IMI2 Project 802750 - FAIRplus

FAIRification of IMI and EFPIA data

WP5 – Communication and outreach to FAIR data user community

D5.4 KPI Dashboard

Lead contributor	Premysl Velek (1– EMBL (ELIXIR Hub))
	premysl.velek@fairplus-project.eu
Other contributors	Hannah Hurst (1 - EMBL (ELIXIR hub))
	Martin Cook (1 - EMBL (ELIXIR hub))
	Tony Burdett (1 - EMBL)
	Wei Gu (7 - University of Luxembourg)
	Melanie Courtot (1 - EMBL)
	Philippe Rocca-Serra (3 - UOXF)
	David Henderson (21 - Bayer)

Due date	31 July 2020
Delivery date	28 July 2020
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	28 July 2020

Document History

Version	Date	Description
V0.1	27 April 2020	First Draft
V0.2	2 July 2020	Feedback collected and final version agreed on
V0.3	13 July 2020	Final Draft
V1.0	28 July 2020	Final Version

Table of Contents

1. Executive Summary

This report presents the FAIRplus Key Performance Indicator (KPI) Dashboard that tracks the progress of the project activities, the delivery of FAIRplus resources and services and their impact on the wider research community.

The FAIRplus KPI Dashboard enables tracking the progress of the FAIRplus project consortium in the following areas: (1) FAIRification of the datasets from selected IMI projects, (2) Development of the FAIR cookbook[1], (3) Content and usage of the IMI FAIR Data Catalogue[2], and (4) Impact of the FAIR SME and Innovation Forums.

The Dashboard is publicly available through the FAIRplus website[3] and will be updated in regular intervals as new services and resources developed by FAIRplus become available.

2. Methods

The development of the Dashboard involved three different categories: (1) Technical considerations (2) Collection of content and visualisation and (3) Establishment of processes to regularly update the dashboard.

The technical development of the Dashboard was closely linked to the technical maintenance of the FAIRplus website and was carried out by the Project Coordinator (ELIXIR Hub).

Early on in the development, it was decided to use the FAIRplus website to host the Dashboard. Following a desk research to explore and assess different options for displaying and managing the content of the Dashboard, Google charts[4] in combination with Google spreadsheets were selected as the most flexible solution.

The technical architecture of the KPI Dashboard is therefore very light-weight and involves storing the raw KPI data in a Google spreadsheet. The data are subsequently presented in a series of charts which are then published and embedded on a dedicated page on the FAIRplus website.

The set of data to be regularly collected from across the FAIRplus consortium was agreed with each specific WP that owns and produces the data: the FAIRplus Squad teams for progress on data FAIRification, WP2 for progress on the FAIR cookbook and

[1] https://fairplus.github.io/cookbook-dev/intro.html
[2] https://datacatalog.elixir-luxembourg.org
[3] https://fairplus-project.eu/impact/KPI-dashboard
[4] https://support.google.com/docs/topic/1361474?hl=en

the FAIRification recipes, WP3 for data on the usage and impact of the FAIRified datasets, and WP4 on the impact of the FAIRplus networking events (the FAIRplus Innovation and SME Forums).

The process for updating each specific type of data, as well as the exact type of data visualisation was determined with each individual data owner, following the discussion on the collection of data.

The Dashboard itself was developed through an iterative process. A series of mockups and early drafts of the Dashboard allowed for gathering of feedback from each individual data owner. The final version of the Dashboard was shared with the Work Package leaders.

3. Results

The Dashboard was released on 28 July 2020 and is now available at https://fairplus-project.eu/impact/kpi-dashboard. The initial version includes six visualisations presenting four different areas of activities in FAIRplus (FAIRification of IMI datasets, FAIR cookbook, IMI FAIR Data Catalogue, and the FAIRplus Innovation and SME programme).

Number of FAIRified datasets

The first chart shows the proportion of the datasets that have been through all stages of the FAIRplus FAIRification process, out of the total number of IMI datasets selected for FAIRification to date.

