⋮

2.3.11

- Incorporating changes made in the Plant Sciences workshop with RDMkit

⋮

2.3.7

- Added references and questions from an RDMkit/DS Wizard working session on Human Data

Knowledge Model ID
dsw:root:2.4.2

Published by
Data Stewardship Wizard
dsw

License
Apache-2.0

Version
2.4.2

Other versions
- 2.4.4
- 2.4.3
- 2.4.1
- 2.4.0
- 2.3.14
- 2.3.13
- 2.3.12
- 2.3.11
- 2.3.10
- 2.3.9

**Figure 5.** Common DSW Knowledge Model change log.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No **871075**.

The changes made in the KM after the workshops are shown in Table 1 and detailed in Annex 1. After the Human data use-case workshop, 6 questions were added to the KM, with its 10 corresponding answers. After the Plants sciences workshop, 2 questions were added with its 4 corresponding answers. After the Toxicology, Epitranscriptomics and Biomolecular simulation data use-cases workshops, a total of 29 questions were added with the corresponding 63 answers. In Annex 1, the links to all questions and answers added are listed. Additionally, edits on the guidance text for the KM questions were also done after all the workshops that resulted in DSW being more effective in providing clear instructions, explanations, or prompts to users when they encountered KM questions. This improved assistance will enhance the user experience, making it easier and more efficient for users to interact with the software and obtain the information they need.

**Table 1.** Changes made to the Common Knowledge Model after the sessions with the demonstrator use-cases.

| Version | Released after workshop | Added questions | Added answer entities | Edits of guidance text for questions |
|---|---|---|---|---|
| 2.3.7 | Human data | 6 | 10 | 3 |
| 2.3.11 | Plant sciences | 2 | 4 | 6 |
| 2.4.2 | Toxicology, Epitranscriptomics and Biomolecular simulation data | 29 | 63 | 47 |

For the Human data use-case, additional work is being conducted to adapt the Common DSW Knowledge Model with an improved version of the sensitive data section.

### 5.2.1. Use-case #2: Epitranscriptomics data

This session discussed the two methods of dealing with heterogeneous data in epitranscriptomics - RDA editing and chemical modification. RDA editing involves the use of existing standards and tools to edit non-transient data, while chemical modification is difficult and unstable.

Also, it was discussed that the sources of the data are mostly public databases. The question is how to describe and handle intermediate files generated during the project. If starting from public data, only the intermediate data needs to be kept in the project, and final results can be stored in REDIportal. However, in epitranscriptomics, the intermediate data can be very large. If the raw data can be submitted to public repositories, doing so early in the project can also avoid the storage responsibility in the project.

It is possible to ignore very large intermediate files since processing is fast (i.e. results are easily and faithfully reproducible, i.e. trade-off storage/reproducibility), but the workflow and containers used should still be recorded through workflows/containers (i.e. Docker) to ensure reproducibility. If a workflow uses a lot of compute power, it can be useful to use a workflow that can make use of parallel processing using many core systems. However, not all workflows are capable of using parallel processing if they have not been designed with this in mind from the beginning.

### 5.2.2. Use-case #4: Human data

As seen in Figure 3, there is work in progress to update the KM with several new additions to ensure the appropriate handling of sensitive data. First, there is a new administrative question to determine whether ethical approval is required. Second, new questions have been added to establish the legal basis for reusing personal data. Third, a GDPR-related section has been added in the event that new personal data is collected. Fourth, an ethical "consent" question has been added for any new personal data collection. Lastly, there are questions related to restrictions on the use of newly collected data, as well as text changes and external references to support the updated information. These additions aim to ensure the appropriate handling and protection of sensitive data.

Also related to the Human Data use-case, the ELIXIR Norway instance of DSW[7] forked the “Life Science DSW Knowledge Model” from the registry and added explicit references to national services that are relevant to Norwegian users. On top of this, several specific questions regarding sensitive data are added in a distinct chapter. Questions in this chapter are based on the Tryggve Checklist on ELSI issues and GDPR compliance[8]. This list is also available as a distinct KM on the SciLifeLab (SE) instance of DSW[9] and there is ongoing work to merge it into the common KMs available from the DSW registry.

