Check for updates

RESEARCH PAPER

Building Expertise on FAIR Through Evolving Bring Your Own Data (BYOD) Workshops: Describing the Data, Software, and Management-focused Approaches and Their Evolution

César H. Bernabé[1‡], Lieze Thielemans[2‡], Rajaram Kaliyaperumal[1‡], Claudio Carta[3‡], Shuxin Zhang[4,5], Celia W.G. van Gelder[6], Nirupama Benis[4,5], Luiz Olavo Bonino da Silva Santos[1,7], Ronald Cornet[4,5], Bruna dos Santos Vieira[8,9], Nawel Lalout[9,10], Ines Henriques[9], Alberto Cámara Ballesteros[11], Kees Burger[1], Martijn G. Kersloot[4,5,12], Friederike Ehrhart[13], Esther van Enckevort[14], Chris T. Evelo[13,15], Alasdair J. G. Gray[16], Marc Hanauer[17], Kristina Hettne[18], Joep de Ligt[19], Arnaldo Pereira[20], Núria Queralt-Rosinach[1], Erik Schultes[21,22], Domenica Taruscio[3], Andra Waagmeester[23], Mark D. Wilkinson[11], Egon L. Willighagen[13], Mascha Jansen[6], Barend Mons[1,21], Marco Roos[1*], Annika Jacobsen[1‡*]

[1]Biosemantics Group, Department of Human Genetics, Leiden University Medical Centre, Leiden, The Netherlands

[2]NIHR Academic Fellow in General Practice, University College London and The Royal Free NHS Foundation Trust

[3]National Centre for Rare Diseases, Istituto Superiore di Sanità, Rome, Italy

[4]Department of Medical Informatics, Amsterdam UMC location University of Amsterdam, Meibergdreef 9, Amsterdam, The Netherlands

[5]Amsterdam Public Health, Methodology & Digital Health, Amsterdam, The Netherlands

[6]Health-RI, Utrecht, The Netherlands

[7]Semantics, Services and Cybersecurity group, University of Twente, Enschede, The Netherlands

[8]Department of Medical Imaging, Radboud Institute for Health Sciences, Radboud University Medical Center, Nijmegen, The Netherlands

[9]Department of Medical BioSciences, Radboud University Medical Center, Nijmegen, The Netherlands

[10]The Duchenne Parent Project, Veneendaal, The Netherlands

[11]Departamento de Biotecnología-Biología Vegetal, Escuela Técnica Superior de Ingeniería Agronómica, Alimentaria y de Biosistemas, Centro de Biotecnología y Genómica de Plantas. Universidad Politécnica de Madrid (UPM) - Instituto Nacional de Investigacíon y Tecnología Agraria y Alimentaria-CSIC (INIA-CSIC). Campus Montegancedo, Pozuelo de Alarcón (Madrid), Spain.

[12]Castor, Amsterdam, The Netherlands

[13]Department of Bioinformatics - BiGCaT, Maastricht University, Maastricht, The Netherlands

[14]University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

[15]Maastricht Centre for Systems Biology (MaCSBio), Maastricht University, Maastricht, The Netherlands

[16]Department of Computer Science, Heriot-Watt University, Edinburgh, UK

[17]INSERM, US14—Orphanet, Plateforme Maladies Rares, Paris, France

[18]Centre for Digital Scholarship, Leiden University Libraries, The Netherlands

[19]Center for Molecular Medicine, Cancer Genomics Netherlands, Department of Genetics, University Medical Center Utrecht, University Utrecht, The Netherlands

‡ These authors share first authorship

* These authors share last authorship

† Corresponding author: Marco Roos (E-mail: m.roos@lumc.nl), Annika Jacobsen (E-mail: a.jacobsen@lumc.nl).

Building Expertise on FAIR Through Evolving Bring Your Own Data (BYOD) Workshops: Describing the Data, Software, and Management-focused Approaches and Their Evolution

[20]University of Aveiro, DETI/IEETA, Aveiro, Portugal

[21]GO FAIR Foundation, Leiden, The Netherlands

[22]The Leiden Academic Center for Drug Research, Leiden, The Netherlands

[23]Micelio, Ekeren-Antwerp, Belgium

ABSTRACT

Since 2014, "Bring Your Own Data" workshops (BYODs) have been organised to inform people about the process and benefits of making resources Findable, Accessible, Interoperable, and Reusable (FAIR, and the FAIRification process). The BYOD workshops' content and format differ depending on their goal, context, and the background and needs of participants. Data-focused BYODs educate domain experts on how to make their data FAIR to find new answers to research questions. Management-focused BYODs promote the benefits of making data FAIR and instruct project managers and policy-makers on the characteristics of FAIRification projects. Software-focused BYODs gather software developers and experts on FAIR to implement or improve software resources that are used to support FAIRification. Overall, these BYODs intend to foster collaboration between different types of stakeholders involved in data management, curation, and reuse (e.g. domain experts, trainers, developers, data owners, data analysts, FAIR experts). The BYODs also serve as an opportunity to learn what kind of support for FAIRification is needed from different communities and to develop teaching materials based on practical examples and experience. In this paper, we detail the three different structures of the BYODs and describe examples of early BYODs related to plant breeding data, and rare disease registries and biobanks, which have shaped the structure of the workshops. We discuss the latest insights into making BYODs more productive by leveraging our almost ten years of training experience in these workshops, including successes and encountered challenges. Finally, we examine how the participants' feedback has motivated the research on FAIR, including the development of workflows and software.

1. INTRODUCTION

The FAIR data principles address critical factors to make the analysis of multiple sources more efficient by improving their Findability, Accessibility, Interoperability, and Reusability for humans and computers [1]. The process of making data FAIR ("FAIRification"), although partially supported by software (e.g. data transformation tools) and standards (e.g. ontologies), relies on expert knowledge about the data generating

domain (domain experts and data owners) and about FAIR-related aspects such as metadata design, conceptual modelling, licensing definition, and use of identifiers (FAIR experts). Given the high demand to make resources FAIR and a shortage of FAIR expertise, a series of "Bring Your Own Data" workshops (BYODs) have been organised and supported by ELIXIR-EXCELERATE, RD-Connect, and later the European Joint Programme on Rare Diseases (EJP RD) since 2014 to bring together expert knowledge to accelerate the practical FAIRification of resources (e.g. datasets, registry information, ontologies). The bidirectional learning experience between domain and FAIR experts results in making the BYODs a mutually beneficial experience. Attendees (domain experts) receive hands-on guidance in making their data FAIR, while trainers (FAIR experts) gain valuable insights to improve their own training skills, topics, and materials, and develop more effective FAIR support tools, processes, and guidelines.

The first BYOD workshop took place six months after the "Jointly Designing a Data FAIRPORT" workshop [2], which marked the inception of the FAIR principles. Subsequently, the first BYOD for rare disease registries and biobanks, held in November 2014, initiated an annual series of BYODs for this specific domain. Initially, data-focused BYODs emerged from the need to train people on FAIRification and therefore focused on making data resources FAIR. Over time, as the FAIR community matured, it became clear that different types of BYODs were necessary to meet different contexts, needs and backgrounds of attendees. Consequently, two additional types of BYOD structures were designed: management- and software-focused BYODs. The former aims at informing managers and policy-makers on the added benefits of FAIR and the requirements for FAIRification. The latter focus on developing software tools that support the process of making data FAIR, or tools and standards that enhance the FAIR level of resources.

The remainder of this paper is organised into four sections. The next section describes the three types of BYODs. The section "the evolution of BYODs" lists the BYODs organised since 2014 and reports on the first BYOD workshop and the series of BYODs for rare disease registries, with emphasis on how these workshops have led to the improvement of the content and didactical aspects of the BYODs. Then, we discuss the impact of the BYODs on the FAIR community and on the expert domains (i.e. plant breeding, and rare disease registries and biobanks). In this paper, we mention different types of experts. For clarity, we refer to "X expert" as an expert specialised in a certain tool, standard, or knowledge. For instance, "FAIR expert" refers to an expert with knowledge and experience in FAIR.

2. THE THREE TYPES OF BYOD STRUCTURE

Despite catering for different contexts and types of participants, all BYODs share the same overarching goal of fostering expertise in FAIR-related topics while cultivating community confidence in the benefits of having FAIR resources. Table 1 summarises the main aspects of the three different BYOD structures, which are further described in the subsections that follow.

