IMI2 Project 802750 - FAIRplus

FAIRification of IMI and EFPIA data

WP2 – Standards definition and process development

D2.2 BYOD Guidelines

Lead contributor	Nick Juty (4 – The University of Manchester)
	nick.juty@manchester.ac.uk
Other contributors	Melanie Courtot (1 – EMBL-EBI)
	Tony Burdett (1 – EMBL-EBI)
	Carole Goble (4 – The University of Manchester)
	Alasdair Gray (12 – Heriot-Watt University)
	Chris Evelo (5 – Maastricht University)
	Egon Willighagen (5 – Maastricht University)
	Nuno Nunes (5 – Maastricht University)
	Philippe Rocca-Serra (3 – University of Oxford)
	Dominique Batista (3 – University of Oxford)
	Sussana-Assunta Sansone (3 – University of Oxford)
	Manfred Kohler (2 – Fraunhofer E.V)
	Danielle Welter (7 – University of Luxembourg)
	Ibrahim Imam (11 – Imperial College London)
	Phil Gribbon (2 – Fraunhofer E.V)

Due date	30 Jun 2022
Delivery date	7 July 2022
Deliverable type	R
Dissemination level	PU

This project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No. 802750. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation and EFPIA companies.

Description of Work	Version	Date
	V1.0	7 July 2022

Document History

Version	Date	Description
V0.6	26 Jun 2019	First Draft
V0.7	18 Jun Jun 2022	Comments
V0.8	20 Jun 2022	Draft incorporating comments from Yojana, Danielle, Phil Gribbon (reviewer)
V0.8.5	26 Jun 2022	Nick incorporating comments from Alasdair, Ibrahim and Tony
V0.9	5 July 2022	Changes from WPLs
V1.0	7 July 2022	Final Version

Table of Contents

Executive Summary

Until now, there has been no standard or generalisable process for the 'FAIRification' of any given dataset. As a consequence, over the course of the FAIRplus project, it was necessary to develop such a methodology so that a) any given dataset could be onboarded for FAIRification (IMI or EFPIA); b) passaged through a defined and repeatable process, and c) could be shown to be demonstrably more FAIR at the conclusion of such a process. Originally, we conceived of a process whereby a FAIRplus 'FAIR-CMMI' team would collaborate with IMI projects through a number of "Bring your own data" (BYOD) workshops. However, early in the project, we refined our strategy such that these teams were known as 'squads' and collaborated closely with IMI projects representatives on an ongoing basis, rather than in discrete one-off workshops, but otherwise executed precisely the same functions (see 'Background').

We report here the development of a methodology, which has been iteratively refined over the lifetime of the project, to describe how the personnel responsible for the FAIRification will work, essentially providing a 'user manual' for those people engaged in the practical work.

For the FAIRplus project we formed 'squads', inspired by but differing from 'agile' and 'sprint' practices. Here, we populated teams across project-organisational and reporting boundaries, based on required expertise for specific FAIRification tasks. Squads were fluid in population, with members switching as required for specific tasks, and tactically responsive, with additional teams 'spun up' or down as required.

Squads worked in 3-monthly 'release cycles', targeting specific FAIRification tasks around which they would 'swarm'. Accompanying the defined steps (e.g. inputs and outputs) of the FAIRification, we have developed a variety of checkpoints and templates to assist in developing a repeatable process.

The squads have been an important mechanism in achieving FAIRplus project objectives, being heavily embedded in many major outputs, and acting as a point of contact with other FAIRification stakeholders such as RDMkit and Pistoia Alliance.

Background

The original grant agreement (pg. 165)[1] stated that "The FAIR-CMMI team will drive the technical implementation of the work plan, composed by cross WP representatives who lead the implementation of the strategy defined by the MB applying the FAIR-CMMI model". Furthermore, the following key areas to be monitored by the 'FAIR-CMMI team' we identified: 1) Selection of IMI projects and access to data; 2) FAIRification process, implementation and deposition; 3) FAIR dissemination and communication; 4) Data management; 5) Innovation management and sustainability. In essence, the 'FAIR-CMMI teams' were practically implemented as 'squads', and this document focuses on the way these squads worked directly towards, or contributed to, the objectives above.

At the outset of the project, the BYOD ('Bring Your Own Data') workshops were envisaged to be face to face events for the practical FAIRification of targeted IMI or EFPIA projects, over the course of a two to three day event (Objective 2 above). As our work began, we evolved this approach to be more collaborative on an ongoing basis. We determined that preparatory work, discussions and gathering of materials were time-consuming, particularly early in the project where no end-to-end process existed, and to maximise the impact of these discussions they needed to be continuous, rather than only taking place approximately once every three months. As the project has matured and best practices for engagement, through FAIRification and dissemination, have evolved, running a strictly time-boxed (2-3 day) BYOD event following the overall FAIRplus FAIRification process is something that we believe may be desirable: we have targeted such an event as being a key output from FAIRplus (see workshops).

Instead of discrete, occasional BYOD events, we converged upon the strategy to run 'squads' as persistent entities, engaged in constant FAIRification work, but then using the BYODs as communication, dissemination and capacity building events.

[1] https://drive.google.com/file/d/124C0egenKg3kiuOos6gfER6qo5jxJjg3/view?usp=sharing

Hence, these BYOD guidelines focus on the objectives above, and describe how to build and run 'squads' (aka FAIR-CMMI teams'). Ultimately, the value from the long-lived squads and the process by which they FAIRify datasets, rather than the BYODs; having persistent squad teams with expertise and a defined process makes BYODs effective, not vice versa.

This document describes the rationale for creating squad teams, determining their composition, how they function and interact inter-squad and intra-project, as well as externally to the project, how they implement the FAIRification process, contribute to the FAIR Cookbook, and share learnings.

Rationale

The aim of FAIRplus is to develop FAIRification processes for selected (IMI and EFPIA) datasets. A more detailed description of this selection process for datasets is described elsewhere[2], as it is not directly related to squad methodology.

The FAIRplus FAIRification process, developed and refined over the course of the project, is founded on:

- The FAIR principles[3] and indicators[4]
- A Capability Maturity Model (CMM)[5]
- Metrics associated with the CMM
- Methods for identifying and using standards
- A "Cookbook"[6] capturing developed guidelines that encapsulate how to FAIRify data
- Bring your own data events (BYODs) to design-build-test-learn the CMM, driven by real use cases

To test the FAIRness of the CMM and BYODs methodology, the project must ensure:

- The development is open to all FAIRplus consortium members, where the resulting process will be disseminated widely (externally) and open to all
- The developed processes and/or models are findable and accessible

[2] https://doi.org/10.1016/j.drudis.2022.05.010
[3] https://www.nature.com/articles/sdata201618
[4] https://datascience.codata.org/articles/10.5334/dsj-2020-041/
[5] https://en.wikipedia.org/wiki/Capability_Maturity_Model_Integration
[6] https://docs.google.com/document/d/1Az-HKYQPJPYyz6aNgwp0XF6o7Knr6hUYCqhKhWiNhds/edit#heading=h.iqlpgic7znm3

- The developed process and/or models are widely applicable outside the specific use case and dataset for which they were developed, where possible
- The processes and/or models are tool agnostic (i.e they can be instantiated by different toolsets) wherever feasible
- The processes and/or models are reusable by data and tool owners/providers independently
- Availability of concrete examples using IMI and EPFIA datasets and answer competency questions developed in collaboration with WP1

To make the FAIRification process credible for data providers, developers and managers:

- It must be developed and reviewed incrementally in an agile way, driven by datasets, user scenarios and competency questions developed with WP1
- It needs to be tractable
- It needs to be graduated and progressive.

Two development approaches are commonly encountered in software engineering: 'graceful degradation' and 'progressive enhancement'. Graceful degradation provides an alternative version of functionality even when a large portion of it has been rendered inoperative. Progressive enhancement starts with a baseline of usable functionality, then increases the richness of the user experience step by step by testing the support for enhancements before applying them. The idea is that the process will allow for movements across the capability maturity spectrum.

The FAIRification process used within FAIRplus transcends and crosscuts the project's work package administrative structure; many of the work package tasks are focused on specific areas and in the leadership of those tasks. Project work, however, requires organisation around specific problems, and will not necessarily or realistically align with those tasks described within individual WPs. Therefore, it is necessary in these instances to 'swarm' around a project and task, whilst maintaining accountability. This is common in software engineering; "Sprint" teams iteratively, holistically and incrementally manage product development as a unit to reach a common goal, with a very close and timeboxed collaboration of all team members across all disciplines. "Skunk works" teams are groups within an organisation given a high degree of autonomy and unhampered by bureaucracy, with the task of working on a project or task. FAIRplus adopted a hybrid of these approaches, developing its own derivative 'squad' methodology; this document describes the formation, composition and working practices of the FAIRplus

squads. This document does not report extensively on project selection[7] and the FAIRification process itself[8], which are reported elsewhere. This report is intended to provide sufficient information to allow one to follow the squad methodology, and to provide as many of the templates and supplementary materials as possible for that purpose. There may, however, be some materials that are not available until the project completes.