### 5.2.3. Use-case #5: Toxicology data

During the session, the importance of guiding users to find existing data resources relevant to their field was emphasised. It was suggested that a clear and concise question needs to be added to facilitate this process under the "reusing data" section. In addition, it was highlighted that users often face difficulties in finding the appropriate data formats and terminologies/ontologies for their field, hence a similar question is needed to provide guidance to these users. The discussion also touched upon the significance of ensuring that the questions are easy to understand and cater to the needs of users from different domains, and that something similar would need to be considered for the rest of the demonstrator use-cases.

## 5.3 Creation of DMP templates in DSW for the demonstrator use-cases

For this deliverable, several PTs are being set up in the DSW based on the Life Sciences DSW Knowledge model, representing the recommendations from the demonstrator projects. A list of the available PTs follows:

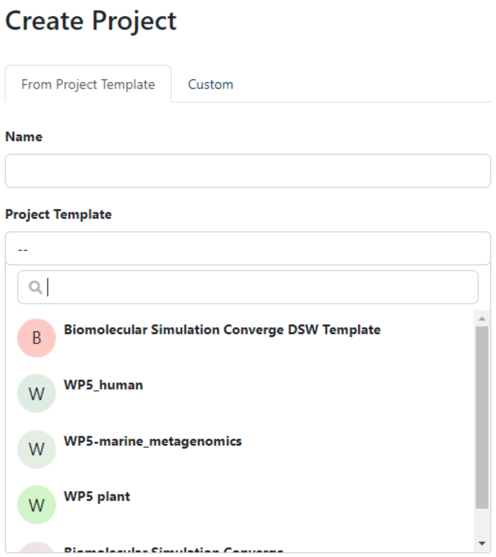
- Use-case #1: Plant sciences

[7] https://elixir-no.ds-wizard.org/
[8] https://neic.no/tryggve/links/
[9] https://dsw.scilifelab.se

This project has received funding from the European Union’s Horizon 2020 Research and Innovation programme under grant agreement No **871075**.

- Use-case #3: Marine metagenomics
- Use-case #4: Human data
- Use-case #6: Biomolecular simulation

Using this link (https://converge.ds-wizard.org/projects/create/from-template), a DSW user (login is required) would directly go to the page for generating a DMP based on any of the templates. As seen in Figure 6, a list of all the available project templates appear and any can be selected to start the work on the DMP.



Create Project

From Project Template | Custom

Name

Project Template

--

B Biomolecular Simulation Converge DSW Template

W WP5_human

W WP5-marine_metagenomics

W WP5 plant

Biomolecular Simulation Converge

**Figure 6.** List of available project templates in DSW.

The protocol to create and import project templates to DSW can be found in Annex 2.

### 5.3.2. Use-case #1: Plant sciences

To take as an example, the creation of the plant sciences PT[10] (you need to be logged in to access the PT) started with a reflection on what kind of data is common to all plant science projects. This reflection revealed that regardless of whether it's a phenotyping or genotyping project, the commonality lies in:

- Types/formats of data that will be used and associated standards
- Encodings/terminologies/vocabularies/ontologies
- Minimal Metadata Standards
- Repositories where to publish the data

These conclusions form the foundation for the plant sciences PT, ensuring standardised practices and facilitating collaboration and data interoperability across plant science projects.

## 5.4 Data Protection Impact Assessment in DSW

[10] https://converge.ds-wizard.org/projects/65a4e5fe-90b2-4573-aa98-01f8b4f65410

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No **871075**.