In addition to the two- or three-day duration of the BYODs, some also contain preparatory and follow-up phases for attendees and trainers. Attendees are invited to participate in introductory webinars to prepare for the workshop, and post-BYOD follow-up meetings, for support on subsequent activities. The introductory

Building Expertise on FAIR Through Evolving Bring Your Own Data (BYOD) Workshops:
Describing the Data, Software, and Management-focused Approaches and Their Evolution

Table 1. An overview of the main characteristics of the BYODs structure. The first row describes the technical knowledge (e.g. on data transformation algorithms) required from the attendees. The second row lists the learning formats that can be used in the BYODs. The third and fourth rows depict the main goals and tasks of the BYODs, respectively. The fifth and sixth rows show the expected backgrounds of the attendees and trainers of the BYODs, respectively. Note that project managers refer to those who manage projects that can benefit from FAIR(ification), such as data collection projects, while standard specialists refer to experts in any standard that enables FAIR (e.g. FAIR Data Point [3]). The seventh row lists the standards and tools commonly used in the FAIRification, as well as development tasks or lectures given during the BYODs. The last row describes the expected outputs for each BYOD format.

	Data	Management	Software
Required technical knowledge for attendees	Intermediate	Low	High
Learning Format	Knowledge exchange Hackathon-Lectures	Knowledge exchange Lectures	Knowledge exchange Hackathon
Main goals	Answer research questions using FAIR data Make resources FAIR Train domain experts on FAIR(ification)	Inform participants about the added benefits of FAIR Inform managers on the characteristics of FAIRification projects	Develop software that support enabling FAIR on resources (e.g. FAIR Data Point) Develop resources that support the FAIRification process (e.g. data transformation tools)
Main tasks	Semantic modelling of (meta) data Hosting and querying of FAIR data	Hands-on scenarios simulating the benefit of FAIR resources Plenary sessions to discuss the characteristics of FAIRification projects	Solution brainstorming Solution implementation Testing of implemented solution
Profile of attendees	Researchers Domain experts	Project managers Registry managers Patient representatives Policy makers Funders	Researchers on FAIR Software Developers FAIR data stewards Domain experts
Profile of trainers	Experts in FAIR FAIR data stewards Ontologists Standards specialists	FAIRification project managers FAIR data stewards Decision/policy makers with knowledge on FAIR	Researchers on FAIR Developers FAIR data stewards Domain experts
Commonly used tools and standards	Ontologies and metadata models FAIR Data Point Domain specific standards (e.g. CDE Semantic Model [ref]) FAIR enabling standards (e.g. DCAT [40])	FAIRification workflows Collaborative brainstorming tools (e.g. mind maps, black boards)	Collaborative brainstorming tools Software development resources (e.g. programming languages such as Python)
Expected outputs	FAIR resource Answer to research question(s)	Audience knowledgeable about the characteristics of FAIRification and the added benefits of FAIR	FAIR enabling resource FAIRification supporting resource

webinars aim to familiarise attendees with FAIR and initial FAIRification needs (e.g. identification of goals and required domain expertise). In the post-BYOD follow-up meetings, attendees are advised on other FAIRification challenges that might appear after the BYOD. The preparatory and follow-up phases provide opportunities for trainers to prepare and evaluate the workshop agenda, training materials, and methods of instruction. Follow-up phases are also used by trainers to plan for improvements in future editions of the workshop in response to feedback from participants.

2.1 Data-focused BYOD Workshops

During the data-focused BYODs, attendees are divided into groups, with at least one trainer allocated per group. Each group can use their own data or request synthetic data. Collaboratively, the groups transform their data into FAIR data by following a step-by-step FAIRification process.

In most workshops, we have followed a FAIRification workflow where data is made FAIR retrospectively collection - *post hoc* and semi-automatically (see Figure 1). This workflow was developed based on emerging FAIRification steps observed in early BYODs. It should be noted, however, that previous BYODs may have deviated from this structure while evolving towards the current format. Additionally, more recent workshops focused on making data FAIR by design (automatically during data collection - *de novo*) (e.g. [5]). Figure 1, which is adapted from [4] and [5], illustrates the FAIRification workflow used in recent BYODs. The workflow is divided into three phases: 1) pre-FAIRification, 2) FAIRification, and 3) post-FAIRification, which are each subdivided into clear steps. These hands-on phases are usually accompanied by lectures about FAIR related topics and plenary sessions where participants can share their experiences with FAIRification, including challenges and success cases (as listed in Table 1).

The Pre-FAIRification phase is composed of three steps. In step 1, to identify FAIRification objectives and (meta)data elements to be collected, the groups define driving objectives and research question(s) focusing on using their sample data in combination with other FAIR data. Next, following steps 2 and 3, the groups
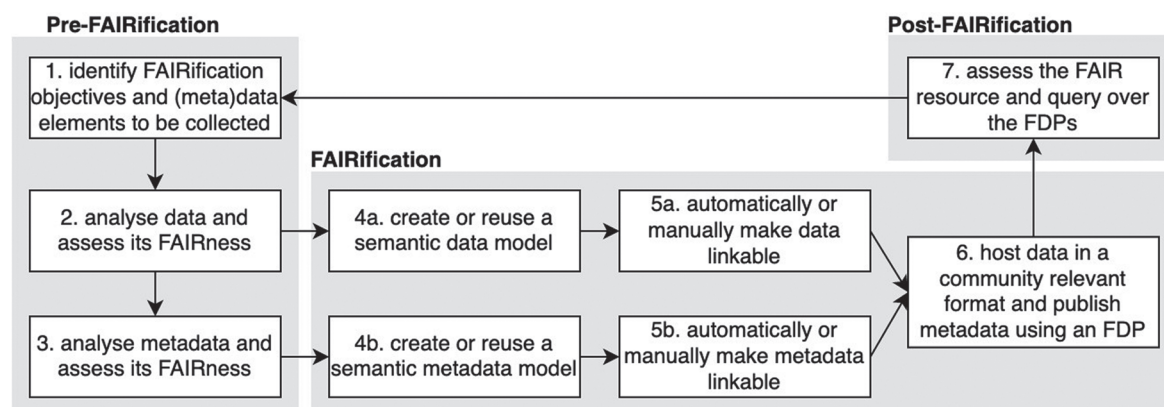

Figure 1. Illustration of the FAIRification workflow used in the data-focused BYODs, modified from [4] and [5].

closely investigate the representation (syntax) and meaning (semantics) of their data and the metadata (i.e. description of data). In our experience, metadata such as the (meta)data's licence and provenance information is often not available a priori and therefore needs to be gathered during the BYOD workshop. Finally, before doing any actual FAIRification, the FAIR status of the data is assessed by using tools such as the FAIR Evaluation Services [6] (step 3) (see [7] for other FAIR assessment services).

In the FAIRification phase, the groups create or reuse a conceptual model to describe the data elements and their relationship (step 4a), and a metadata model to provide information about the data (step 4b). These conceptual models must contain, at a minimum, the data elements required to answer their driving research question(s). These conceptual models are semantically enriched by binding the models' concepts to terms from reference ontologies. In step 5a, the data is made machine-readable (i.e. in a format that can be processed by a computer) by using the semantic data model and existing tooling to generate an ontologised version of the data manually (e.g. FAIRifier [8]) or automatically (e.g. Castor EDC [9], MOLGENIS [10]). The metadata is also made machine-readable (step 5b) by using metadata standards (e.g. Data Catalogue Vocabulary (DCAT) [11]). Finally, the machine-readable metadata is made available using the FAIR Data Point (FDP) (step 6) [3] and the machine-readable data is hosted using a community relevant file format (e.g. RDF [12]).

In the post-FAIRification phase, the driving research question(s) defined in step 1 are answered using the newly created FAIR data (step 7). Here, the FAIR status of the data resource is reassessed and compared to the assessment done in the pre-FAIRification phase to verify if the improvement of the FAIR level of the data meets the goals previously defined.