By creating teams which cross-cut WP and organisational boundaries, it is possible to focus on a specific problem, product, or to address a set of competency questions, rather than on the team's technical capability or duties. This also has the effect of creating a unifying purpose for a distributed and multi-disciplinary goal or project, as well as getting a sense of actual achievement and producing early results: the so-called "Minimal Viable Product" referred to in the FAIRplus proposal. An equivalent approach was taken in the 'Open PHACTS' project by organising around "task forces" instead of work packages, thus relegating the work packages to be a reporting framework.

We have developed the FAIRification process using this approach. This approach is further refined and defined by 'Spotify', where the task teams are called "squads". For further information[9,10]. A time stamped version (26/6/19) of this document forms the basis of the BYOD draft[11] submitted for MS2.1.

Methodology development

Within FAIRplus, we initially organised available personnel across teams as 2 squads, each dealing with 2 pilot projects each. Obviously, however, the number of squads can scale up or down with the number of projects and tasks to be addressed, as well as with available personnel with whom to populate them. One could also imagine their division across other axes, for instance across the 4 pillars of FAIR (F, A, I and R). Furthermore, a single squad may self divide into different task teams, each focusing on a discrete task towards a single overarching (FAIRification) goal.

FAIRplus initially identified 4 pilot projects (See D1.1)

- Oncotrack

[7] https://doi.org/10.1016/j.drudis.2022.05.010
[8] FAIRIfication process publication in process
[9] https://labs.spotify.com/2014/03/27/spotify-engineering-culture-part-1/
[10] https://labs.spotify.com/2014/09/20/spotify-engineering-culture-part-2/
[11] https://docs.google.com/document/d/1WpNYWFYF86uAwAo6PwlF8apTqGdudClUh_JlLhy09I0/edit#

- eTOX
- ReSolute
- ND4BB

Given available resources and personnel, those pilot projects were assigned across TWO squads, which themselves divided into task teams:

- Squad 1: Oncotrack and eTOX
- Squad 2: Resolute and ND4BB

The overarching aim was to develop methods for FAIRification that were applicable beyond the project task that they were developed for; the utility of a FAIRification process that applies only to specific data types, formats or proprietary tools/resources would be extremely limited, and not of general use to the wider community.

Having two squads allowed different techniques to be trialled, over multiple pilots per squad, and generated some "compare and contrast" learnings, as well as facilitating the management of logistics. Each squad was responsible for the entire FAIRification process for their assigned datasets. A brief description of the 'Squads Methodology' is provided in a planned publication[12], which provides greater detail on the FAIRification process steps, which are summarised below (see FAIRification process).

Squads Action Plan

The sections below, specifically 'Squad Scope' and 'Squad composition', are presented as a set of instructions to define the remit of the squads, and how they should be populated. These should be considered as a set of instructions or onboarding instructions for personnel being recruited to squad working.

Squad Scope

A **squad** is responsible for iteratively developing and refining FAIRification techniques that are used in the FAIRification of a specific dataset. Each squad will run for one or more **release** cycles, with each release refining the techniques used and aiming to ensure the datasets become incrementally more FAIR. Each release

12 https://docs.google.com/document/d/17wprJrzQrDlH8AqO_TfO0mGSp06L6-x66jybXoZ3MuY/edit#heading=h.olk6qo6y1zxu

cycle is time-boxed initially as being 3 months long, allowing them to be potentially executed between face-to-face (F2F) meetings (4 per year), where work and learnings are reviewed in person. These F2F meetings are also known as 'Bring Your Own Data' workshops (**BYODs**). At the end of each release, each squad must produce several outputs, to be shared with the other squad(s) after each release (see Squad Documentation and Monitoring). Outputs will be compared and techniques contrasted in a **shared retrospective**. Retrospectives occur between releases, and inform the approach taken by the squad(s) for its next release cycle. Where possible and appropriate, these retrospectives align with face-to-face meetings.

Squad Composition

Squads are populated from personnel available across WP structures, and selection is based on the skills and expertise deemed necessary for each particular project or task. The skill sets of FAIRplus participants were collected beforehand in an expertise spreadsheet[13] to facilitate decision making. It should be noted that the initial spreadsheets were not fine-grained enough to capture sufficient information to enable decision-making, and this process also required discussion with potential candidates.

Squads are **self-organising** around their dataset(s), which define the problem space for the squad. Each squad has operational freedom to define the tasks and subtasks necessary to FAIRify the dataset, as well as how these tasks will be run, for instance, whether they will use sprint cycles, how long sprints may be, task tracking (e.g. github issues, google sheets with hyperlinks, etc) and managing any code that is developed. Typically, github has been used to track individual tasks, usually accompanied by an overarching spreadsheet or google doc table to track multiple projects/tasks, linking to those individual github[14] issues. Over time, these initial squad-internal decisions were captured as part of the release cycle progression in output templates that evolved over time (see Squad Documentation and Monitoring).

Within each squad, there is an assigned **problem holder** who is the squad's representative from WP1 and has a deep understanding of the data produced by the pilot project. WP1 within FAIRplus is assigned with initial engagement and onboarding of IMI projects into the FAIRplus FAIRification process. This work is

[13] https://docs.google.com/spreadsheets/d/1S33X3J0ymtgv_TFpuFuRvxKbH90e7zPJCD8jBZnjzrQ/edit#gid=1871594603

[14] https://github.com/FAIRplus

described in a publication[15]. The problem holder is part of the team **responsible for specifying the desired outcomes** of FAIRification and **defining the success criteria**. After project engagement, the onboarding team (WP1, including the problem holder, and squad representatives) hands off to the squad team, who in collaboration with project representative(s) will define competency questions, and a ranked list of potential tasks. When defining success criteria, there is a consultation with the squad task team to design an approach that considers a broad range of potential users of the data and ensures FAIRification approaches cater to the broadest possible set of stakeholders achievable within a release cycle. As the squad iterated over releases, and as our methodologies improved, it was expected that more and more user groups would be supported by the solutions that we developed. The problem holder plays a key role in transitioning and 'hand holding' the project through the WP1 selection process, and into the squad working methodology. The problem holder helps liaise with project representatives throughout the lifetime of the squad release cycle(s).

Initially, within our methodology, we defined a role for a **skeptic** who would assume a skeptical and "friendly fire" view of the data. The skeptic should not be invested in the data or the final outcome, and would **challenge** the outcomes and success criteria. This role was later supplanted by that of an '**honest evaluator**' (see Squad Documentation and Monitoring).

Each squad also has a **squad leader**, who is responsible for driving the progress of the squad, as well as for overseeing the work, administration, completion of tracking tasks and squad output materials. The squad leader will recruit other personnel into their squad, and set the scope of development efforts (in consultation with the problem holder). They are ultimately **responsible for reporting progress** to the rest of the project (both internally to FAIRplus, and externally to project representatives) at a regular cadence and for **producing the final report**; squads nominally meet weekly to report progress; external (IMI project liaison) meetings are organised *ad hoc*, and as required or needed with project representatives. Squad leads are also responsible for logistics within their squad, for example, defining the agenda for squad calls, organising task tracking mechanisms, and communicating internally across WPs.

Initially, each squad had **representatives** from WP1, 2 and 3 covering as broad a range of expertise/ tasks as possible, and ensuring a good split between industry and academia. WP1 (onboarding of projects), WP2 (documentation and development of methodologies) and WP3 (implementation) have been ever present

[15] https://doi.org/10.1016/j.drudis.2022.05.010

participants, with WP4 (dissemination) and WP5 (fellowship) added later to facilitate output dissemination and fellowship[16] teaching modules. Task teams within squads directly facilitated incorporation of information into the cookbook[17], a major output of the FAIRplus project.

As the squad methodology matured over the course of the project, it was found to be necessary to also assign **deputy squad leads**, to cover some of the leadership and organisational aspects of the work.

The ideal squad is composed of 8 people. This is made up of one (possibly 2) people from WP1 (depending on expertise), 5 from WP 2 and 3 (at least one covering each identified task) and 2-3 additional "floating" squad members recruited depending on specific tasks or areas of need. Where necessary, for example if squads contained more than 10 members, additional squad teams could be spun off to address additional projects or tasks. For the duration of the project, a google spreadsheet was used to monitor squad composition[18], over time.