As mentioned in section 4.4, the integration of DSW and Data Information System (DAISY) enabled the facilitation of DPIA reporting, which was previously not possible through DSW alone. The integration has been designed in two independent tasks:

- By introducing a new KM[11] (you need to be logged in to access the KM) and a corresponding document output template, DSW enables the creation of a DPIA project directly within the platform. This content-based integration allows users to fill in the necessary information and generate a DPIA document seamlessly. Additionally, users can leverage DSW's built-in features such as version history, collaboration, comments, and more while working on DPIA projects.
- The technical integration of DSW and DAISY enables the linking of projects between the two platforms and the querying of DAISY project details through an API within DSW. This integration aims to leverage the GDPR-specific features of DAISY and enable users to use them while generating documents in DSW. With this feature, it becomes possible to add an appendix to the data management plan (DMP) that focuses solely on GDPR compliance.

# 6. Conclusions

This deliverable aims to provide valuable resources for life sciences researchers and data stewards to create reference Data Management Plans (DMPs) for their research projects. To achieve this objective, several activities were undertaken with a focus on the needs of the different domains covered by the demonstrator use-cases.

One of the main accomplishments of this deliverable is the extension of the domain pages in the RDMkit to include domain-specific information that can be used as a reference when developing DMPs for different research projects. In addition, a new Tool Assembly was added to the RDMkit, corresponding to the Plant Sciences demonstrator use-case.

To facilitate the generation of DMPs, the DSW was used as a collaborative tool that enables data stewards and researchers to efficiently create DMPs for their research projects. The DSW's KMs were adapted to address the specific DMP questions needed for each demonstrator use-case, ensuring that the relevant information for DMPs can vary across different domains and can be customised accordingly. During the workshops and feedback collection, insights were gained regarding the phrasing of the questions used to guide users in creating their DMPs and the wording of the questions was refined. Then, the KM became more user-friendly, making it easier for users to understand and respond to the prompts related to DMP development.

Moreover, the creation of DMP templates in DSW for the demonstrator use-case was accomplished, to provide domain-specific recommendations by answering the questions of the KM. Finally, a new KM was introduced enabling the creation of DPIAs directly within the DSW. With this new feature, users can easily generate a DPIA document by filling in the necessary information within the DSW.

[11] https://converge.ds-wizard.org/knowledge-models/elixir.lu:dpia-research:0.1.0

This project has received funding from the European Union’s Horizon 2020 Research and Innovation programme under grant agreement No **871075**.

In conclusion, this deliverable provides several valuable resources for life sciences researchers and data stewards, including extended RDMkit pages, customised DSW KMs, domain-specific DMP templates, and a new KM for creating DPIAs. These resources are designed to encourage good research data management practices across the EU research landscape, ensuring that valuable research data is effectively managed before, during, and after a project.

# 7. Impact

The resources provided in this deliverable play a crucial role in raising awareness and promoting the adoption of effective data management practices among life sciences domains that are represented in the ELIXIR-CONVERGE demonstrator use-cases.

By offering extended RDMkit pages and custom DSW KMs, these resources equip researchers with the necessary information and tools to develop comprehensive DMPs, which serve as strategic roadmaps for data management throughout the research project lifecycle as well as with guidance on good data management practices for projects consortia. These resources are starting to be used in new projects.

Good research data management ensures that research data is well-organised, easily accessible, and properly documented, allowing for transparency, reproducibility, and reusability of scientific findings. Effective data management practices also contribute to reducing errors, minimising data loss, improving data quality, and enhancing collaboration within the research community.

Finally, adopting standardised and domain-specific data management practices facilitates data sharing and collaboration across different research domains. It enables researchers to understand and comply with relevant regulations, policies, and ethical considerations related to data management and privacy.

# 8. Next Steps

The upcoming steps present a challenge as ELIXIR-CONVERGE is concluding, but they should involve gathering feedback from users and stakeholders from the new RDMkit pages to drive improvements. The project partners have disseminated these advancements to the corresponding ELIXIR communities, including the new emerging Data Management community. These communities will have a crucial role in collecting feedback from users and improving the RDMkit pages as well as DSW features. By encouraging community engagement and collaboration, we can ensure wider adoption and usage of the extended RDMkit.