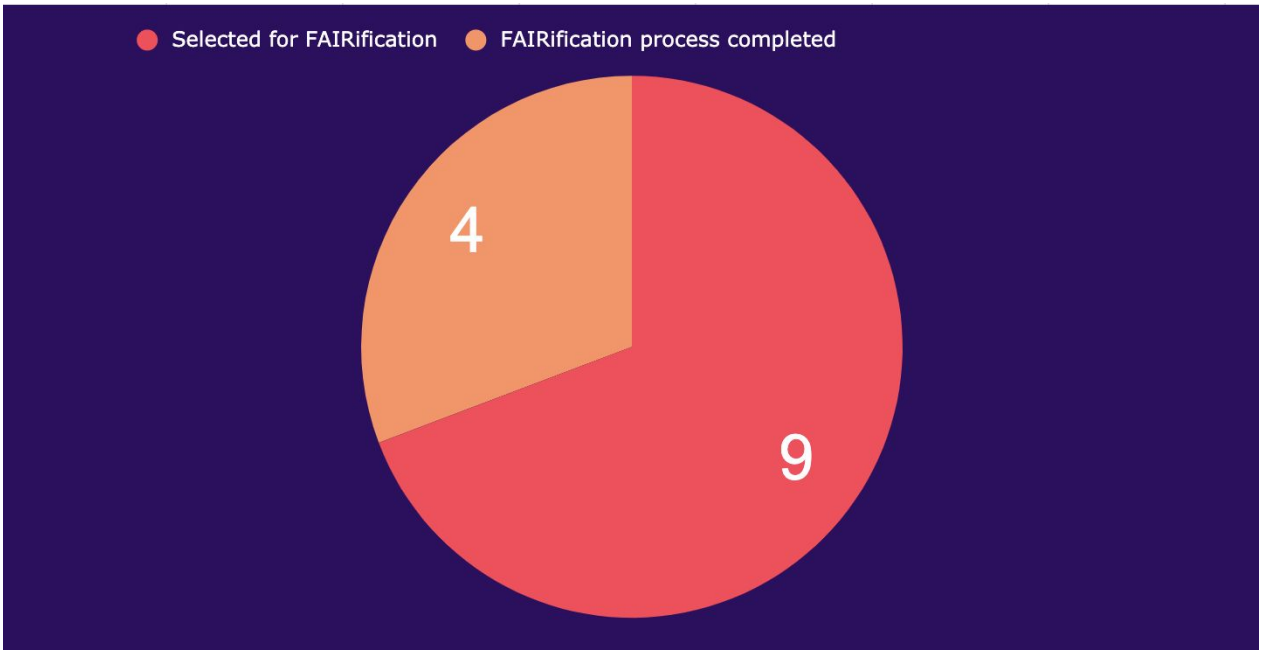

FIgure 1. Number of FAIRified datasets

As the FAIRification of IMI datasets operate in yearly cycles, this chart will be updated once a year, following the start of a new FAIRification cycle.

The next update is scheduled for January 2021.

Number of datasets that have completed each of the FAIRification process five milestones

Chart two provides more details on the FAIRification process, showing how many projects have gone through five successive stages of the process: Selection, Access, Design, Implementation and Dissemination.

It should be noted that the specific datasets from IMI projects selected for FAIRification are representative of the data overall data produced by the individual IMI projects but don't constitute their entire data production. The specific datasets selected for FAIRification were identified in collaboration with the data owners to demonstrate feasibility and value of the FAIRification process.[5]

The FAIRification of the datasets of IMI projects is continuous and runs in three-month cycles. The chart presents cumulative data grouped by year. The current year group (2020) therefore presents preliminary data and will be updated after each cycle.