Squad Documentation and Monitoring

As mentioned previously, as our working practices evolved over the course of the project, we organised our working materials more thoroughly, as well as revisiting previous materials to bring them in line with the most recent methodologies, in an effort to provide consistent and thorough documentation for all projects that had progressed through our FAIRification process.

Squad materials were placed in the appropriate project specific google drive folder here[19], with the **squad leader being responsible for ensuring availability of materials and documents**, and for necessary communication tools, if any (e.g. email, slack, clickup, Freedcamp etc). Intra- and inter-squad communication, as well as across work packages and externally, played a crucial role in success of the project; this cross-pollination of ideas and strategies across the various teams engaged in the practical FAIRification of datasets, and well as engaging with external stakeholders, fostering collaborations as with RDMkit and Pistoia Alliance.

[16] https://fairplus-project.eu/get-involved/fellowship
[17] https://faircookbook.elixir-europe.org/content/home.html
[18] https://docs.google.com/spreadsheets/d/17i7w0OVvdiLDeQNtET60hmtvOTK2alm7YyPCwUKq3cI/edit#gid=0
[19] https://drive.google.com/drive/u/1/folders/1jZDbkeKM6O1ZoWr78jEqyfYyBSTzEYog

Individual project documents, checklists and forms created by squads are elaborated upon in the relevant sections below, with templates or examples linked from those sections to the Appendix, where appropriate.

FAIRification process

Over the course of the FAIRplus project, we have refined our FAIRification process and captured the relevant steps of that process, where work is delegated entirely to squads, into two distinct phase categories, designated as **orange** phase and **purple** phase. Orange phase essentially consists of the onboarding (with WP1), and focuses on 'FAIRification goal setting', and preliminary examination stages for a specific project, while the purple phase involves task identification, prioritisation and implementation stages. A visual representation of these two phases is provided below (Figure 1 - FAIRification process). Note that the orange phase incorporates a 'FAIR assessment' to evaluate the FAIRness of the data prior to squad interventions, while the purple phase terminates with a subsequent 'post-FAIRification' assessment, to determine the improvements made. This FAIRness 'delta' is presented for all projects that have progressed through our pipeline[20].

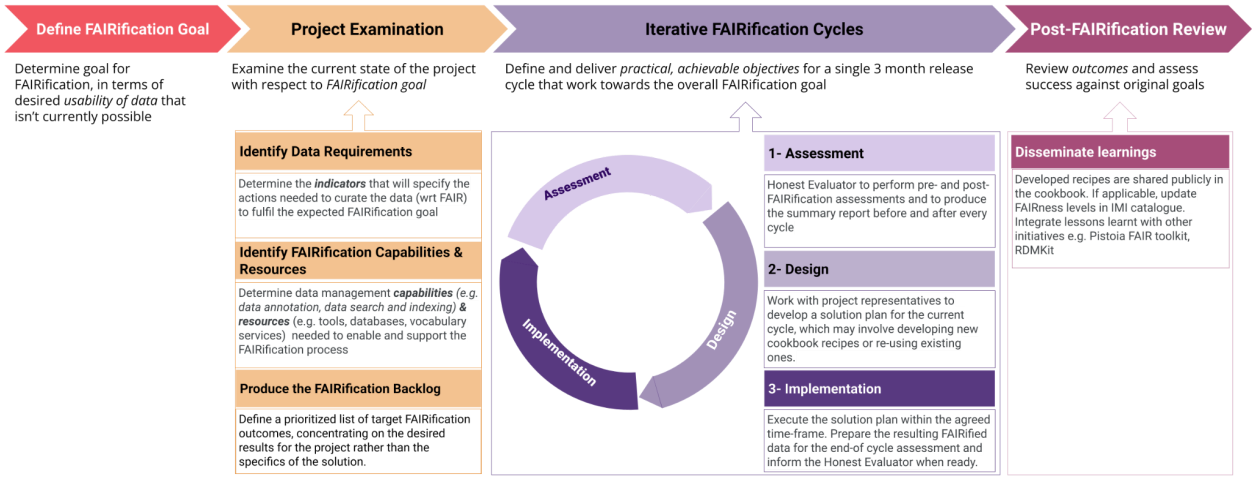


Figure 1. FAIRplus FAIRification process, highlighting the 2 phases (orange and purple), with preceding 'goal setting' (red), and subsequent post-fairification assessment and review (maroon).

Each phase has a defined list of outputs that must be generated to progress through the FAIRification process. These are summarised below, and links are provided to the template in the Appendix (Appendix A.1 - A.6). For instance, these

[20] https://fairplus-project.eu/impact/kpi-dashboard

documents provide varying levels of granularity and different perspectives or focus with respect to overarching FAIRification goals, task objectives and an overview of the total FAIRIfication process as a visual depiction. These documents are captured on a 'per project basis', and stored in the relevant project directory[21] of the FAIRplus drive. As each project passes through the FAIRification process, they must pass specific 'checkpoints' to ensure appropriate materials have been gathered to enable task progression, and that the tasks identified will address the FAIRification goal. This work is done in concert with WP1 and external project representatives, to ensure that those tasks identified meet both the needs of the FAIRplus project and meet the requirements of the external project.

Project selection and FAIR goal definition (orange phase)

Initially, WP1 conducted a selection process for appropriate projects, based on a number of criteria, and following an interactive survey[22] with project representatives. Over the course of the FAIRplus project, this process was refined to include squad members (usually squad leads and deputies), to more closely examine the data types, and the extent to which data would be available for each project. Project selection was driven by a number of criteria, including societal and scientific value; this work is described elsewhere[6,23] in more detail.

Orange (examination) phase

Following goal definition, potential tasks towards the FAIR goal(s), of which there may be more than one, are defined under different categories, including 'identifiers', 'metadata' and 'data standards', and which are categorised at 'Data requirements'. An early version of these categories provides more description, and is provided in Appendix A.1. This early version was later refined to that seen in Figure 1 above. Potential tasks are identified through discussion between appropriately qualified and selected squad team representatives, the squad lead responsible for that specific project, and project representatives.These discussions with project owners facilitates a shared and deeper understanding of the data, enable identification of possible and existing FAIRification capabilities and resources, and help define the data types that will be targeted in the process.

Discussions with project owners are generally carried out through a combination of teleconference calls and email; the means of engagement are agreed between

[21] https://drive.google.com/drive/folders/1cYh6gRSw5pErWCsljEyusI3kvGyrZtO1?usp=sharing
[22] https://doi.org/10.1016/j.drudis.2022.05.010
[23] https://faircookbook.elixir-europe.org/content/recipes/introduction/priorization.html

partners to come to a mutually convenient cadence and mechanism. Once the tasks have been identified, they are assessed jointly with project representatives on the basis of potential impact, resource required, personnel availability (on both sides as needed), and time required for implementation. Special care is taken to ensure that the tasks can be achieved in the release cycle timeframe (3 months from entry into orange phase proper) and that there are no external dependencies out of the control of the task team. In addition, a higher rating is given to the achievable solutions and implementations that are a) more generally exploitable outside of the project for which they were created, and b) would preferably result in a recipe for the FAIR Cookbook, or at least contribute to recipe improvement. More details on the FAIR Cookbook are available elsewhere (see D2.1; FAIR cookbook[24]).

All information relating to task identification are documented, usually as meeting minutes or in slides, but are also summarised in a tailored work plan document. While it is not necessary to share all the minutes for numerous discussions, the work plan document is used throughout the remainder of the FAIRification process, including through the implementation phase, and serves as a visual summary of the entire FAIRplus FAIRification journey. An example output at the end of the EBISC project is shown in Appendix A.2.

Project pre-assessment

An important preliminary step prior to any implementation work being undertaken is the assessment of the existing FAIR level of the datasets with which the squads will work. Initially, such assessments were made using the 'indicators' reported[25,26], following community discussion, by the Research Data Alliance (RDA)[27] FAIR Data Maturity Model Working Group. Unfortunately, these indicators were found to not be specific enough for our purposes, and more practically, are ambiguous in many cases; it was often the case that independent evaluation of a given dataset by multiple assessors would result in different scores. For this reason, over the course of the FAIRplus project, we have transitioned to a more appropriate set of 'Dataset Maturity Model' indicators (see Capability Maturity Model (CMM)). Since this transition resulted in 2 different types of results, we later revisited all projects and performed assessments using both indicator sets, to give consistency particularly for presentation of results.