# 9. Deviation from Description of Action

Based on the Description of Action, deliverable 5.2 was intended to cover information related to the first two use-cases, while deliverable 5.5 aimed to expand upon that by incorporating the remaining use-cases. However, the progress made with the use-cases has been non-linear, with certain use-cases contributing more input than others. As a result, for deliverable 5.2, information from all six use-cases was included. Then, in deliverable 5.5, the remaining information pertaining to all six use-cases is presented and discussed.

# Annex 1 - Added questions and answers to the Common Knowledge Model

Version 2.3.7 (After Human data workshop):

- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.3.7/preview?questionUuid=3782d32b-91b5-432f-8f93-92bb22868a22
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.3.7/preview?questionUuid=d0e029ee-aee0-420f-bc6f-ad471410ad42
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.3.7/preview?questionUuid=5d51a103-ec07-4c20-b269-f2f59f26d2cd
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.3.7/preview?questionUuid=8915bd25-db22-4ed6-bcc8-b1bbdc52989e
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.3.7/preview?questionUuid=de1b1a0f-58ea-4859-91b0-aa2b52090395
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.3.7/preview?questionUuid=f9572d3c-c91f-4945-a56b-ccfb32750fe9

Version 2.3.11 (After Plant sciences workshop):

- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.3.11/preview?questionUuid=f67a9910-4ea3-4280-b503-386dff3b8305
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.3.11/preview?questionUuid=cc57d7c0-49db-4546-958f-e9407eaf93ee

Version 4.2.4 (After Toxicology, Biomolecular Simulations and Epitranscriptomics data workshops)

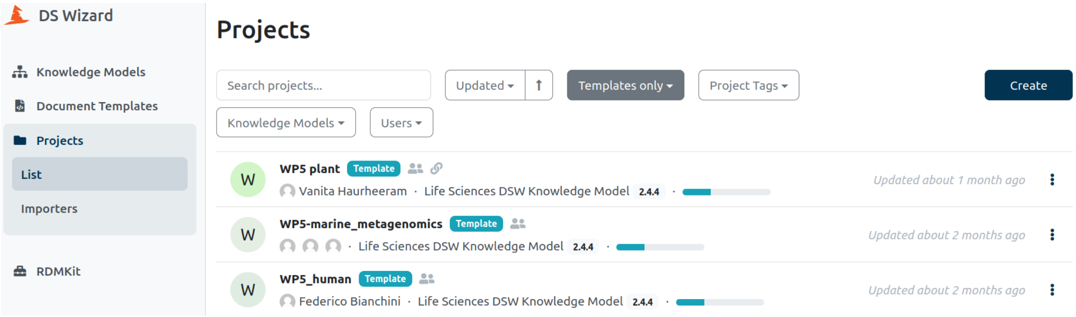
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=726e4a24-6d81-440b-8a70-5d36fa65c3a9
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=d4191666-21f3-4875-8426-d12a3aa1ffce
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=07f599c3-592c-479b-906c-36478ae792b9
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=a781badf-f7ee-4588-9478-d31470f00c38
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=190f7e34-c4f1-41c0-98e9-1f556b5b37b0
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=d4a879f1-0bad-4d6e-8ecc-07ae28e0848b
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=39a9f1e7-29ba-4c0d-8432-1b67b01fd7e6
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=ef5884e6-e61f-46e2-a940-ce64dd4c381f

- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=0c99e650-cbfe-445b-8e27-408aa5bd46f0
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=510aca7e-2844-4b27-8b89-83d0ed189312
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=6b122010-3847-4e56-8b0d-13f272407ddd
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=884585e3-85f7-4620-bf1e-5f302f9599b8
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=49b3ad26-d127-40ae-ab8d-fd256110c23d
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=6f2248f2-a65a-4ee8-9e87-0b7b6027c139
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=a0d52f2d-6828-42b0-a1f1-cda4ed7dad3b
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=b6d0922c-6e9c-45f8-b687-79310744601a
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=bc7538cd-2b28-446a-a448-2b316c4f36a2
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=f2bcb8d2-6a73-43a4-a2d3-16f50e4577a5
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=17464e0f-f477-4a45-a556-687a97b0e826
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=b19eab58-3b7f-4b66-b6f2-7e9c2c3c3e27
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=849cb43b-8622-4da0-9ef3-ae845dfe58f1
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=f714fa99-6c14-436c-ac62-2e7789322ecb
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=76a5112d-229f-4de2-8274-af20137f9553
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=d34f2b16-77f0-425a-9e83-7b28095a58ca
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=9deb2446-1a90-4837-af38-d9c33e67b0a3
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=881f6c68-53e0-4c05-b070-bb5a21a2ce5f
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=7cd59ba8-90c5-43f8-b5af-03a73fb301dc
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=340d766a-69ec-4805-8f73-6d2631dafe90
- https://researchers.ds-wizard.org/knowledge-models/dsw:root:2.4.2/preview?questionUuid=8dc09427-dcee-4df7-89c1-bb2515cfd95b

# Annex 2 - Protocol to add Project Templates to DSW

## A2.1. Create a Project Template in DSW

To create a Project Template (PT), you first need to have a Data Steward role in the instance of DSW in which you are working on. In the "Projects" tab, you can click on "Create" on the top right of the screen as shown in Figure A1.


DS Wizard
Knowledge Models
Document Templates
Projects
List
Importers
RDMKit

Projects

Search projects...
Updated ▾ ↑
Templates only ▾
Project Tags ▾
Create
Knowledge Models ▾
Users ▾

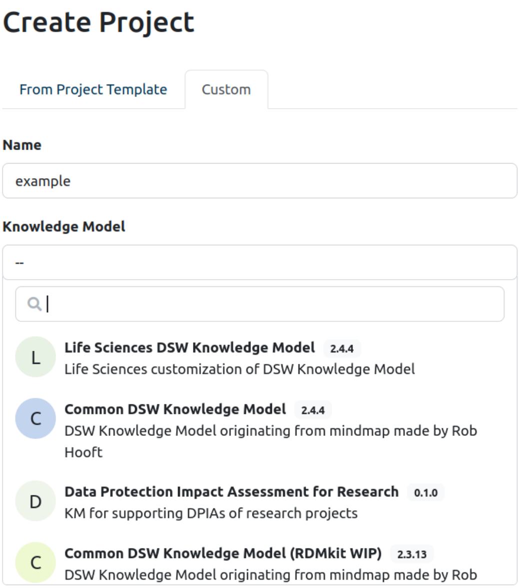
W WP5 plant Template
Vanita Haurheeram · Life Sciences DSW Knowledge Model 2.4.4 ·
Updated about 1 month ago

W WP5-marine_metagenomics Template
· Life Sciences DSW Knowledge Model 2.4.4 ·
Updated about 2 months ago

W WP5_human Template
Federico Bianchini · Life Sciences DSW Knowledge Model 2.4.4 ·
Updated about 2 months ago

**Figure A1.** Projects screen in the DSW.

Next, as seen in Figure A2, it is essential to give your PT a descriptive name that will attract a larger audience of data managers and researchers. Additionally, a KM can be selected, which is a checklist outlining the necessary information for constructing a DMP that your PT will be built upon. Multiple KMs are available, and you can find detailed descriptions of each one in the "Knowledge Models" tab of DSW.