To illustrate, in the pre-FAIRification phase, a group defines "finding new treatment candidates for untreated rare disease patients" as a driving goal, and reuses data from different rare diseases registries to achieve this goal. The data includes information on diagnosis, symptoms and treatments, and metadata includes information such as the (meta)data's licence (e.g. CC BY-NC 4.0 [13]) and provenance (e.g. from which registry the data originates). In the FAIRification phase, the group adopts the Common Data Elements (CDE) Semantic Model [14] (step 4a) as the semantic data model, and the EJP RD metadata model [15] as the semantic metadata model (step 4b). Reusing the ontologies from the models adopted in steps 4a and 4b supports making the data and metadata linkable (steps 5a and 5b). Finally, the newly linked (meta)data is hosted and published using an FDP (step 6). During the post-FAIRification phase, the group leverages the FAIR data they have created to address their research question by writing federated queries. For instance, they may query their FAIR data with other public resources to identify treatment candidates for patients with similar symptoms.

2.2 Management-focused BYOD Workshops

The management-focused BYOD workshops are geared towards informing registry and project managers, patient representatives, and decision-makers about the characteristics of FAIRification, including the associated costs, time, expertise, and effort required. The need for this type of BYOD emerged due to the

Building Expertise on FAIR Through Evolving Bring Your Own Data (BYOD) Workshops:
Describing the Data, Software, and Management-focused Approaches and Their Evolution

growing adoption of FAIR in various institutions, which has required personnel in high-level positions to become familiar with the benefits and prerequisites of data FAIRification. As a result, these workshops place less emphasis on technical work and more on general considerations of FAIRification. The management-focused BYOD is divided into three phases: (i) understanding the problem of not having FAIR data, (ii) acquiring knowledge about FAIR and FAIRification, and (iii) training on FAIRification project management.

The first phase of a management-focused BYOD is executed in an interactive manner, typically through the use of simulated case scenarios that recreate the challenge of dealing with incomprehensible and non-interoperable data. To highlight the importance of FAIR data, attendees are tasked with challenges that require connecting data from multiple sources, while being presented with non-standardised and multilingually annotated data in different formats, making the task more difficult to accomplish.

In the second phase, attendees learn about the benefits of FAIR and the main steps of FAIRification. The learned benefits aim to address the challenges identified in phase one. Plenary and hands-on sessions provide practical experience in FAIRification related tasks, including conceptual modelling, making metadata findable, and using ontologies and FAIR compliant Electronic Data Capture (EDC) systems. This phase is typically concluded by revisiting the mock case from the first phase, but this time using FAIR data, thus demonstrating how the previously identified challenges can now be solved more easily and efficiently.

In the third phase, participants discuss the implications (e.g. budget, time, required expertise and infrastructure) of FAIR for project managers and policymakers. After the plenary sessions, attendees have a hands-on session on how to create their FAIRification team.

A real-world example of this structure can be visualised on the agenda of recent management-focused BYODs organised for rare disease registries and biobanks (e.g. [16]). For instance, in the one held online in 2022 [16], attendees experienced the problems of not having FAIR data through a digital game where they had to find treatments for new patients in different datasets organised in a non-standardised manner (e.g. using synonyms for equivalent concepts) and presented in several languages (e.g. Mandarin, Dutch, and Spanish). Thereafter, lectures and discussion sessions on topics such as FAIRification steps, conceptual modelling, ontologies, and querying informed the attendees about FAIR-related aspects. On the second day, the attendees played the same digital game, only this time with FAIR data, which allowed them to accomplish the goal of finding treatment in distributed datasets. After lectures and discussions on the benefits of FAIR, participants exchanged experiences about the implications of data FAIRification for registry managers.

2.3 Software-focused BYOD Workshops

The main goal of software-focused BYOD workshops is to create software that supports FAIRification, or software and standards that increase the FAIR level of resources, as shown in Table 1. Participants of software-focused BYODs include researchers working on FAIR-related projects, FAIR data stewards [17], developers and, in certain cases, domain experts. In this type of workshop, trainers and attendees come from similar backgrounds, working together to exchange knowledge while tackling the same goal. This type of BYOD is organised in a hackathon setting with five phases:

Building Expertise on FAIR Through Evolving Bring Your Own Data (BYOD) Workshops:
Describing the Data, Software, and Management-focused Approaches and Their Evolution

1. Understanding the problem: participants discuss the need or problem to be solved (e.g. facilitating metadata publication)
2. Proposing solutions: participants are invited to brainstorm solutions (e.g. using brainstorming tools such as mind maps) to the problem described in the previous phase (e.g. developing software to support the creation and publication of FAIR metadata)
3. Prioritising tasks: the implementation tasks are ordered by importance, and then selected for implementation (e.g. developing a proof-of-concept software that creates machine-readable metadata from an Excel sheet and publishes it via an FDP)
4. Coding and experimenting: the prioritised tasks are implemented, and the resulting implementation is tested (e.g. implementing and testing the proof-of-concept software with mock data).
5. Reporting: the implementation is reported and published (e.g. a paper or website documenting the script developed during the hackathon)

Most software-focused BYODs are typically structured around iterative cycles. After a set period of time, participants convene to report on their group status and the findings from their tasks. They can then decide to switch or merge groups, get advice from others and/or continue their tasks. As a result, the agenda of the software-focused BYOD is adaptable to the requirements that emerge during the workshop. At the end of the BYOD, conclusions on which solutions to follow up are made. Adaptations of current tools, prototypes, proof-of-concept implementations, or architectural designs are examples of outcomes of software-focused BYODs.

The "hackathon to make MOLGENIS FAIR" [18], which took place in 2016, is a real-world example of a software-focused BYOD. During this event, software developers and FAIR experts worked collaboratively to create a proof-of-concept of making MOLGENIS FAIR. MOLGENIS [10] is an open-source data platform for the management of scientific data. By the end of the hackathon, the team had implemented an application programming interface (API) to publish datasets in MOLGENIS as FDPs.

3. THE EVOLUTION OF BYODS

Table 2 highlights the BYODs held from 2014 to 2023. For context, the table includes the date of the "Jointly Designing a Data FAIRPORT" Workshop, where the FAIR principles were initially conceived, and the publication of the paper describing the FAIR principles [1]. A list with more detailed information on the workshops is available as supporting information[①].

BYODs listed by Table 2 include the first workshop on genetic biodiversity [19] and the Bring Your Own Rett Syndrome Data workshop [20]. The former was a data-focused BYOD where participants worked on linking different datasets (e.g., the Centre for Genetic Resources (CGN) tomato collection and phenotypic observations, and variants from the 150 tomato genome re-sequencing project [21]) that were then queried as combined data. The latter focused on producing FAIR nanopublications[②] about Rett Syndrome, the results

---

[①] https://doi.org/10.5281/zenodo.8155154

Building Expertise on FAIR Through Evolving Bring Your Own Data (BYOD) Workshops:
Describing the Data, Software, and Management-focused Approaches and Their Evolution

Table 2. Overview of 'Bring Your Own Data' (BYOD) workshops organised from 2014 to September 2023.

Title	Date and Location	Focus
FAIR Principles Idealisation Jointly Designing a Data FAIRPORT	13-16 Jan 2014-Leiden, the Netherlands	-
The first BYOD workshop	24-25 Jun 2014-Leiden, the Netherlands	Data
The first Bring Your Own Data (BYOD) Workshop to Link Rare Disease Registries-First RD BYOD Workshop	26-27 Nov 2014-Rome, Italy	Data
The first "green genetics" BYOD	21-22 Jan 2015-Wageningen, the Netherlands	Data
The Bring Your Own Template (BYOT) workshop	20 Feb 2015-Utrecht, the Netherlands	Data
The second RD BYOD Workshop	24-25 Sep 2015-Rome, Italy	Management
FAIR Principles Paper Published	15 Mar 2016	-
The third RD BYOD Workshop	29-30 Sep 2016-Rome, Italy	Data
The FAIR Data Hackathon	19-20 Oct 2016-Utrecht, the Netherlands	Software
The Software Solution Provider BYOD	25-27 Oct 2016-Leiden, the Netherlands	Software
The Bring Your Own Rett Syndrome Data workshop	1-3 Nov 2016-Maastricht, the Netherlands	Data
How to Make Data FAIR for Open Science	15-19 May 2017-Leiden, the Netherlands	Data and management
The plant phenotype BYOD and hackathon	30 May-1 Jun 2017-Ghent, Belgium	Software
The cancer genomics BYOD	6-8 Jun 2017-Utrecht, the Netherlands	Data and software
The fourth RD BYOD workshop	21-22 Sep 2017-Rome, Italy	Data
The DSM BYOD workshop	25-26 Sep 2017-Delft, the Netherlands	Data
The fifth RD BYOD workshop	13-14 Sep 2018-Rome, Italy	Data and management
The RIKILT/WUR BYOD workshop	22 Nov 2018-Wageningen, the Netherlands	Data
BYOD FAIRification workshop at Leiden University Library	18 Jun-2019-Leiden, the Netherlands	Data
The sixth RD BYOD workshop	26-27 Sep 2019-Rome, Italy	Data and management
The seventh RD BYOD workshop	1-2 Oct 2020-Online	Management
The eighth RD BYOD workshop	30 Sep-1 Oct 2021-Online	Management
The ninth RD BYOD workshop	29-30 Sep 2022-Online	Management
The World Duchenne Organization's FAIR Training Program	7-9 March 2023-Online	Management
The tenth RD BYOD workshop	28-29 Sep 2023-Rome, Italy	Management

of which led to an ELIXIR implementation study on the interoperability of molecular data in rare diseases (MolData2) [22] and contributed to the development of the cross-omics data analysis work package of the EJP RD.