[24] https://zenodo.org/record/6783564#.Yr213OzMKAk
[25] https://github.com/rd-alliance/FAIR-data-maturity-model-WG/issues/30
[26] https://datascience.codata.org/articles/10.5334/dsj-2020-041/
[27] https://www.rd-alliance.org/

Orange to Purple phase checkpoint

Following task identification and prioritisation in the orange phase, there is a checkpoint before the squad work moves into the purple phase. This requires a checklist of outputs from the orange phase. This checklist is provided in Appendix A.3. The required documents are gathered together in a project specific folder:

- Tailored FAIRification process diagram[28], as described above (completed example Appendix A.3)
- Tasks/requirements list[29], providing a list of tasks identified, category of task, and ranking (template, Appendix A.4)
- FAIR pre assessment[30], indicating the FAIR level of the dataset at engagement (see Capability Maturity Model (CMM))
- Table of responsibilities[31], listing squad personnel for the task team (template, Appendix A.5)
- An optional but recommended descriptive document relaying relevant details of discussions during orange phase, contextualising the decisions made
- Honest evaluation form[32] (Appendix A.6)

These materials are given to at least 2 Honest Evaluators, who will complete an appraisal to judge whether a) the relevant materials are present and correctly completed, b) whether the tasks identified in orange phase make sense logically, are realistic and achievable in the given time frame, and fit with the overall FAIR goal. With these criteria in mind, Honest Evaluators will complete their form, and return the appraisal to the appropriate squad lead leading that project team. Overall, this checkpoint arrives at a 'go', 'no go' or 'requires revision' decision. Both appraisals must agree on a 'go' position for the work to proceed. If a 'go' decision is

[28] https://docs.google.com/drawings/d/1BdQpRLk07sUDs5d1ucApFDeTa03HTROFVJxBHrYeyFg/edit

[29] https://docs.google.com/spreadsheets/d/1gHV1lPkwcaHYMVjGxGPGIF3HNhG3iuDn1yd-MqzNfgs/edit?usp=sharing

[30] https://docs.google.com/spreadsheets/d/1vaRf4PXJ7FvxLZpIKXRIkiEYDSZP5ajsb3KnETz7Em0/edit?usp=sharing

[31] https://docs.google.com/document/u/1/d/1XXNuQxvw_IOTwAB4_aRiEbcQ7H3FbhtHne9h0U-yv84/edit

[32] https://docs.google.com/document/d/1V58Gk5TGqALwfHFQVpLDEN5uStAvdabRgweibZ2Y6gI/edit?usp=sharing

not indicated, the work will return back to the squad team to rerun the process or revise the tasks in line with the guidance given by the Honest Evaluators. At this stage, particularly if the decision is 'no go', it may be necessary to repopulate the team, for instance, if the 'no go' decision arose due to a lack of necessary expertise or skill set within that team. Practically during these projects, there were no 'no go' decisions. There were a couple of instances where revisions were requested, which were minor, largely to do with updating project documentation.

Purple (Implementation) phase

There are a variety of possible tasks that can be involved in FAIRification, and for the purposes of the process description, we categorised them as 'identifier', 'metadata', 'ontology' and 'data sharing' strategies (see Appendix A.1). As part of the FAIRplus work, we have designed a 'reference' FAIRification process diagram (Figure 1), a 'template' (Figure 2), which lists all the categories of the orange and purple phases, as well as a work plan (for completed example, see Appendix A.2) which is generated by completing the work plan template for a specific project. This work is described in more detail in a planned publication (internal link). As part of the FAIRplus project, we have generated a FAIRification template, which specifies clear steps for the purple phase of the FAIRification Process.

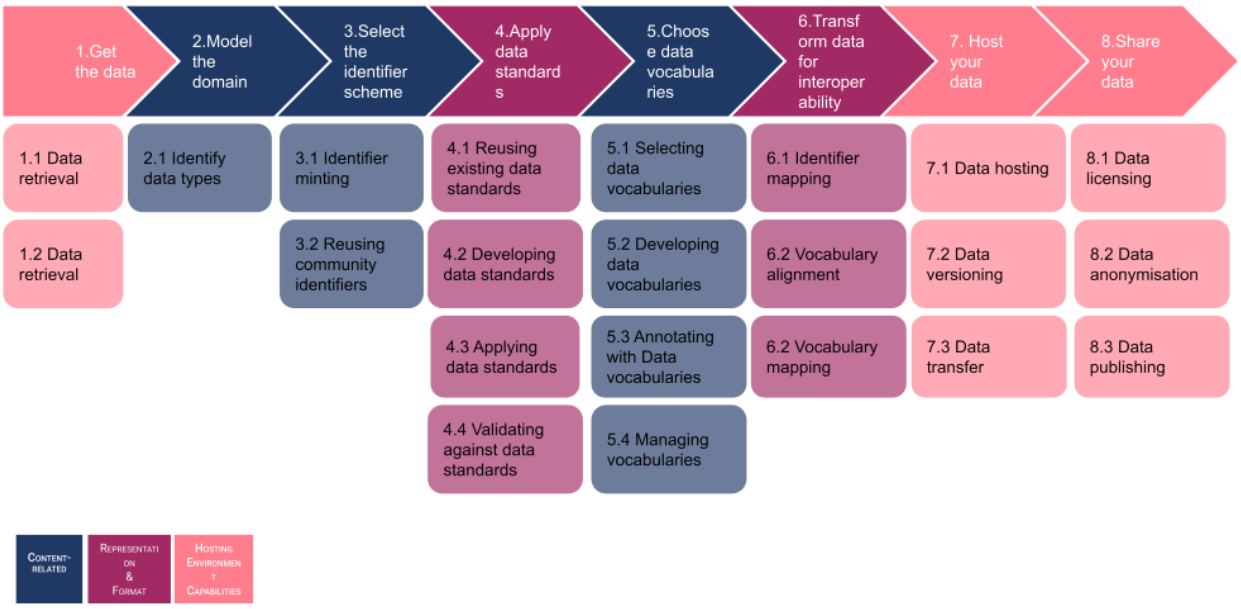


Figure 2. FAIRification Template

As described earlier, progress for a particular project is captured visually in the tailored FAIRification process diagram (or work plan), as well as in the 'outputs' document that is first used in the orange to purple phase checkpoint (Appendix A.3). This document is updated at the end of the purple phase and sent once again, with updated information including a post-FAIRification assessment, for Honest Evaluation (see above). This checkpoint determines whether the project can move on to the 'closing out' process, or else will be sent back for a further release cycle. It should be noted that projects may undergo a 2nd release cycle without repeating the orange (project examination) phase, if it is deemed appropriate or necessary, provided the appropriate documents are revised. While this did not happen over the course of FAIRplus (to date), in many cases, there was continued discussion with collaborating partners, as engagement tails off after the close out phase. These discussions largely focussed on clarifications as to how subsequent tasks by partners may be undertaken, following their work with FAIRplus.

Closing out process

During the early stages of the FAIRplus project, there was no formal closing out process to end collaborative work with IMI/EFPIA projects. This often resulted in confusion on both sides, and a more formal process was established in the latter stages of the project, in collaboration with WP4 and WP1.

The focus of this process is to:

- Get feedback from project partners on their general experience of their FAIRplus collaboration
- Get feedback from project partners for any specific resource that could be improved (tools, recipes, web pages, etc)
- Determine if initial engagement with WP1 was sufficient and clear, and whether any improvements could be made
- Discuss collectively whether the FAIR goal(s) had been achieved
- For FAIRplus squads to prepare summary slides (or narrative text) for dissemination of interaction details for both parties
- For squads to describe further steps for FAIR improvements to partners, with an indication of likely time requirements and impact
- To collectively identify 'success stories' or 'use cases' which could be used on web pages or form pamphlets for use on both sides and to assess the suitability of a join press release

This wide ranging set of objectives required participants in those closing out meetings from project partners, squad representatives (squad lead, and task team reps), WP1 representatives, and WP4 (outreach/dissemination).

The materials routinely used for this meeting are as follows:

1. Slide deck (~7 slides), summarising the project, FAIR goals, FAIR improvements and next steps for the partners
2. Updated work plan, with all task title and status of each at the end of the collaboration (eg. Appendix A.2)
3. The total set tasks identified with more detail (category, requirements, status) (eg Appendix A.4)
4. A spreadsheet containing all metrics pre- and post- engagement for the project
5. (Optional) Narrative description of the work done (approximately 1 side A4)

Subteams

Over the course of the work, some tasks were identified that could be run independently of the squad teams engaged in project work. Specifically, the following subteams were established in the early stages of the project to drive focused efforts.

Capability Maturity Model (CMM)

Since the RDA indicators were somewhat generic, sometimes ambiguous and overlapping, this hindered both their interpretation, application and implementation towards a consistent means of evaluation. To enable consistent and coherent processing of the diverse datasets acted upon through FAIRplus, we therefore need a better reference set of metrics and hence indicators.