Create Project

From Project Template | Custom

Name
example

Knowledge Model
--

L Life Sciences DSW Knowledge Model 2.4.4
Life Sciences customization of DSW Knowledge Model

C Common DSW Knowledge Model 2.4.4
DSW Knowledge Model originating from mindmap made by Rob Hooft

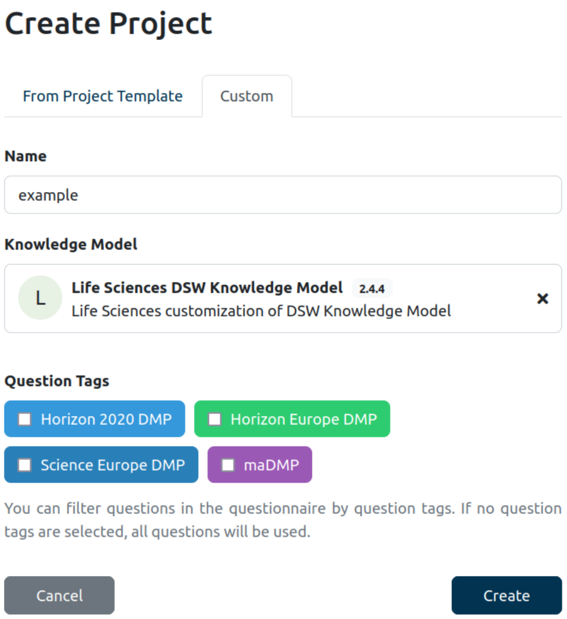
D Data Protection Impact Assessment for Research 0.1.0
KM for supporting DPIAs of research projects

C Common DSW Knowledge Model (RDMkit WIP) 2.3.13
DSW Knowledge Model originating from mindmap made by Rob

**Figure A2.** Knowledge Model selection in the "Create Project" screen in DSW.

Once you have chosen a KM, you can filter questions in the questionnaire by question tags (Figure A3). If no question tags are selected, all questions will be shown. Four question tags are available according to what type of project the DMP is for:

- Horizon 2020: for EU's research and innovation funding programme from 2014 to 2020.
- Horizon Europe: for EU's research and innovation funding programme from 2021.
- Science Europe: for major public organisations that fund or perform research in Europe.
- maDMP: for machine-actionable data management plans.



Create Project

Name

example

Knowledge Model

L Life Sciences DSW Knowledge Model 2.4.4
Life Sciences customization of DSW Knowledge Model ×

Question Tags

□ Horizon 2020 DMP □ Horizon Europe DMP
□ Science Europe DMP □ maDMP

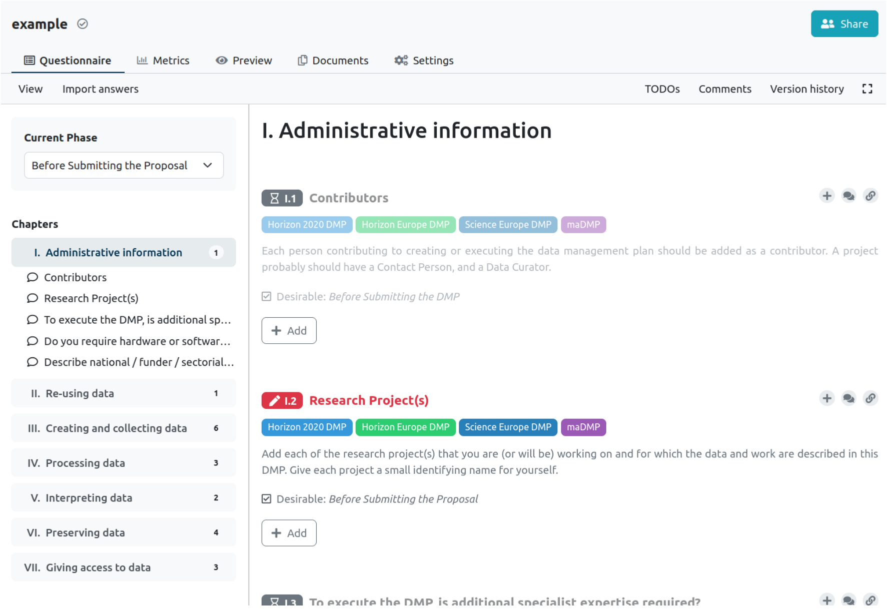
You can filter questions in the questionnaire by question tags. If no question tags are selected, all questions will be used.