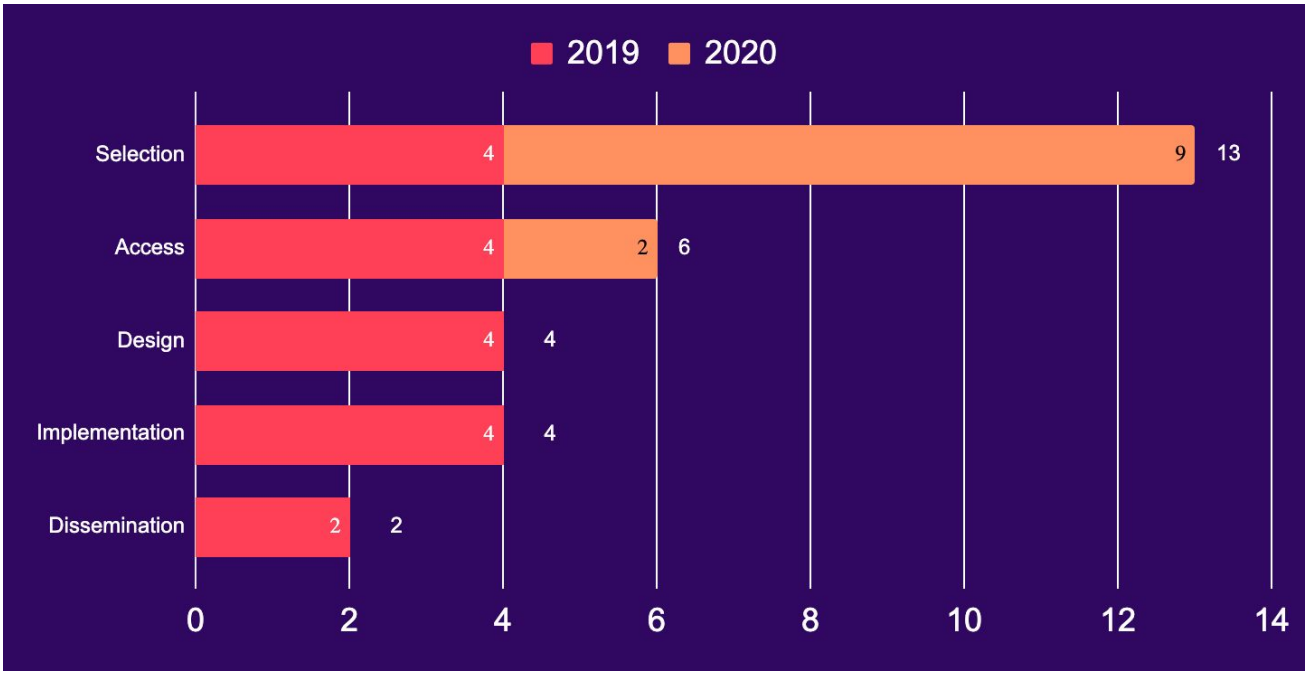

FIgure 2. Number of datasets that have completed each of the FAIRification process milestones

[5] For the description of the selection process see Deliverable D1.1 First 3 data sets from Pilots selected and available, doi: 10.5281/zenodo.3274229

The description of the five stages are provided below:

- **Selection:** Identification and selection of IMI datasets, including assessment of the scientific value of the datasets.
- **Access:** Evaluating and addressing ethical and legal requirements of the dataset FAIRification. Datasets that have completed this stage can be shared with the FAIRplus Squad teams who can then start with the technical part of the FAIRification process.
- **Design:** Technical planning of the FAIRification process. This stage typically includes describing the goals of the FAIRification, deciding on metadata standards or selection of appropriate ontologies.
- **Implementation:** The actual implementation of the FAIRification process carried out by the FAIRplus Squad teams. This stage is completed when the initial FAIRification of the dataset is completed and the FAIRified data are findable and accessible by third parties outside the project consortium.