---

[2] Nanopublications are defined as the smallest publishable unit of facts with full information where the knowledge comes from.

All BYODs have played an important role in iteratively improving the structure of subsequent ones, as well as in facilitating the adoption and research on FAIR. As representative examples, the following subsections describe the inaugural BYOD-the Human Protein Atlas and MycoBase BYOD-and the series of BYOD workshops focused on linking rare disease registries and biobanks. The BYODs evolved based on the trainers' perception during the workshops and based on informal feedback from attendees.

3.1 The First BYOD Workshop: Human Protein Atlas and MycoBase

The first data-focused BYOD workshop was held in Leiden, the Netherlands, on the 24th and 25th of June, 2014 [23, 24]. It was organised by a group of researchers from across Europe and sponsored by the Dutch Techcentre for Life Sciences (DTL) [25], and ELIXIR [26]. The BYOD, which focused on making data interoperable, brought together data owners from the Human Protein Atlas [27] and MycoBase [28] with Linked Data experts. It is important to note that the FAIR principles and, consequently, the concept of "FAIR" data were still under development by then. Therefore, this BYOD focused on creating "Linked Data", which is a suggested step towards having FAIR data.

The data owners needed to be familiar with their current internal data management structures, i.e. the database schema and data pipelines for creating and displaying entries. The Linked Data experts had a variety of backgrounds such as semantic web services [29] and integration platforms [30]. The main aim was to develop sample Linked Data to demonstrate the added value of interoperable data for facilitating answering research questions by reusing information from multiple resources.

The BYOD event started with a plenary training session providing an overview of Linked Data. Then, the attendees were split into two working groups, each of which aimed to develop a proof of concept centred around one of the data resources, driven by their own research questions. The Human Protein Atlas group focused on developing a subset of Linked Data from the Human Protein Atlas. This was then connected to WikiPathways data [31]. The MycoBase group linked their data with the content of ChEMBL through the Open PHACTS API [32, 33]. The Human Protein Atlas developers have used the experience of the event to develop their own RDF data release, heavily reusing the ontological model of neXtProt [34].

In this BYOD, it became clear there was a need to include preparation and follow-up meetings in the agenda of subsequent workshops. The experience gained by organisers provided insights for planning pre-BYOD training about the FAIR principles and for organising post-BYOD supporting sessions. Additionally, it highlighted the importance of publishing training materials that could be used at other BYOD events, promoting knowledge sharing and dissemination.

3.2 A Series of Annually Recurring BYODs: Rare Diseases Registries

Making rare disease resources interoperable and, thereby, preparing them for multi-source analysis is crucial since rare diseases occur at low frequency. In Europe, a disease is considered rare when it affects less than 5 in 10,000 individuals [35]. Ensuring the interoperability of rare diseases data is important because non-integrated data would likely be insufficient to support research or improvements in outpatient

Building Expertise on FAIR Through Evolving Bring Your Own Data (BYOD) Workshops:
Describing the Data, Software, and Management-focused Approaches and Their Evolution

care. Therefore, each local data resource is of relatively limited value on its own, but may be highly valuable in combination with other data.

The first BYOD for rare disease registries and biobanks® (RD-BYOD) was held in Rome, Italy, at Istituto Superiore di Sanità on the 26th and 27th of November, 2014. The RD-BYOD was attended mainly by RD-Connect partners [36], including rare disease data owners and software engineers with Linked Data expertise. The main focus was to train data owners in making rare disease patient registries and biobanks interoperable, while also identifying tools to be developed.

With support of the BYODs, the rare disease community quickly acknowledged the importance of data interoperability, and later the FAIR principles. In 2017, the International Rare Disease Research Consortium (IRDiRC) declared the FAIR guiding principles as a 'recognised resource' to "accelerate the pace of translating discoveries into clinical applications" [37]. Since 2019, the series of annually recurring RD-BYODs has been an intrinsic part of the annual summer school on Rare Disease Registries. Editions of the course have been approved by the International Conference on Rare Diseases and Orphan Drugs (ICORD) [38].

Over the years, the RD-BYODs evolved to alleviate the steep learning curve of FAIRification. For instance, trainers were instructed to avoid very technical terms that could confuse beginners or participants with different expertise. Additionally, the RD-BYOD has evolved in response to feedback from participants and advancements in FAIR procedures and technologies. For example, training on FAIRification project management for registry managers was added in 2016 and expanded in subsequent editions of the workshop. As a result, from 2017, priority was given to attendees who were involved in or actively planned to establish a rare disease registry, primarily within a European Reference Network (ERN) [39], shifting the focus from a data-focused to a management-focused structure.

The RD-BYOD has also been used to experiment with, get feedback on and disseminate the technical developments that support the rare diseases community. It also informs registry managers about the available tools and standards. Recent RD-BYODs have been adapted to reflect practical aspects of the rare diseases domain, such as including topics to address needs from patient organisations and ERNs. As an example, the EJP RD ontological model for "Common Data Elements", its supporting tool [14], and the EJP RD ontological metadata model for rare disease patient registries, biobanks and catalogues [15] were presented to participants in the latest editions, together with hands-on sessions for demonstration. The experience acquired by RD-BYOD trainers has been embedded in guidance resources, such as a guide for data stewards to make European rare disease patient registries FAIR [40].

4. DISCUSSION

The BYODs have benefited attendees and trainers in many ways. For trainers, the workshops have created a collaborative environment where the FAIR community gains new insights from the attendees while helping

---

® For the sake of readability, the "BYOD for rare disease registries and biobanks" is referred to as "RD-BYOD" in this subsection.

them deal with their FAIR(ification)-related needs. For example, researchers on FAIR use the open and flexible BYOD environment to test FAIR-related tools with attendees. Similarly, feedback and questions raised during BYODs have supported research on FAIR and FAIRification methods. For instance, research on goal-based FAIRification planning methods [41], assessment of RDF data [42], large-scale implementation of FAIR principles [43] and quality of modelling [44] has benefited from experience from recent BYODs.

Furthermore, lessons learned from success and shortcomings of BYODs provide guidance on future research paths. To illustrate, the pre- and post-BYOD activities have underscored the iterative nature of FAIRification, where the target resource is initially addressed and then expanded by subsequent FAIRification efforts. For example, in the first FAIRification iteration, a subset of data concepts within a given dataset may be addressed, with subsequent iterations expanding the scope to encompass a larger set of concepts. Other challenges, such as solving the communication gap due to the interdisciplinary nature of FAIR and the diverse expertise of BYOD attendees, highlight the need for further research on such topics. Additionally, the difficulty in reaching consensus during conceptual modelling tasks [45], which are crucial in FAIRification [43], is another obstacle frequently encountered in BYODs.

The three different BYOD formats described in this paper are intended to guide others in organising their own BYODs. These formats can be freely adapted by any community to suit their own learning goals, needs and constraints. We suggest that BYODs are organised with a multidisciplinary training group, including at least a FAIR expert, a conceptual modelling expert and an expert in the domain of the resource to be made FAIR.