**Figure A3.** Choose question tags in the “Create Project” screen in DSW.

KMs are structured in Chapters, Questions, Answers, and supplementary resources, and it forms the basis for the questions presented to users that want to create the DMP. For example, in the Life Sciences KM, you can find 7 chapters, as seen in Figure A4:

- Administrative information
- Re-using data
- Creating and collection data
- Processing data
- Interpreting data
- Preserving data
- Giving access to data

This project has received funding from the European Union’s Horizon 2020 Research and Innovation programme under grant agreement No **871075**.



Current Phase

Before Submitting the Proposal

Chapters

I. Administrative information 1

Contributors

Research Project(s)

To execute the DMP, is additional sp…

Do you require hardware or softwar…

Describe national / funder / sectorial…

II. Re-using data 1

III. Creating and collecting data 6

IV. Processing data 3

V. Interpreting data 2

VI. Preserving data 4

VII. Giving access to data 3

I. Administrative information

I.1 Contributors

Horizon 2020 DMP Horizon Europe DMP Science Europe DMP maDMP

Each person contributing to creating or executing the data management plan should be added as a contributor. A project probably should have a Contact Person, and a Data Curator.

☑ Desirable: *Before Submitting the DMP*

+ Add

I.2 Research Project(s)

Horizon 2020 DMP Horizon Europe DMP Science Europe DMP maDMP

Add each of the research project(s) that you are (or will be) working on and for which the data and work are described in this DMP. Give each project a small identifying name for yourself.

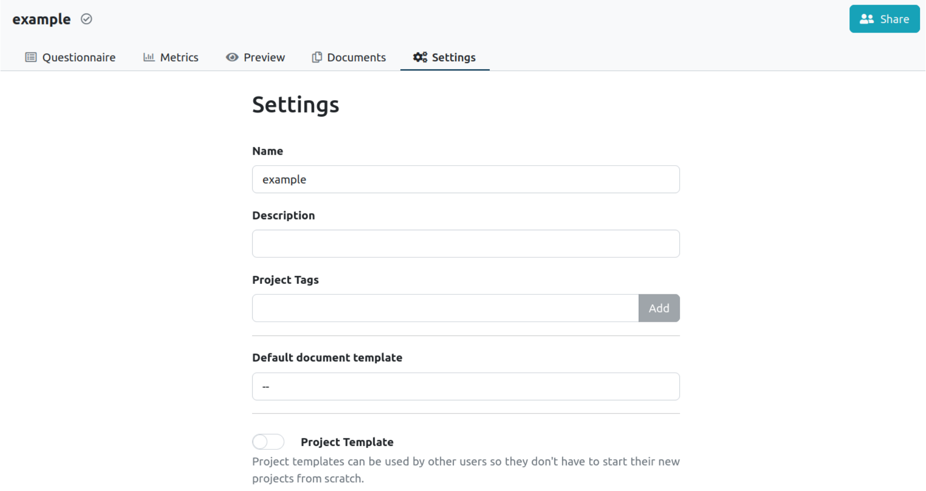
☑ Desirable: *Before Submitting the Proposal*

+ Add

I.3 To execute the DMP is additional specialist expertise required?

**Figure A4**. Questions in the Life Sciences KM.

To create your project template, the project template button needs to be activated. This button is in Settings as seen in Figure A5. Then, a description of what kind of project it can be applied to can be added (which is recommended).



Settings

Name

example

Description

Project Tags

Add

Default document template

--

○ Project Template

Project templates can be used by other users so they don't have to start their new projects from scratch.