  The implementation step may only cover certain aspects and only reach an initial maturity level, as defined by the FAIRplus developed `Capability Maturity Model`[6]. In that sense, one has to understand that FAIRness has fractal properties[7]. Some datasets may require or may undergo several rounds of implementation as new capabilities are established. This is also constrained by resources allocated by the requesting projects, EFPIA partners and academic staff availability.
- **Dissemination:** Dissemination of the FAIRified datasets and strengthening its impact. This stage includes developing FAIRification recipes based on the experience gained from the FAIRification or showcasing how the FAIRified datasets have been used.

The data in this chart will be updated quarterly, following the Squad teams face to face meetings. After each meeting, the FAIRplus Squads will assess the progress they have made and update accordingly the number of stages each dataset has completed.

This chart is complemented by a table that provides more information about the individual datasets in each of the five stages. The table also includes links to the websites of the relevant IMI projects, and links to access the FAIRified datasets (for Implementation and Dissemination stages).

[6] https://drive.google.com/file/d/1fKR_2jub8tMN9hZbhUtoNQK3SaPn9ZSX/view
[7] http://blog.thehyve.nl/blog/fair-is-like-a-fractal

KPI dahsboard : IMI projects datasets

Selection	Access	Design	Implementation	Disseminatio
Resolute	Resolute	Resolute	Resolute	Resolute
eTox	eTox	eTox	eTox	eTox
OncoTrack	OncoTrack	OncoTrack	OncoTrack	
ND4BB	ND4BB	ND4BB	ND4BB	
UltraDD	UltraDD			
UBIOPRED	UBIOPRED			
IMIDIA				
RHAPSODY				
EQIPD				
ABIRISK				
EBISC I				
EBiSC2				

IMI projects datasets

FIgure 3. Table listing individual IMI projects and the FAIRification stages of them completed

Participation at FAIRplus Innovation and SME Forums

The fourth chart provides an overview of the participants of the FAIRplus Innovation and SME Forums. It divides the participants into five groups and shows the proportion of participants from each group. The groups are as follows:

1. Academia
2. EFPIA companies
3. Not-for-profit organisations (NGOs)
4. Small and Medium-sized Enterprises (SMEs)
5. Policy makers

It also indicates the proportion of senior managers (directors, principal investigators or senior researchers), who attended the event.