For attendees, the workshops have aided the participating community by fostering the convergence of standards and tools. In this way, BYODs have become a valuable resource for advancing FAIR data practises. The BYODs' structure has inspired various FAIR training activities and courses, some of which are already offered by universities, research institutes, or consortia as part of their research data management programmes (e.g. [46–48]). The Metadata for Machines workshop (M4M) [49] and the Three-Point FAIRification Framework (3PFF) [50] are also examples of training frameworks that were inspired by the BYODs. Similarly, other FAIRification workflows and frameworks have embedded knowledge acquired by researchers who participated in the BYODs (as trainers or attendees). Examples of these include the generic workflow for the Data FAIRification process [4], the de novo FAIRification process of a registry for vascular anomalies [5], the FAIR in action framework for guiding FAIRification [52] and the FAIR Hourglass model for FAIRification and FAIR orchestration [46].

It is also noticeable that BYODs reflect the maturing of the FAIR community. While early BYODs used prototype tools designed to handle specific FAIRification tasks, recent BYODs have introduced more comprehensive tools that can cover a greater part of FAIRification. For instance, while the first BYODs for rare diseases reused generic tools for converting small datasets to linked data, recent ones introduced software systems that can automatically make data FAIR upon collection (e.g. Castor, MOLGENIS). Moreover, recent BYODs were capable of presenting more complex real-world FAIRification cases that lead to new insights and facilitated the retrospective FAIRification of a patient-led registry (e.g. The Duchenne Data Platform [53]).

Data Intelligence Just Accepted MS.
https://doi.org/10.1162/dint_a_00236

Building Expertise on FAIR Through Evolving Bring Your Own Data (BYOD) Workshops:
Describing the Data, Software, and Management-focused Approaches and Their Evolution

For future training activities, we recommend combining different types of BYODs to tackle various tasks needed at different stages of FAIRification projects. Practically, a FAIRification project starts by creating a homogeneous basic knowledge about FAIR among all people involved, thus making the commitment and investment efforts clear to the whole FAIRification team. This can be done with management-focused BYODs. After the FAIRification project has been set up and its objectives have been identified, it is necessary to gather sufficient technical expertise, which can be supported by organising data- and software-focused BYODs.

For future BYODs, we plan to explicitly align the contents with the knowledge units mentioned in Appendix E of the FAIRsFAIR Teaching and Training Handbook [54]. Furthermore, we aspire to make our teaching materials themselves FAIR, in order to contribute to overcoming the shortage of FAIR expertise and to continue to keep learning as instructors. We suggest that different communities share training materials and lessons learned, so that BYODs continue to evolve as a whole. We have also observed that trainers, who are usually FAIR enthusiasts, are willing to support other groups in organising their own BYODs, for example by attending certain BYOD sessions as invited speakers or by giving advice on organising the BYODs.

Finally, we note that BYODs should not be equated with FAIRification projects, as their primary emphasis is on participants rather than the output. Nevertheless, BYODs can act as a catalyst for such projects, for example by providing a launch pad for dedicated teams to continue the FAIRification process initiated in a BYOD.

5. CONCLUSION

Initially, BYODs aimed at making data interoperable by using available Linked Data technologies. Since their inception, BYODs have evolved and provided a collaborative space to develop FAIRification tools and more robust technologies. Additionally, BYOD workshops have become an important means of exchanging views and knowledge on FAIRification, and on informing researchers and managers on the benefits of FAIR.

Experience has shown that FAIR implementations are an effective approach to enable multi-source analysis, and the BYODs are a valuable asset in promoting the adoption of the FAIR principles in various domains. We will, therefore, continue to organise BYODs to accelerate the adoption and promotion of the FAIR principles.

ACKNOWLEDGMENTS

Building Expertise on FAIR Through Evolving Bring Your Own Data (BYOD) Workshops:
Describing the Data, Software, and Management-focused Approaches and Their Evolution

AUTHOR CONTRIBUTIONS

The work presented in the manuscript is a result of many years of experience by all authors. C. H. Bernabé and L. Thielemans are the lead in writing the manuscript. All authors contributed to the writing and provided critical feedback to help shape the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

REFERENCES

[1]   Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., et al.: The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data 3, 1–9 (2016)

[2]   Lusher, S., Mons, B.: Jointly designing a data FAIRPORT (2014). Lorentz Center at Snellius. Available at: https://www.lorentzcenter.nl/jointly-designing-a-data-fairport.html

[3]   da Silva Santos, L.O.B., Burger, K., Kaliyaperumal, R., Wilkinson, M.D.: FAIR Data Point: A FAIR-oriented approach for metadata publication. Data Intelligence, p. 1–21 (2022)

[4]   Jacobsen, A., Kaliyaperumal, R., da Silva Santos, L.O.B., et al.: A generic workflow for the data FAIRification process. Data Intelligence 2(1-2), 56–65 (2020)

[5]   Groenen, K.H., Jacobsen, A., Kersloot, M.G., et al.: The de novo FAIRification process of a registry for vascular anomalies. Orphanet Journal of Rare Diseases 16(1), 1–10 (2021)

[6]   Wilkinson, M.D., Batista, D.: FAIR Evaluation Services (2023). Available at: https://fairsharing.github.io/FAIR-Evaluator-FrontEnd/

[7]   FAIRSharing Team and Data Readiness Group. FAIRAssist (2023). Available at: https://fairassist.org/#!/

[8]   FAIR Data Team. FAIRifier (2023). Available at: https://github.com/FAIRDataTeam/FAIRifier

[9]   Kersloot, M.G., Jacobsen, A., Groenen, K.H., et al.: De-novo FAIRification via an Electronic Data Capture system by automated transformation of filled electronic Case Report Forms into machine-readable data. Journal of Biomedical Informatics 122, 103897 (2021)

[10]   van der Velde, K.J., Imhann, F., Charbon, B., et al.: MOLGENIS research: Advanced bioinformatics data software for non-bioinformaticians. Bioinformatics 35(6), 1076–1078 (2019)

[11]   W3C. Data Catalog Vocabulary (DCAT) - Version 2. World Wide Web Consortium (W3C) (2019). Available at: https://www.w3.org/TR/vocab-dcat-2/

[12] (W3C) WWWC. Resource Description Framework (RDF); Accessed 2023. Available at: https://www.w3.org/RDF/

[13] Creative Commons. Creative Commons Attribution-NonCommercial 4.0 International License (2023). Available at: https://creativecommons.org/licenses/by-nc/4.0/

[14] Kaliyaperumal, R., Wilkinson, M.D., Moreno, P.A., et al.: Semantic modelling of common data elements for rare disease registries, and a prototype workflow for their deployment over registry data. Journal of Biomedical Semantics 13(1), 1–16 (2022)

[15] EJP-RD. Metadata for EJP rare disease patient registries, biobanks and catalogs (2023). Available at: https://github.com/ejp-rd-vp/resource-metadata-schema

[16] EJP-RD Consortium. International Summer School on Rare Disease Registries and FAIRification of Data (2023). Available at: https://www.ejprarediseases.org/event/international-summer-school-on-rare-disease-registries-and-fairification-

[17] dos Santos Vieira, B., Bernabé, C.H., Zhang, S., et al.: Towards FAIRification of sensitive and fragmented rare disease patient data: challenges and solutions in European reference network registries. Orphanet Journal of Rare Diseases 17(1), 436 (2022)

[18] DTL. Successful Hackathon to Make MOLGENIS FAIR (2016). Available at: https://www.dtls.nl/2016/10/30/successful-hackathon-make-molgenis-fair/

[19] Dutch Techcentre for Life Sciences. First Green BYOD - Yet Another Successful Bring Data Workshop (2015). Available at: https://www.dtls.nl/2015/02/11/first-green-byod-yet-another-successful-bring-data-workshop/

[20] Dutch Techcentre for Life Sciences. Bring Rett Syndrome Data workshop: a summary (2016). Available at: https://www.dtls.nl/2016/11/10/bring-rett-syndrome-data-workshop-summary/

[21] Wageningen University & Research. CGN Tomato Collection (2006). Available at: https://www.wur.nl/nl/onderzoek-resultaten/wettelijke-onderzoekstaken/centrum-voor-genetische-bronnen-nederland-1/plant/genebank/cgn-crop-collections/cgn-fruit-vegetables-collection/cgn-tomato-collection.htm

[22] Dutch Techcentre for Life Sciences. MolData2 implementation study: Facilitating molecular studies of rare diseases (2018). Available at: https://www.dtls.nl/2018/07/06/moldata2-implementation-study-facilitating-molecular-studies-of-rare-diseases/