**Figure A5**. Setting screens in DSW.

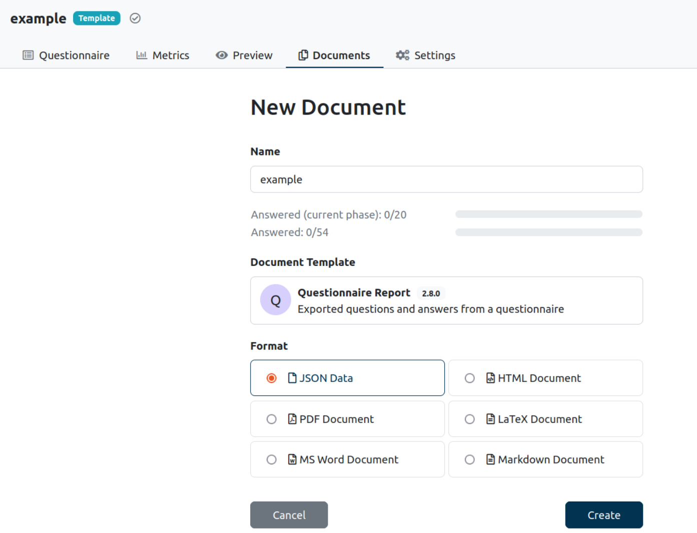
Then, in each chapter you can prefill some answers to the questions which can have the same answer for several projects, specially domain related questions (ontologies, vocabularies, document types, …).

## A2.2. Import a project template on a different instance of DSW

It is possible to transfer a project template into another instance of DSW (which enables the questionnaire report importation). For that, the KM used must also be available on it. The first step is

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No **871075**.

to download the prefilled answers of your PT, to do so, you need to go to the “Documents” tab of your PT and click on “New document”, and choose “Questionnaire Report” for the “Document Template” field and “JSON Data” for the “Format” field and click on “Create” (Figure A6).



New Document

Name: example

Answered (current phase): 0/20

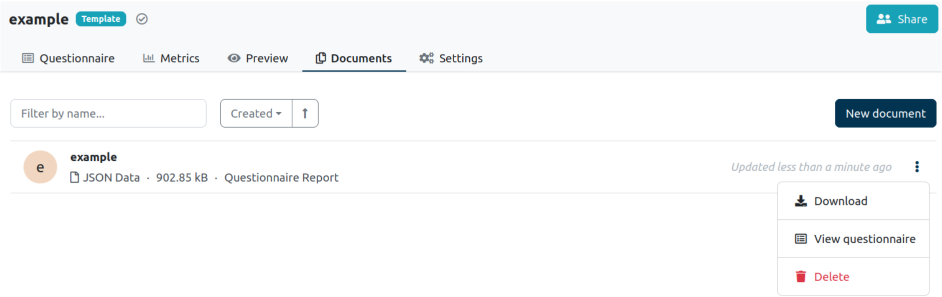
Answered: 0/54

Document Template: Questionnaire Report 2.8.0 — Exported questions and answers from a questionnaire

Format: ● JSON Data ○ HTML Document ○ PDF Document ○ LaTeX Document ○ MS Word Document ○ Markdown Document

**Figure A6.** “New Document” screen in DSW.

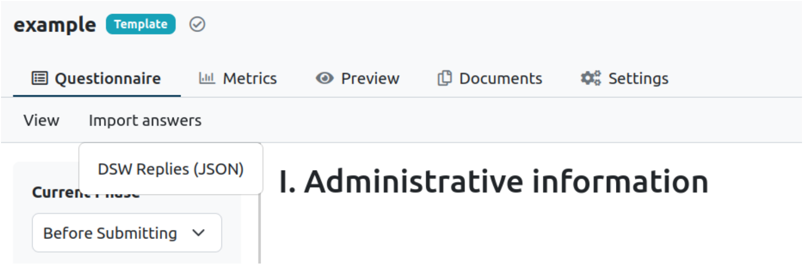
Then, it is possible to download it and use it in the second step.



example — JSON Data · 902.85 kB · Questionnaire Report — Updated less than a minute ago

**Figure A7.** Download button of a Document in DSW.

The second step is to create a PT on the DSW instance where the PT is imported with the same characteristics. On the questionnaire tab of your PT, click on “Import answers” and “DSW Replies (JSON)”, this will open a pop-up in which you can choose the previously downloaded file (Figure A8).



I. Administrative information

**Figure A8.** Import answers to a PT in DSW.