The CMM[33], also known as the DataSet Maturity Model (DSM) is an enhanced interpretation of the FAIR Maturity Indicators created through community consultation by the RDA (see Project pre-assessment). This interpretation of the original FAIR indicators improves their application by categorising the area of focus of the individual indicators (representation & format', 'hosting environment' and 'content-related'). These indicators are additionally 'grouped', such that compliance to a 'set' of these indicators is sufficient to comply with a level of maturity. There are 5 maturity levels in this model, where '0' indicators no compliance, and '5' indicates 'enterprise' level compliance (Figure 3).

[33] https://fairplus.github.io/Data-Maturity/

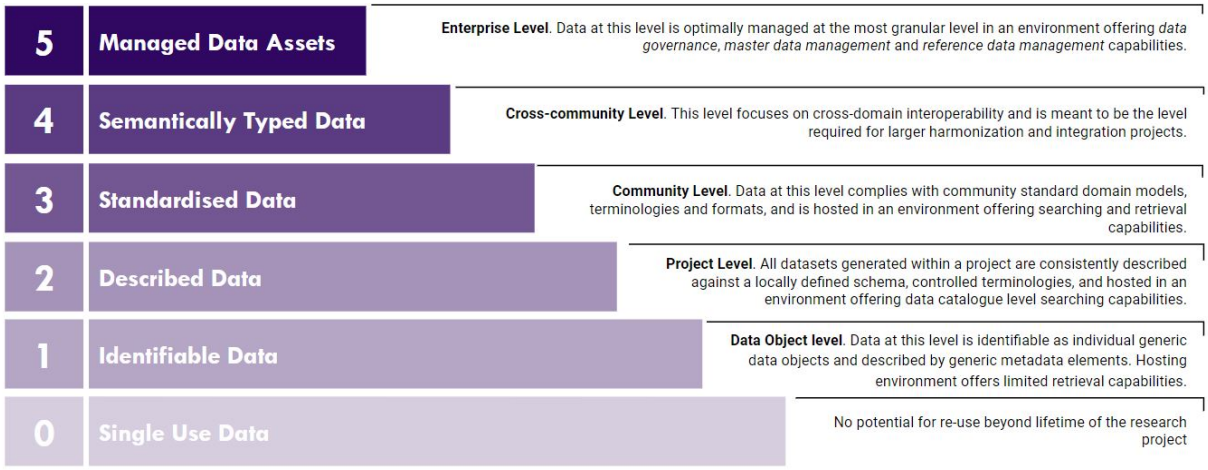


Figure 3. Compliance levels for the FAIRplus CMM, categorising dataset maturity as defined by compliance to sets of indicators developed by FAIRplus.

To simplify FAIR assessment using this maturity model, the underlying maturity model is being transformed into a guided tool, which is accessible through a web interface. Further details on this work will be reported in a deliverable (D2.6; FAIR-CMM) towards the end of the project.

Tooling & FAIR Wizard

This team was given a remit to tie together those priority outputs of the FAIRplus project that should be sustained following completion of the project, specifically the Cookbook and the CMM Maturity indicators. These works were pulled together into a tool, targeting project managers and data scientists, allowing them to define FAIRification goals, and highlight possible solutions, and directing them to specific Cookbook recipes. The FAIR Wizard[34] (Figure 4) collates those FAIRification goals from prior FAIRplus engagements. It is implemented as a web based survey, underpinned by a hierarchical decision tree through which a user is directed using a series of questions. The output directs the user to key areas they have identified through that survey that would benefit from FAIRification recipes. In addition, work is currently underway to link the CookBook to other external resources, such as the RDMkit[35] and the Pistoia Alliance FAIR Toolkit[36], to provide a more comprehensive

[34] https://wwwdev.ebi.ac.uk/ait/fair-wizard/
[35] https://rdmkit.elixir-europe.org/
[36] https://fairtoolkit.pistoiaalliance.org/

set of guidance, at different levels of granularity, depending to the user role (eg. lab scientist versus research manager).

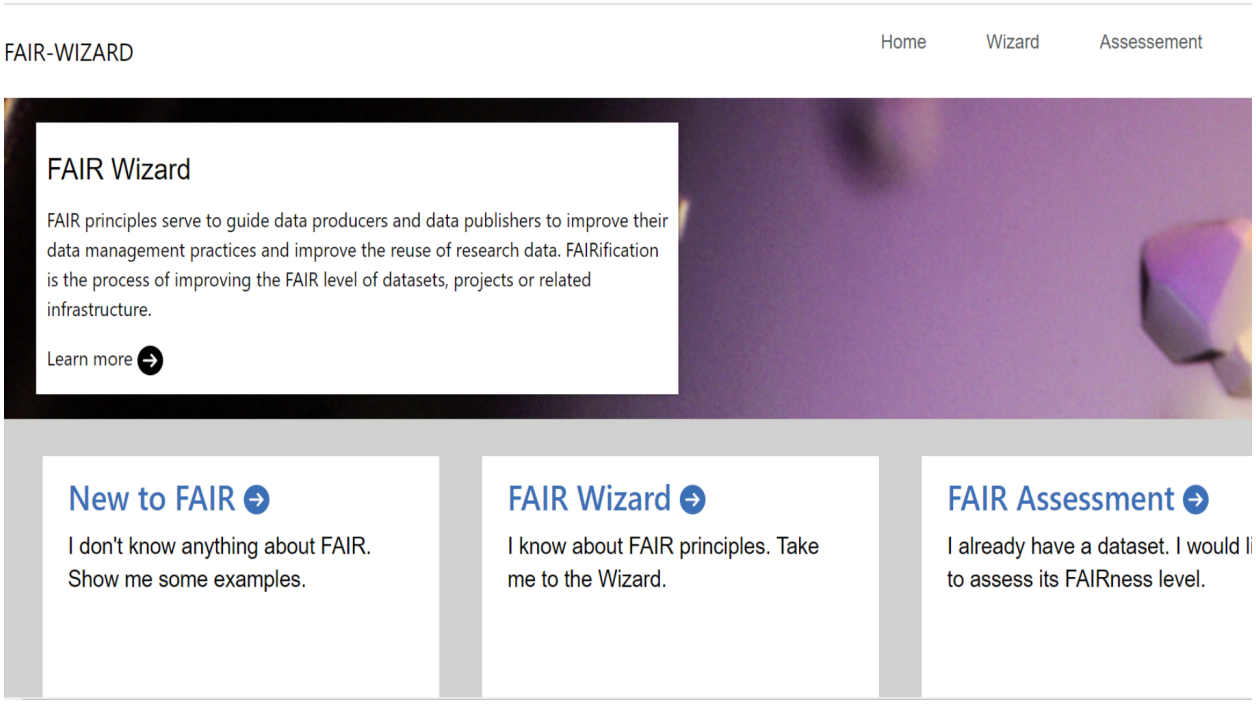


Figure 4. FAIR Wizard homepage, which provides access to introductory FAIR materials, FAIR assessment interface and the FAIR WIzard itself.

Squad Workshops (BYODs) & Retrospectives

As described in the original Grant Agreement[37], WP2 was tasked with defining the FAIRification process by implementing the standards and the metrics through three kinds of BYOD workshops: CMMI-Construction, Capacity-building, and FAIR implementation. Each BYOD targets different stakeholders, encompassing diverse data types, tools and different levels of initial FAIRness:

- FAIR-CMMI Construction - Targeted the development, testing and iteratively refining the FAIRification methodology with participants. Specifically, these also provided an opportunity to assess our FAIR-CMMI products (guidelines, metrics, FAIR Cookbook) and processes.

[37] https://drive.google.com/file/d/124C0egenKg3kiuOos6gfER6qo5jxJjg3/view?usp=sharing

- Capacity-building - Focused on the knowledge transfer of the developed FAIRification methodology within participants. These BYODs provide the opportunity to train data providers in the use of our FAIRification process.
- Implementation - Aimed to focus on: 1) the annotation of specific IMI datasets supported by FAIRplus participants; and 2) datasets in EFPIA organisations primarily run by industry participants. These BYODs mainly focused on hackathons, validation and assessment exercises.