[23] Roos, M., Gray, A.J., Waagmeester, A., et al.: Bring Your Own Data Workshops: A Mechanism to Aid Data Owners to Comply with Linked Data Best Practices. In: SWAT4LS (2014)

[24] Gray, Alasdair, J.G.: First FAIRPORT-ELIXIR BYOD workshop (2014). [Online]. Available at: https://alasdairgray.github.io/posts/2014-07-11-first-byod-workshop

[25] Dutch Techcentre for Life Sciences. About DTL (2023). Available at: https://www.dtls.nl/

[26] ELIXIR Europe. ELIXIR Europe (2023). Available at: https://elixir-europe.org/

[27] Uhlen, M., Oksvold, P., Fagerberg, L., et al.: Towards a knowledge-based human protein atlas. Nature Biotechnology 28(12), 1248–1250 (2010)

[28] Robert, V., Vu, D., Amor, A.B.H., et al.: MycoBank gearing up for new horizons. IMA Fungus 4, 371–379 (2013)

[29] Wilkinson, M.D., McCarthy, L., Vandervalk, B., et al.: SADI, SHARE, and the in silico scientific method. In: BMC bioinformatics. vol. 11. BioMed Central, p. 1–5 (2010)

[30] Williams, A.J., Harland, L., Groth, P., et al.: Open PHACTS: Semantic interoperability for drug discovery. Drug Discovery Today 17(21-22), 1188–1198 (2012)

[31] Slenter, D.N., Kutmon, M., Hanspers, K., et al.: WikiPathways: A multifaceted pathway database bridging metabolomics to other omics research. Nucleic Acids Research 46(D1), D661–D667 (2018)

[32] Gray, A.J., Groth, P., Loizou, A., et al.: Applying linked data approaches to pharmacology: Architectural decisions and implementation. Semantic Web 5(2), 101–113 (2014)

Data Intelligence Just Accepted MS.
https://doi.org/10.1162/dint_a_00236

Building Expertise on FAIR Through Evolving Bring Your Own Data (BYOD) Workshops:
Describing the Data, Software, and Management-focused Approaches and Their Evolution

[33] Willighagen, E.L., Waagmeester, A., Spjuth, O., et al.: The ChEMBL database as linked open data. Journal of Cheminformatics 5(1), 1–12 (2013)

[34] Lane, L., Argoud-Puy, G., Britan, A., et al.: neXtProt: A knowledge platform for human proteins. Nucleic Acids Research 40(D1), D76–D83 (2012)

[35] Simone Baldovino, M., Domenica Taruscio, M., Dario Roccatello, M.: Rare diseases in Europe: From a wide to a local perspective. Isr Med Assoc J 18, 359–63 (2016)

[36] RD-Connect. RD-Connect (2022). Retrieved May 3, 2023, Available at: https://rd-connect.eu/

[37] IRDiRC. Current IRDiRC Recognized Resources (2023). Available at: https://irdirc.org/resources-2/irdirc-recognized-resources/current-irdirc-recognized-resources/

[38] RD E, editor. International Summer School on Rare Disease Registries and FAIRification of Data. Rome, Italy: Istituto Superiore di Sanità (2022)

[39] European Commission. European Reference Networks (2023). Available at: https://health.ec.europa.eu/european-reference-networks/overview_en

[40] van Damme, P., Moreno, P.A., Bernabé, C.H., et al.: A Resource for Guiding Data Stewards to Make European Rare Disease Patient Registries FAIR. Ubiquity Press (2023)

[41] Bernabé, C., Sales, T.P., Schultes, E., et al.: A goal-oriented method for FAIRification planning. Research Square (Preprint) (2023)

[42] Zhang, S., Benis, N., Cornet, R.: Automated approach for quality assessment of RDF resources. BMC Medical Informatics and Decision Making 23(1), 1–16 (2023)

[43] dos Santos Vieira, B., Bernabé, C.H., Henriques, I., et al.: Critical steps towards large-scale implementation of the FAIR data principles (2023). Available at: https://doi.org/10.5281/zenodo.7867293

[44] Update citation: Bernabé, César H., et al. "The use of foundational ontologies in biomedical research." Journal of Biomedical Semantics 14.1 (2023): 21

[45] Neuhaus, F., Hastings, J.: Ontology development is consensus creation, not (merely) representation. Applied Ontology (Preprint), 1–19 (2022)

[46] Hettne, Kristina, M and Verhaar, Paul and Companjen, et al.: Bring Your Own Data FAIRification workshop 18 June 2019; 2023. Available at: https://osf.io/8ecg2

[47] University of Twente. FAIR Principles and the FAIRification process course (2021). Available at: https://osiris.utwente.nl/student/OnderwijsCatalogusSelect.do?selectie=cursus&cursus=202100258&collegejaar=2021&taal=en

[48] Aktau, A., Gambardella, A., Hettne, K., Ulzen van, N.: Final outcomes of the taskforce FAIRification as as Service (2023). Available at: https://doi.org/10.5281/zenodo.7546767

[49] GO FAIR. Making FAIR Metadata (2023). Available at: https://www.go-fair.org/today/making-fair-metadata/

[50] How to GO FAIR. Available at: https://www.go-fair.org/how-to-go-fair/

[51] Welter, D., Juty, N., Rocca-Serra, P., et al.: FAIR in action-a flexible framework to guide FAIRification. Scientific Data 10(1), 291 (2023)

[52] Schultes, E.: The FAIR hourglass: A framework for FAIR implementation. Fair Connect 1(1), 13–17 (2023)

[53] van Lin, N., Paliouras, G., Vroom, E., et al.: How patient organizations can drive FAIR data efforts to facilitate research and health care: A report of the virtual second international meeting on Duchenne data sharing, march 3, 2021. Journal of Neuromuscular Diseases 8(6), 1097–1108 (2021)

[54] Engelhardt, C., Biernacka, K., Coffey, A., et al.: D7.4 How to be FAIR with your data. A teaching and training handbook for higher education institutions. Zenodo (2022). Available at: https://doi.org/10.5281/zenodo.6674301

Building Expertise on FAIR Through Evolving Bring Your Own Data (BYOD) Workshops:
Describing the Data, Software, and Management-focused Approaches and Their Evolution

AUTHOR BIOGRAPHY

César Bernabé César Henrique Bernabé received a BSc degree in Computer Science from the Federal University of Espírito Santo in 2017, followed by an MSc degree in Informatics in 2020 from the same university. His master's research focused on the use of ontologies to improve requirements elicitation methods for software engineering. In 2020, he joined the Biosemantics group at Leiden University Medical Centre as a PhD student, where he is researching FAIR principles and the use of foundational ontologies for translational research, and how methods from Software Engineering can be reused and aid Bioinformatics research.

Lieze Thielemans Lieze Thielemans completed her medical degree at Imperial College London in 2020. During her time there she completed an intercalated BSc in Cardiovascular Sciences with a project FAIRifying a rare disease registry dataset at the Leiden University Medical Centre, supervised by Dr Marco Roos. Following this she has supported BYODs held in the Istituto Superior di Sanita, Italy, as a trainer and FAIR Steward. She is now an NIHR Academic Fellow, combining research into Primary Care and Medical Education with clinical work as a General Practice trainee.

Rajaram Kaliyaperumal Rajaram Kaliyaperumal was born in Pondicherry, India. He received a B.Tech degree in Biomedical Engineering from Pondicherry University, India in 2008 and an M.Sc degree in Biomedical Engineering from Linköping University, Sweden in 2011. In 2012 he joined the department of Computer and Information Science, Linköping University as a software engineer. During this time, he developed methods and tools to align and repair ontologies. In 2013 he joined the Biosemantics group, Leiden, in the Netherlands as a software developer. His current research activities include investigating the use of semantic web technology in the context of FAIR data and developing prototypes to demonstrate generating and the use of FAIR data.