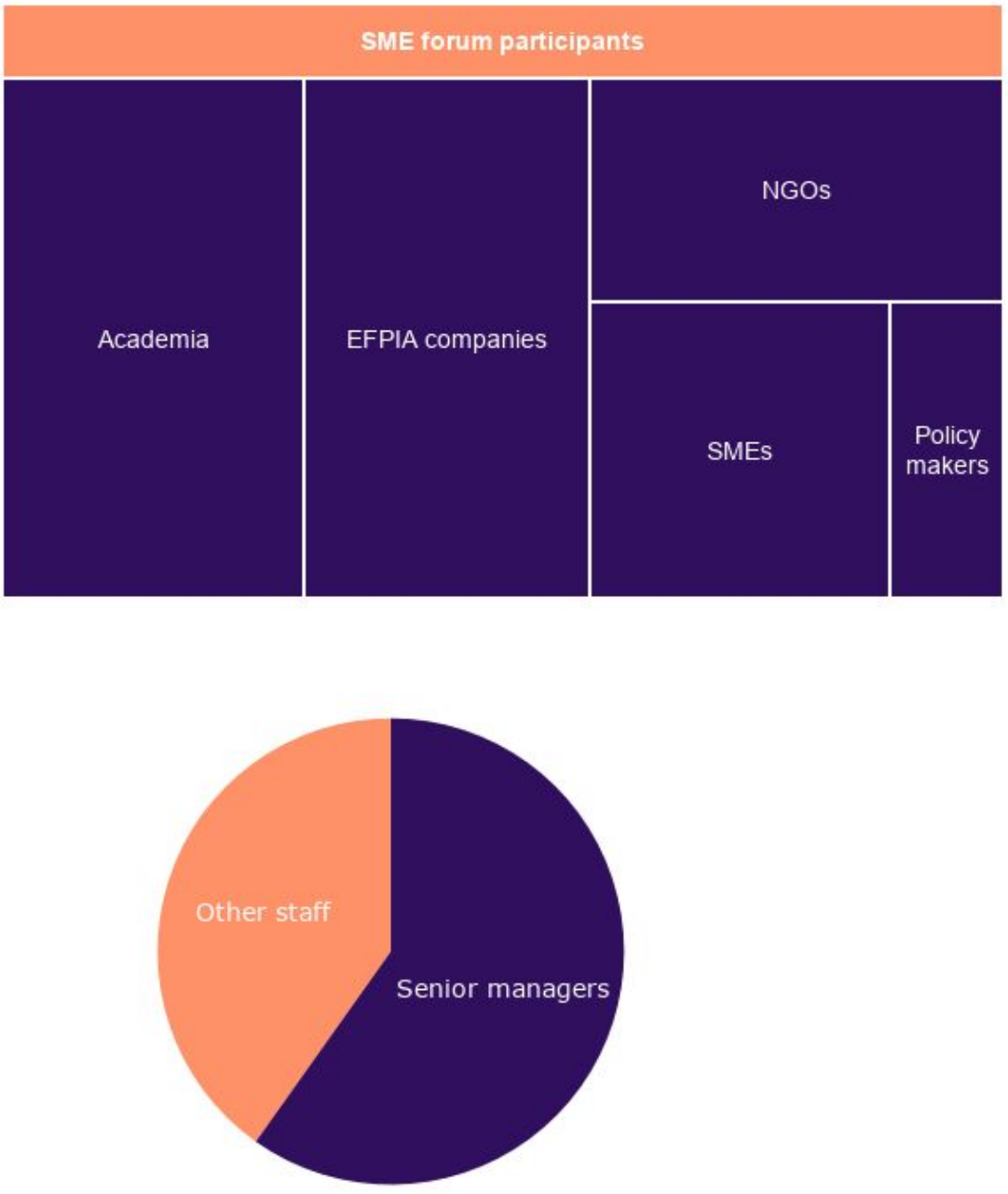

Figure 4. Participants from the FAIRplus Innovation and SME Forum: where they come from

The information is based on the information provided by each participant during the registration of the event. The data are only presented on an aggregate level - no personal information about the participants is available.

As the FAIRplus Innovation and SME Forums are organised once a year, this chart will be updated on a yearly basis. The current version of the chart presents data about the FAIRplus Innovation and SME Forum in January 2020; the next update will take place in January 2021 and will combine data from both 2020 and 2021 events.

Number of FAIRification recipes published in the FAIR cookbook

To present the dissemination of FAIRifiation best practices and experience from the FAIRification of the IMI datasets, the fifth chart shows the number of FAIRification recipes published in the FAIR cookbook.