As described earlier (see background), since the impact level of one-off BYODs was determined to be suboptimal, we implemented BYODs differently, and more efficiently, than they were originally described. The 'construction', capacity-building' and 'implementation' BYODs were encompassed by regular squad work, with BYODs manifested as knowledge sharing, testing, validation and retrospective opportunities. As such these varied in focus from developing FAIRification guidelines and processes, developing and refinement of indicators to test FAIRness, discussions around possible 'next round' projects (in collaboration with WP1 representatives), discussions on dissemination (with representatives from WP4 and WP5), and hands on FAIRification of individual datasets. Furthermore, to facilitate capacity-building, squads representatives participated in the WP4-led fellowship programme, where EFPIA delegates were trained on FAIRplus FAIRification process targeting their own contributed dataset. Specifically, squad representatives provided training, participated in WP4 fellowship events, and were included in communication channels, such as slack and email, to help guide Fellow through the process. This work is reported more extensively in D4.3[38].

Since squads operated largely independently, with weekly catch up meetings to track progress, it was necessary to share deeper details on strategies employed, as well as to get a higher level of the progress achieved across all squad and tasks teams, and to share learnings. For this reason, we scheduled a face to face meeting, targeted to coincide with the end of each (3 month) release cycle. This meeting essentially marked the transition from one release cycle to the next, where results from completed project works were shared, new projects were selected for the subsequent release cycle, and broader strategies and timelines were evaluated, to ensure the FAIRplus project would achieve its objectives. It was also an opportunity to share learnings between the squads and task teams, as well as refine and test project outputs, such as the FAIR Wizard, maturity model, and the Cookbook itself. The workshops which have taken place to date (June, 2022), described as 'Bring

[38] https://zenodo.org/record/3935396#.YrisXajMI2x

Your Own Data' (BYODs) in the original DoA, are found in Table 1. An example agenda from such a meeting (meeting 11, March 2022) is given in Appendix A.7.

F2F	Location, Host	BYOD type (focus)	Dates
1	Hinxton, EBI	B, C (Dataset, Retrospectives)	16th to 19th Apr, 2019
2	Erl Wood Manor, **Eli Lilly**	B, C; I (Dataset, Retrospectives; EFPIA Use cases)	9th and 10th Jul, 2019
3	Müllerstraße, **BAYER AG**	B, C (Dataset, Retrospectives)	10th and 11th Oct, 2019
4	Hinxton, EBI	B, C; I (Dataset, Retrospectives; EFPIA Use cases)	27th and 28th Jan, 2020
5	Virtual	B, C (Cookbook recipes, Maturity model)	28th to 30th Apr, 2020
6	Virtual	B, C (Cookbook recipes, Maturity model)	22nd to 24th Jul, 2020
7	Virtual	B, C (Dataset, Retrospectives, Tooling)	18th to 20th Nov, 2020
8	Virtual	B, C (Refining Process, Retrospective)	12th to 14th Apr, 2021
9	Virtual	B, C (Toolings, Maturity model)	19th to 21st Jul, 2021
10	Virtual	B, C (Dataset, Retrospectives, Process)	1st to 3rd Dec, 2021
11	Virtual	B, C (Documentation, Tooling)	22nd to 24th Mar, 2022
12	Berlin, **Bayer AG**	I (End to end process trial)	5th to 7th Jul, 2022
13	*planned*		TBC Oct, 2022

	Barcelona, BSG		

Table 1. List of Squad Face to Face workshops (until June 2022). Hosting venues were rotated between academic and EFPIA (bold) institutes, COVID-19 pandemic permitting. BYODs were focused on Implementation ('I'), Capacity-building ('B') or Construction ('C'). Note: most Implementation ('I') is run as a continuous squad activity, with BOYDs running with particular focus to refine methodology, share learnings, and to test and validate process and tooling. As the FAIRplus FAIRification process has evolved, at BYOD #12 we aim to trail our end-to-end process, something that was not possible at the outset of the project.

Discussion

The 'FAIR-CMMI teams' in the original Grant Agreement have been implemented as 'squads'. The squads, composed of cross-WP representatives, have been a major player in the FAIRplus project, being the practical means by which technical implementation of the work plan has been accomplished. Squads have contributed to or been entirely responsible for the identified objectives:
1) Selection of IMI projects and access to data; 2) FAIRification process, implementation and deposition; 3) FAIR dissemination and communication; 4) Data management; 5) Innovation management and sustainability.

The 'squad' methodology described here is the result of the iterative refinement of our initial approach (MS2.1), over the duration of the FAIRplus project. We believe this final form of the methodology has been very impactful with respect to achieving the FAIRplus project objectives, and that this was only possible due to the cross cutting nature of the strategy employed, specifically to reduce administrative and reporting overhead, for squad teams, and allowing the focus to be primarily on practical work of the squads.

Major refinements to our initial approach (MS2.1) had profound impact on our working practices, particularly for:

1. Communication with project representatives. Particularly for early projects with which FAIRplus engaged, we found that project representatives were being contacted by multiple FAIRplus representatives, independently, to discuss a variety of possible tasks in an uncoordinated way, painting a very confused picture. We discovered early on that we needed to appoint primary contacts on both sides of the collaborative partnership through which all such meetings should be organised, and that the lines for dialogue should remain open and active throughout the release cycle.

2. Expectation management. Again, early on, FAIRplus squads launched numerous tasks through squads to flatly address as many interventions as possible in as short a space of time. This of course elevated expectations on our partners' side, which needed to be reigned in. As we progressed through the project, we focused on more specific tasks and communicated accordingly.
3. Impactful task specification. Related to (2), as we progressed through the project, we focused more on 'bang for buck' or high impact tasks, rather than numerous smaller tasks that had little impact of FAIR improvements, since the metrics to measure such improvements we often not granular enough to measure or reflect small 'delta' changes in FAIRness.
4. Output orientated. Related to (2) and (3), squads focused on addressing tasks such that the outputs (for example recipes or guidance) could be used outside of the specific project for which they were created - specifically not tying into a 'point' solution which is not usable by anyone else.
5. Outreach and comms. We found that having visibility external to the project was increasingly important over the course of FAIRplus; it was easier to engage new collaborators (new IMI projects, or other stakeholders developing parallel resources in the FAIR domain) where our work and focus was clearly documented. Later in the project, we targeted the extraction of success stories and where possible published joint press releases.

In developing this methodology, we have been inspired by existing practices, particularly the 'spotify squads' (see 'Rationale' above), and have developed a strategy for coalescing a team around practically FAIRifying data. We have documented here a repeatable process, and have provided templates and guidance to enable others to follow this methodology.

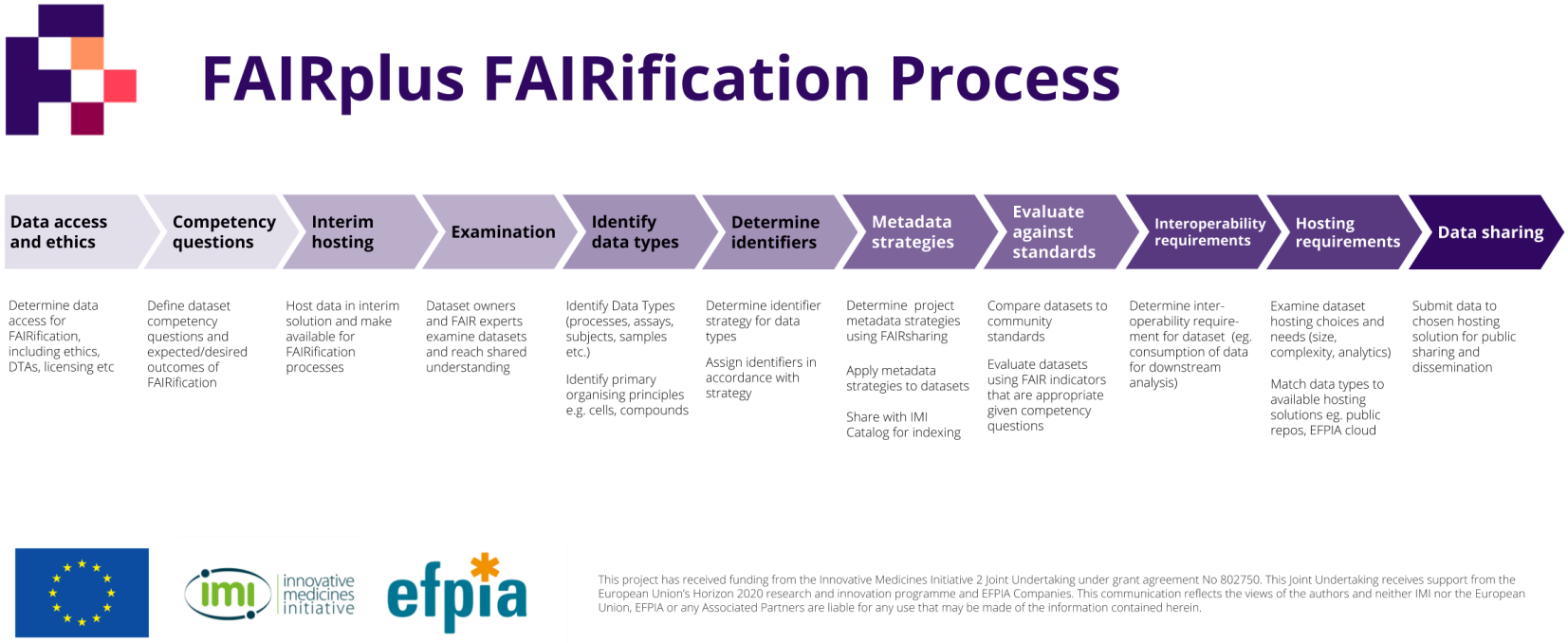
The squads have been an important mechanism in achieving FAIRplus project objectives, being heavily embedded in many major outputs, and acting as point of contact with other FAIRification stakeholders such as RDMkit and Pistoia Alliance.