Claudio Carta Claudio focused his university studies in the field of Biology and has a PhD in Biomedical and Cytomorphological Sciences. Since March 15, 2002 (ongoing), Claudio Carta has been staff personnel at the Istituto Superiore di Sanità, ISS, Rome (Italy), and his research is focused on Rare Diseases (RDs). Since October, 2012 (ongoing), he has been working as a researcher at the National Centre for Rare Diseases, ISS, Rome, Italy. Since 1st January 2019 (ongoing), Claudio Carta became the Local Technical Coordinator of ELIXIR IT in ISS.
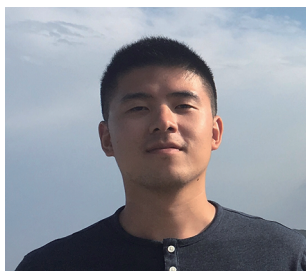
Shuxin Zhang Shuxin Zhang is a PhD candidate who has been working with data interoperability for four years with specific interest in the application of Semantic Web technologies. He obtained a Master of Science in Medical Informatics from University of Amsterdam (the Netherlands) in 2019 and a Bachelor of Science in Pharmacy from Zhejiang University (China) in 2015. Between 2020 and 2023, Shuxin was one of the FAIR data stewards in the European Joint Programme on Rare Diseases to support rare-disease registries in making their data Findable, Accessible, Interoperable, and Reusable (FAIR) for both humans and machines.

Celia W.G. van Gelder Celia van Gelder has a background in chemistry and bioinformatics and currently is Training Programme Manager of the Health-RI national health data infrastructure in the Netherlands, which is also the Dutch Node of ELIXIR (ELIXIR-NL), as well as Network Manager of the national Thematic Digital Competence Center for Life Sciences & Health (TDCC-LSH). Within DTL/ELIXIR-NL and Health-RI, FAIR data stewardship training activities have a prominent place.

Building Expertise on FAIR Through Evolving Bring Your Own Data (BYOD) Workshops:
Describing the Data, Software, and Management-focused Approaches and Their Evolution

Nirupama Benis Nirupama Benis is Assistant Professor and Post-Doctoral Researcher at the Department Medical Informatics, Reusable Health Data. She works on data analysis and data management, specifically the FAIRification of data. Most of her education and work has been in interdisciplinary fields, she thrives in the intersection of technology with biology.

Luiz Olavo Bonino da Silva Santos Luiz is an associate professor at the University of Twente and at Leiden University Medical Centre. His main research topics are the FAIR principles, semantic interoperability, ontology-driven conceptual modelling and intelligent systems design. Since 2014, he has been closely involved in designing and leading the implementation of a number of technologies and approaches to support the realisation of the FAIR principles.

Ronald Cornet Ronald Cornet holds a position as associate professor, principal investigator and principal educator at the department of Medical Informatics in the Amsterdam Public Health research institute, Amsterdam UMC.

He leads a research group on "reusable health data", in which research focuses on semantic interoperability, both from a technical perspective and from a users' point of view, as a key component to establishing FAIR data for healthcare and research. His research has resulted in over 100 journal publications and over 80 peer-reviewed conference proceedings. See the list of publications for a complete bibliography. Ronald has supervised 10 completed PhD-projects.

Building Expertise on FAIR Through Evolving Bring Your Own Data (BYOD) Workshops:
Describing the Data, Software, and Management-focused Approaches and Their Evolution

Bruna dos Santos Vieira Bruna has over 6 years of experience working in hospitals and a background in hospital management and public health. She studied Hospital Management at the University Feevale (Brazil-2014) and a Public Health Master's degree from the University of Porto (Portugal-2018), focusing on data quality of hospital-based cancer registries. In 2019, she joined the Medical BioSciences department at Radboudumc (Netherlands) as a FAIR data steward working on the Registry for Rare Vascular Anomalies (VASCA) and European Joint Programme on Rare Diseases (EJP-RD). At EJP-RD, Bruna collaborates with the Metadata WF, and FAIRification WF and leads the FAIRification Stewards. In 2022 Bruna joined Health-RI, the Dutch national research infrastructure as a semantic/metadata specialist.

Nawel Lalout Nawel is Project Manager FAIR Duchenne Data for the World Duchenne Organization and for the Duchenne Parent Project in the Netherlands where she manages, facilitates and oversees the FAIRification process of Duchenne Muscular Dystrophy patient registries and relevant resources. She coordinates the FAIRification efforts of the European Reference Network for Neuromuscular Diseases (ERN EURO-NMD) in her capacity as FAIR Data Steward at Radboud University Medical Center in the Netherlands. To support other Rare Diseases Patient Organizations and patient engagement in FAIR implementation activities, Nawel is co-chairing the Rare Diseases GO FAIR Patient network, a new initiative to foster understanding and adoption of the FAIR principles for and with patient organizations.
Nawel has extensive experience in international development and worked in a variety of sectors including pharmaceutical, non-for-profit, Life Sciences and high technology. She holds a Masters degree in General Management from the University of West London.

Inês Henriques Inês graduated from the University of Lisbon with a MSc in Pharmaceutical Sciences (2020), her thesis focused on the usage of pharmacogenetic data by community pharmacists to guide personalized therapeutic options for patients. In the same year, she started working as data analyst for the Portuguese National Medicine Agency (INFARMED I.P.), and later joined the European Medicines Agency as Scientific Officer focusing on data mapping of Post-Marketing Authorization Measures (Amsterdam-2021). She joined in September 2021 the Radboud University Medical centre as FAIRification Steward for both the European Joint Programme on Rare Diseases (2021-2022) and the Healthy Data Program (2023).

Building Expertise on FAIR Through Evolving Bring Your Own Data (BYOD) Workshops:
Describing the Data, Software, and Management-focused Approaches and Their Evolution

Alberto Cámara Ballesteros Alberto has a Master's Degree in Computational Biology from the Technical University of Madrid in Spain. After several experiences with the Semantic Web, knowledge representation, acquisition and data management, Alberto joined the European Joint Programme on Rare Diseases as a Pillar 2 contributor and a FAIRification Steward.

Kees Burger Kees Burger is a software engineer at the Leiden University Medical Center. He is leading the development of FAIR infrastructure reference implementations, and has been involved with FAIR developments since the inception of FAIR.

Martijn G. Kersloot Martijn Kersloot is an Assistant Professor in Medical Informatics at Amsterdam University Medical Centers and a Product Owner at Castor. With a PhD on FAIR (Findable, Accessible, Interoperable, and Reusable) data in medical research, he is passionate about semantic interoperability and ensuring that collected data is not wasted, but is optimally reused. Martijn is currently combining the 'best of both worlds' by training the medical informaticians of the future while also actively working to improve Castor's Clinical Data Management System.

Building Expertise on FAIR Through Evolving Bring Your Own Data (BYOD) Workshops:
Describing the Data, Software, and Management-focused Approaches and Their Evolution

Friederike Ehrhart Friederike studied Biology in Würzburg, Germany, with focus on Biotechnology. After graduating in 2004 she continued with a PhD thesis at Fraunhofer Institute for Biomedical Engineering in cooperation with Saarland University, Germany. After receiving the PhD title in 2009 she continued working as a postdoc researcher, later as group lead and head of laboratory at Fraunhofer before changing to Maastricht University, Department of Bioinformatics, in 2015. Since March 2019 she holds a tenured position as assistant professor.

Esther van Enckevort Esther van Enckevort is a Project Manager at the Department of Genetics at the University Medical Center Groningen (UMCG). She is an expert in software development, data harmonisation, catalogues, and FAIR. She has a strong systems engineering background, including ontologies and semantic web. In recent years she has extended her expertise by leading the development of various catalogues (including BBMRI-ERIC Directory, RD-Connect, LifeCycle, EUCAN-Connect and ISBER). She has been involved in many projects and organisations including the EJP-RD, BBMRI-ERIC CS-IT, EUCAN-Connect, ISBER and Health-RI. She has experience leading work packages in EUCAN-Connect (WP2-Federated Catalogue), BBMRI-ERIC (WP1-Directory) and working groups in Health-RI (Portal), ISBER (Locator), IRDiRC (Machine Readable Consent). Esther is a member of the IRDiRC Interdisciplinary Scientific Committee and initiated the IRDiRC working group on machine readable consent that developed the Digital Use Conditions (DUC) framework.