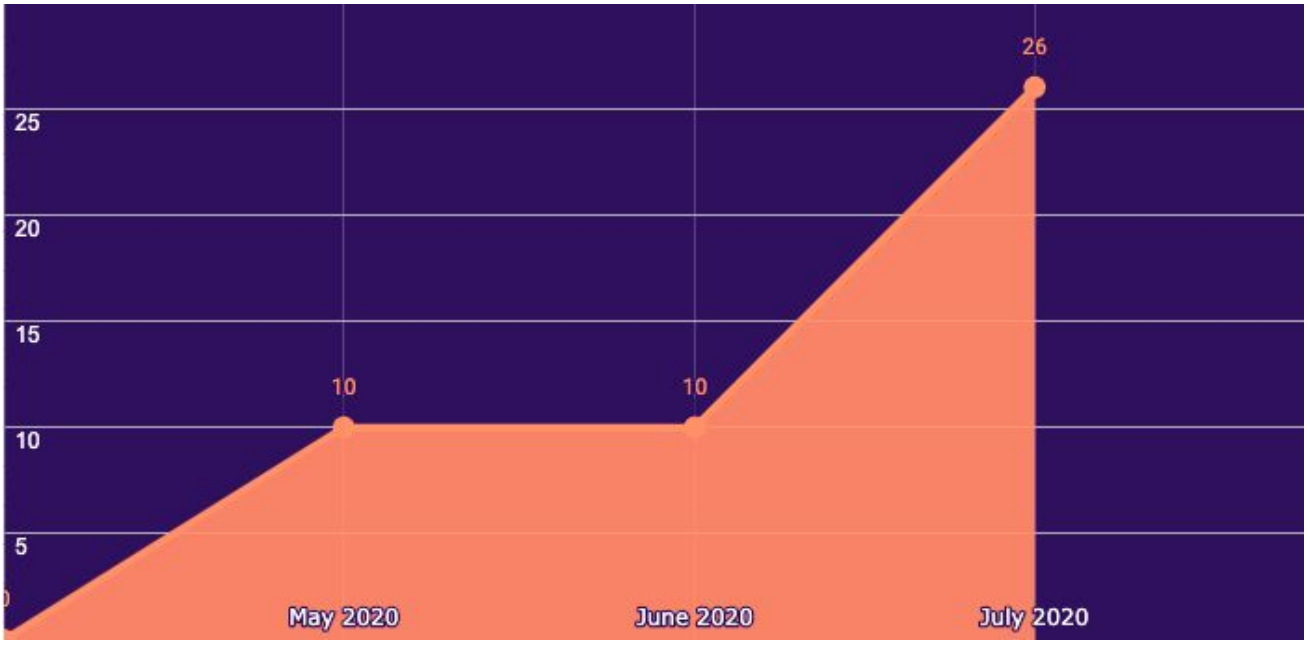

FIgure 5. Number of FAIRification recipes published in the FAIR cookbook

Although - at the time of submission of this report - the FAIR cookbook is still in a beta version, there are already several recipes published. The tracking of the publication of new recipes started in May 2020 and will be updated regularly on a monthly basis.

Number of users accessing the datasets FAIRified by FAIRplus

Datasets that have completed all stages of the FAIRfication process are subsequently made accessible through the IMI FAIR Data Catalogue[8]. To demonstrate the impact of the FAIRification process, the sixth chart in the KPI Dashboard presents the number of visitors that accessed the FAIRified data.

[8] https://datacatalog.elixir-luxembourg.org

FIgure 6. Number of users accessing the datasets FAIRified by FAIRplus

The data presented in the Dashboard come from the web analytics service managed by the University of Luxembourg (Consortium participant no. 7).

It should be pointed out that collecting web traffic relies on collecting personal information from users about what page and how often they visited. The data therefore include only users who gave explicit consent with their data being collected by the University of Luxembourg.

Furthermore, since the IMI FAIR Data Catalogue is a metadata repository, the datasets ,the metadata of which it indexes, are actually hosted by third party repositories (typically national and international domain- or type- specific repositories widely used within particular life science domains).

Therefore, the numbers presented in the chart only reflect the number of accessions to the FAIRified datasets through the Catalogue, not the direct accesses from primary sources. The chart is only indicative of the traffic to the IMI FAIRplus Data Catalogue.

4. Conclusion

With its release, the KPI Dashboard becomes the main tool to track the progress and impact of the FAIRplus project. As the project develops, more metrics and indicators to demonstrate the impact of both FAIRplus and FAIR data management will become available to collect and analyse.

The data to be collected and presented as part of the Dashboard may include the number of IMI project proposals that reference the FAIR cookbook, the IMI Data Catalogue, evidence of re-use of datasets FAIRified by FAIRplus (new analysis, papers, project proposals etc) or citations in scientific literature of (1) FAIRified datasets and (2) FAIRplus resources.

The Dashboard will also track the progress of the FAIRplus Fellowship Programme and its impact.

5. Repository for primary data

All data presented on the KPI Dashboard are stored on the private FAIRplus Google Drive. If you wish to access the raw data visualised in the six charts described in this deliverable report, please contact fairplus-pm@elixir-europe.org.