Repository for primary data

The repository for primary data is the FAIRplus project google drive. Please contact the FAIRplus project manager for access. FAIRplus-PM@elixir-europe.org.

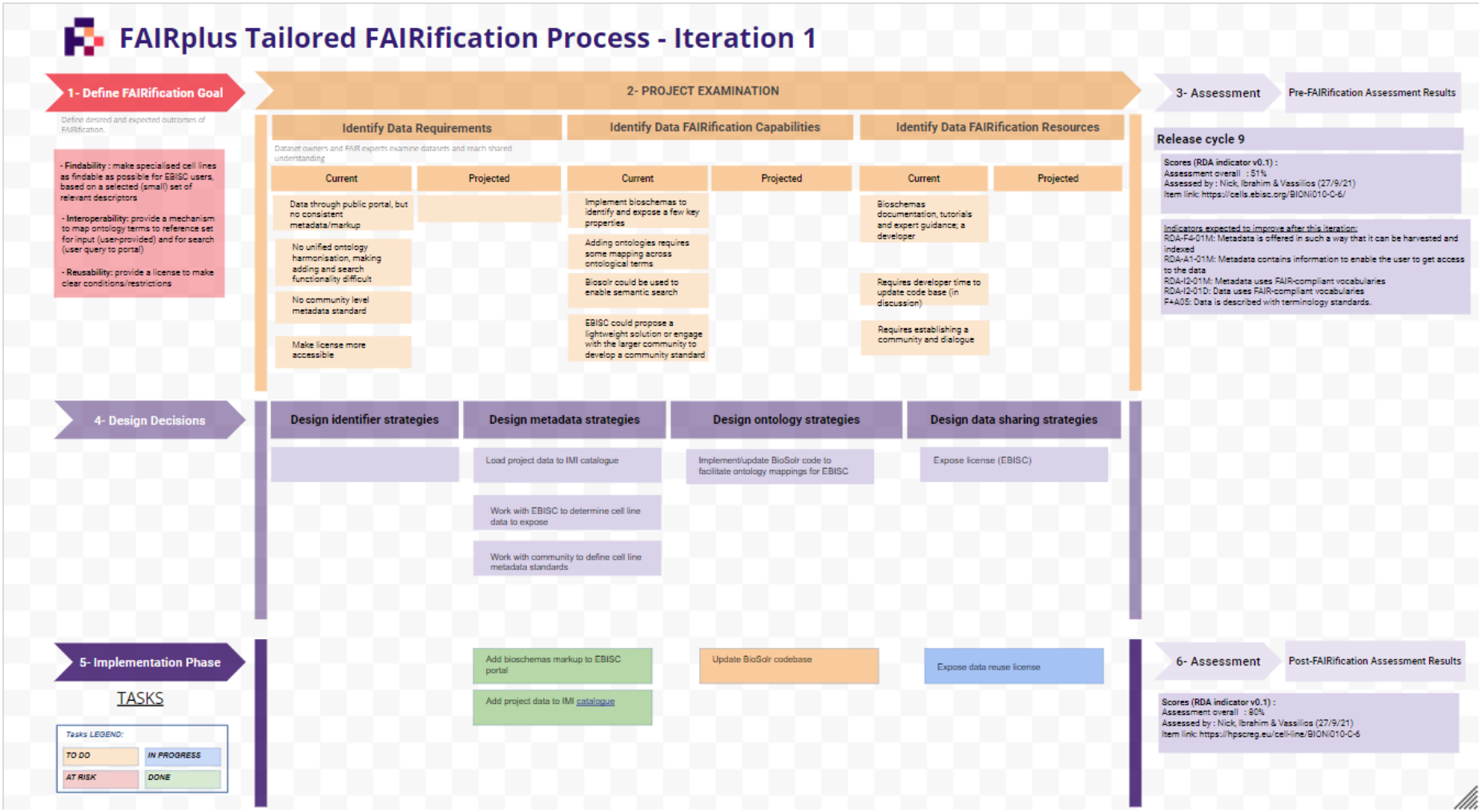
Appendix

Appendix A.1 Early version of the FAIRification process



This project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 802750. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA Companies. This communication reflects the views of the authors and neither IMI nor the European Union, EFPIA or any Associated Partners are liable for any use that may be made of the information contained herein.

Appendix A.2. Example Tailored Workplan (EBISC)



Appendix A.3 Orange and Purple phase outputs checklist

Add a ✅ instead of the 🔘, if the task is done.

Phase	Required Inputs	Required Outputs
Orange	🔘 WP1 Survey results for the selected IMI project (link here to e.g "FAIRplus_WP1_DataSurvey_ELSISurvey_IMIDIA_20191218") 🔘 WP1 Prioritisation scorecard for the selected IMI project (link here to e.g. "FAIRplus_PrioritizationScorecard_IMIDIA") 🔘 Table of responsibilities completed for at least Orange Phase 🔘 Access to project data (real or synthetic), metadata or data dictionaries, plus any other relevant documentation (as appropriate based on project-specific CDA) 🔘 A nominated primary contact within the targeted IMI project who understands the data (e.g. a data steward)	🔘 FAIRification Process Diagram template with section 1 (FAIRification goal), 2 (Project examination), 3 (pre-FAIRification assessment) and 4 (Design decisions) completed for the upcoming cycle 🔘 Requirements & Tasks spreadsheet with description of each use case, its current status, desired outcome and identified tasks and prioritisation by the data owner 🔘 Pre-FAIRification assessment using the RDA indicators of a set of data corresponding to the FAIRification goal 🔘 Table of responsibilities completed for the upcoming cycle 🔘 A link to a project folder in the FAIRplus Google Drive which contains all of the documents mentioned above, and potentially more: 🔘 OPTIONAL: A document outlining accompanying details from discussions between squads and IMI project members
Purple	🔘 FAIRification Process Diagram template with section 1 (FAIRification goal), 2 (Project examination), 3 (pre-FAIRification assessment) and 4 (Design	🔘 Updated FAIRification Process Diagram with section 5 (Implementation) and 6 (post-FAIRification assessment) completed 🔘 A list of new required recipes for feedback to the cookbook

	decisions) completed for the upcoming phase 🔘 Requirements & Tasks spreadsheet with description of each use case, its current status, desired outcome and identified tasks 🔘 Pre-FAIRification assessment using the RDA indicators 🔘 Table of responsibilities updated for the upcoming purple phase 🔘 OPTIONAL: A document outlining accompanying details from discussions between squads and the IMI project A current version of the cookbook (with an overview of existing recipes)	team (WP2) 🔘 An "MVP" implementation of the tailored FAIRification process 🔘 A post-FAIRification assessment of dataset(s) One of: 🔘 New FAIRification Process Diagram and Table of responsibilities for the next iteration of the purple phase (if applicable) 🔘 A summary report of all FAIRification cycles with links to all pre- and post-FAIRification assessments and FAIRification activities undertaken in each cycle
Sign -off	🔘 An implemented tailored FAIRification process 🔘 A summary report of all FAIRification cycles with links to all pre- and post-FAIRification assessments and FAIRification activities undertaken in each cycle	🔘 Update KPIs 🔘 A list of recommended future opportunities for IMI projects to work on 🔘 A presentation on the overall improvements made with IMI stakeholders (probably at squad F2F) 🔘 Add metadata to IMI data catalogue

The role of the phase lead is to ensure the outputs have been delivered, and the role of the honest evaluator is to check the outputs for each phase are delivered and to assess them for quality and appropriateness. This template was used to track inputs and outputs for both orange and purple phases of the FAIRplus FAIRification process.*Note: hyperlinks are to FAIRplus internal project directories and may not be accessible. Where appropriate, templates are provided in the appendix. This template is adapted for publication.*

Appendix A.4 Tasks and responsibilities template

Design Strategy	Summary	Description/Dataset Current status	Dataset Desired state	Recipes or Actions	Status	Assignee
Identifier	Add Identifiers	ZENODO identifier given, ChEMBL identifier to be added			Prioritised	
Metadata	Qualify metadata given in ZENODO	Mapping of metadata needed			Started	
Ontology	add Ontology link into metadata	Identify Ontologies needed and link to metadata			Started	
Data sharing	upload to ChEMBL	Dataset in ZENODO uploaded to ChEMBL will be delivered by Dec 2021	Published		Finished	

Appendix A.5 Phase responsibilities template

FAIR Pathway	**Phase**	**Targeted use cases**	**Squad Lead**	**Data Owner**	**WP1 Honest Evaluator**	**WP2/3 Honest Evaluator**	**Phase Responsibilities**	**Task team**
Molecular - prospective	Orange	*n/a*	*Oliver(+ backup)*	*Joe Blogs (joe.blogs@institute.org)*	*Peter*	*Bobby*	*Marcia*	*Marcia, Carole, Cindy, Mike, Greg, Oliver*
	Purple (cycle 1)							
	Purple (cycle n)							
	Sign-off							

Priority areas were identified in collaboration with stakeholders (IMI/EFPIA), and the nature of activity categorised as retrospective or prospective, with respect to the data available. *FAIR Pathway: examples include (further pathways may be added to this list):

- Molecular - retrospective
- Molecular - prospective
- Clinical - retrospective
- Clinical - prospective

*Targeted use cases: examples include (further use cases may be added to this list):

- metadata standards

Example information is given in italics. Any resemblance to 'Brady Bunch' characters is purely coincidental.
Note: Honest Evaluators should not be active participants in the task team.
Note: This template is adapted for publication.

Appendix A.6 Honest Evaluation template and guidance

Project name: x
Honest Evaluator name: NAME
Date of evaluation requested: DATE
Date of evaluation finished: DATE

Definition of an Honest Evaluator
The Honest Evaluator (HE) is a person responsible for providing oversight of the end-to-end FAIRification process, taking a work-package-specific perspective. Tasks include e.g. checking for alignment of Work Package tasks with squads plans, assessing efficiency of squad work, and ensuring squads do not get side-tracked. Naturally qualified people are those indicating "Active advice" from the participant survey.

Who is an Honest Evaluator, and what do they do?
The Honest Evaluator needs an overview of the project, but stands outside of the process - they are a reviewer who is not a squad leader, neither assigning tasks, but: reviewing the process, and giving (honest) feedback.

Honest evaluators are invoked at:
- Boundary from orange to purple phases
- Boundary at the end of each purple cycle

Requirements on the output of the Honest Evaluator:
- Provide an overall decision and a summary report (max 1 page), for which the Honest Evaluator should not invest more than 1h of their time

FAIRification Goal Guidelines
A good FAIRification goal communicates clear scientific value ("why"), defines a specific scope ("what"), and is actionable (with a specific endpoint). To pass review, FAIRification goals must:
- Define the impact - it should at least benefit the project, better would be to benefit the project and FAIRplus, best is to benefit community
  - Recognise project needs and why they are necessary
- Concentrate on scientific value over technical value; "being more FAIR" is not in and of itself a goal
- Avoid using words like FAIRify, FAIRification, in goal descriptions
- Try to be explicit; it's easier if the scientific value can be communicated rather than assumed (e.g. submit to public repositories has implicit value)
- Multiple FAIRification goals are acceptable and should be documented separately:
  - One goal should drive one purple iteration, making clear which goal is under consideration where appropriate
  - Iterations of purple phase (for individual goals) can be run sequentially, or in parallel

Form for completion by Honest Evaluators

Question	Revisions (major-minor)	Next steps evaluation	Go / No-Go?
What is the overall decision on the squad outputs (ok/minor revisions/major revisions/outright reject)			
Is the overall fairification goal clear (see guidelines below)?			
Is the overall FAIRification goal the right goal (ie Have we prioritized the right thing? resources/efforts-benefit?)			
Was the goal for the last cycle clear?			
Was the goal achieved?			
Is the goal for the next cycle clear?			

Appendix A.7 Squad F2F March 2022 Agenda

Note: embedded links may not be accessible to non FAIRplus participants

Agenda

11th Squad Virtual "Face to Face" meeting

Date: 22-24 March 2022

Virtual meeting rooms: (click on "join in browser" if you don't want to download the app)
Link to Zoom meeting for all 3 days:

Join Zoom Meeting
https://elixir-europe-org.zoom.us/j/82421119390?pwd=dUo4K1h6cFUxbmtpZVJlVnpyTnVidz09

Meeting ID: 824 2111 9390
Passcode: 023134

To dial in via telephone, find your number here: https://us02web.zoom.us/u/kewkzhZ5n

The folder containing all documents related to these meetings can be found here (this includes the minutes).

Objectives

This meeting provides:

- Review of each project processed by FAIRplus so far with respect to
    a. Completeness of documentation
    b. Standardised set of input and output documents for each phase
- Strategic forward planning for the next delivery cycle and the remaining project time

General rules/set up of the virtual meetings:

- Try to join the call a few minutes before the official starting time to make sure that your sound settings are working
- Mute your microphone when you are not talking
- Leave your camera on as much as possible (if your home situation allows), unless this has consequences for the connection
- Indicate when you are away from the keyboard (afk), but stay logged in into the call

- Additional communication channels during the calls are: Slack (indicate if you are not on this), minutes document, Github (if you have no access yet, communicate your Githandle to us).

Organizer / responsible admin person

Please contact these persons via Slack, chat or email for questions regarding organization etc. :

- Ibrahim Emam, Nick Juty, Tony Burdett, Danielle Welter

Attendees

(Day 1-2): *see minutes doc*
(Day 3): *see minutes doc*

Agenda

Time (CET, GMT)		Minutes link
Day 1: Project Review		

10:00
9:00
	• Introduction - Tony
	• Project summaries - Nick
	  ○ What materials are needed by the end of the purple phase, and the end of the orange phase?
	  ○ Exemplar folder from EBISC, introduce the tracking sheet
		Slides (Nick)

11:00 10:00	Coffee break	
11:30 10:30	• Collate all project summary materials in breakouts (2-4 people per group) - Nick	
13:00 12:00	Lunch break	
14:30 13:30	• Short review of progress (5-10 mins) - Nick • Collate all project summary materials in breakouts (continued) (2-4 people per group)	

15:45 14:45	Coffee break	

16:00
15:00
	• Review gaps from tracking sheet - Dani
	  ○ What's the right assessment?
	  ○ How do we trace decisions that were made? (from goal or assessment results)
	• Prioritise gap filling exercises for day 2
	  ○	

17:30 16:30	End Day 1	
Day 2: Capturing Outcomes		
10:00 9:00	• Introduction and recap of day 1 - Tony	
10:30 9:30	• Gap filling task teams in breakouts	
11:30 10:30	Coffee break	
12:00 11:00	• Gap filling task teams in breakouts	
13:00 12:00	Lunch break	
14:30 13:30	• Review and reprioritise • Fill remaining gaps, in breakouts or plenary depending on progress	
16:00 15:00	End day 2	

Day 3: Delivery Planning Day		

10:30
9:30
	• Flash updates (10 minutes plus questions each) - Dani chairing
	  ○ FAIR Wizard (Fuqi)
	  ○ Maturity Model (Ibrahim)	

	○ Cookbook (Philippe) ○ Tool discoverer (Eva)	
11:30 10:30	Coffee Break	
12:00 11:00	• Brainstorming - Defining use cases for technical integration of FAIRplus outputs - Tony	
13:00 12:00	Lunch Break	
14:30 13:30	• CMM exemplar appraisals (Ibrahim) • Review and recap - did we get everything to where it should be? - Ibrahim • What still remains to be done? • How best to deploy materials externally (and will they be FAIR)?	
15:30 14:30	• Project prioritisation by squads - Tony chairing • Actions review / update • Wrap-up and planning for remaining F2F dates	
16:00 15:00	End Day 3	

A Google drive folder is available in which you will find all documents related to these meetings.

Other important documents and links:

list of attendees

Results folders IMI datasets:
ULTRA DD/EUbOPEN
EBiSC
IMIDIA/RHAPSODY
APPROACH
ABIRISK

Cookbook related
GitHub Cookbook repository
Public version Cookbook
Other:
MTR report

Phases Feedback
In- and outputs Phases

Goal of squad work: ***we'll have annex recipes about how we're fairifying existing datasets (those recipes will be specific to the datasets we're working with). Those will inform generic recipes which will be included in the cookbook to be used for future fairification of new datasets.***