Building Expertise on FAIR Through Evolving Bring Your Own Data (BYOD) Workshops:
Describing the Data, Software, and Management-focused Approaches and Their Evolution

Chris T. Evelo Chris Evelo is the founder and head of the department of Bioinformatics-BiGCaT at Maastricht University where he leads an enthusiastic group of researchers and he is a PI in the Maastricht Center for Systems Biology (MaCSBio). He was trained in biomolecular sciences and his early career was in experimental biochemical toxicology where he also applied physiologically based biokinetic modelling. This, combined with subsequent experiences in nutrition and drug research and his broad interest made him a generalist, with a broad focus on the human interaction with the molecular environment. His current research focuses on bioinformatics for integrative systems biology; aiming at a better interpretation of experimental data through integration in data models that build on structuring existing knowledge.

Alasdair J. G. Gray Alasdair's research focuses on practical data management and its application in information systems-utilising and extending advances in knowledge management technologies to improve information systems. This approach has supported advances in science while deepening the understanding of knowledge based systems to support future applications. Alasdair has collaborated with a large range of scientists-from astronomers through the life sciences to environmental scientists-working with both academic research labs and industrial partners.

Marc Hanauer Marc Hanauer is an research engineer at INSERM, Chief Technology Officer at Orphanet since 2007 and Orphanet Deputy director since 2017.

He is also in charge of managing Orphanet's innovation strategy and informatic teams.

Marc Hanauer is in charge of interoperability and FAIRification process amongst Orphanet and Orphadata datasets. He leads the technical production of ontologies such as ORDO (Orphanet Rare Diseases Ontology), HOOM (HPO-ORDO Ontological Module), RDCO (Rare Diseases Cases Ontology) and tools and services derived from the Orphanet Knowledge Base (API, Orphapackets, Data visualisation tool, mapping services...)

He's involved in ELIXIR communities, Elixir Core Resource (Orphadata) and Global Core Biodata Resources.

He participates in several european projects such as RD-code, OD4RD (Orphanet Data for Rare Diseases) European Joint Programm on Rare Diseases (EJP-RD) and SOLVE-RD project.

Kristina Hettne Kristina is a Digital Scholarship Librarian at the Leiden University Libraries' Centre for Digital Scholarship. She helps researchers navigate Open Access and FAIR research data management and enjoys shaping the future of these subjects. She holds a PhD in Bioinformatics and MSc in Computer Science.
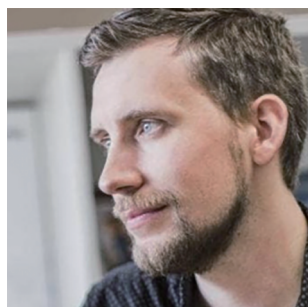
Joep de Ligt Joep is a bioinformatician with extensive experience in translational genomics in human health. He is currently the Lead Database at Hartwig Medical Foundation where he strives to make cancer data better accesible to support research to improve outcomes for current and future patients. He has been an advocate and practical implmenter of FAIR for many years. He has worked on infectious diseases (inclduing the covid response) in New Zealand, as well as on stem ceel biology and cancer genetics in Utrecht (NL) after a PhD in clinical genetics in Nijmegen (NL).

Arnaldo Pereira Arnaldo Pereira is an Assistant Professor of Computer Science at the Polytechnic Institute of Bragança, Portugal. He earned a Ph.D. in Computer Science, specializing in Semantic Data and Natural Language Processing, from the University of Aveiro, Portugal. His current research interests are related to Intelligent Systems, Multi-agent Systems, and Large Language Models.

Building Expertise on FAIR Through Evolving Bring Your Own Data (BYOD) Workshops:
Describing the Data, Software, and Management-focused Approaches and Their Evolution

Núria Queralt-Rosinach Núria Queralt Rosinach is from Reus, Catalonia (Spain). She obtained her MSc in Bioinformatics from Pompeu Fabra University (UB/UPF) in June 2008, her PhD in Computational Chemistry from Rovira i Virgili University (URV) in May of 2010, between travel and travel she did her first postdoc at the IBI Group in the beautiful, Mediterranean and sunny Barcelona (Catalonia, Spain) started in June of 2012, and joined the Su Lab as a research associate in June of 2016. Her current research project is mainly focused on biomedical discovery specifically on rare genetic diseases, with focus on the creation of novel methods and tools based on semantic and artificial intelligence (AI) technologies to understand the mechanisms underpinning diseases and therapies, and the interplay genotype-environment-phenotype.

Erik Schultes Erik Schultes is the FAIR Implementation Lead at the *GO FAIR Foundation*, Senior Researcher at the *Leiden Academic Center for Drug Research at Leiden University*, and is the Scientific Director at *partners in FAIR*. Erik was International Science Coordinator at the GO FAIR International Support and Coordination Office from its inception in 2017 until 2021. In GO FAIR, Erik has been working with a diverse community of stakeholders to develop scalable implementations of FAIR data and services, including the Metadata for Machine Workshop and the FAIR Implementation Profile. Erik's academic interests currently focus on capturing high-throughput metabolomics data with as FAIR Digital Twins.

Domenica Taruscio Dr. Taruscio is director of the National Center for Rare Diseases at the Italian National Institute of Health (Istituto Superiore di Sanità) in Rome, Italy. She completed her medical studies at Bologna University in Italy and her postdoctoral work at the Yale School of Medicine in New Haven, Connecticut. She is an active contributor to the development of the European scientific and institutional framework on rare diseases and orphan medicinal products. Her studies have been focused on rare diseases, where she works at the intersection of science and society on issues ranging from prevention to diagnosis of rare diseases, as well as care of patients. Her activities include training health operators and developing educational information that empowers patients and improves their quality of life. Dr. Taruscio has coordinated or participated in several projects funded by the European Commission. She is a member of national, European, and international committees, including the International Rare Diseases Research Consortium (IRDiRC). She is cofounder of the Undiagnosed Diseases Network International and represents Italy at the European Commission on the Board of Member States for the European Reference Networks. She is also president of the International Collaboration on Rare Diseases and Orphan Drugs (ICORD), an international society whose mission is to improve the welfare of patients with rare diseases and their families worldwide through better knowledge, research, care, information, education, and awareness. In July 2021, she was nominated as a Distinguished Visitor of the Faculty of Medical Sciences (National University of Córdoba) and a Foreign Corresponding Member of the Academy of Medical Sciences (Córdoba, Argentina).

Mark D. Wilkinson Mark D Wilkinson, B.Sc. Hons Genetics, University of Alberta; Ph.D. Botany, University of British Columbia (1996). Currently Isaac Peral Distinguished Researcher, Universidad Politecnica de Madrid. My research is primarily focused on machine-actionable data, focusing on the biological data space, and building automated agents for large-scale data discovery and integration. I am lead author of the FAIR Data Principles, and have built a variety of software tools around "being FAIR", including the first automated FAIR assessment tool. I also have a spin-off company that does consulting and software development around helping data publishers be more FAIR. I am currently co-Chair of the European Open Science Cloud Task Force on FAIR Metrics and Data Quality where we are working on FAIR-compliant standards for all participants in the EOSC ecosystem.

Egon L. Willighagen Egon Willighagen studied chemistry, majored in cheminformatics, and received his PhD in chemometrics (2008), both at the University of Nijmegen. After postdoctoral visits to the University of Cambridge, Cologne University, Wageningen University, Uppsala University, and the Karolinska universities where he applied his approaches to metabolomics, toxicology, and drug discovery, he returned to The Netherlands, and is now assistant professor as Maastricht University. He is known for his Open Science work, including founding the Chemistry Development Kit, maintaining Jmol in the past and BridgeDb now, and contributions to many other open science projects. He has led various international work groups (including the GO FAIR Cheminstry IN), was editor-in-chief for five years of the Journal of Cheminformatics, and active in the ELIXIR Europe in various roles.

Mascha Jansen is a skilled professional currently working as a Programme Manager for Health-RI. With a background in biotechnology and communications, she has over 10 years of experience in the life sciences domain, including roles at Go FAIR, GO FAIR Foundation, DTL, and BBMRI. Mascha has led several projects and been actively involved in numerous FAIR (Findable, Accessible, Interoperable, and Reusable) data initiatives, such as the BYOD (Bring Your Own Data) programme. She excels in coordinating roles and driving impactful projects.

Barend Mons Barend Mons is a molecular biologist by training and a leading FAIR data specialist. He is currently a professor in Leiden and most known for innovations in scholarly collaboration, especially nanopublications, knowledge graph based discovery and most recently the FAIR data initiative and GO FAIR. Since 2012 he is Professor in biosemantics in the Department of Human Genetics at the Leiden University Medical Center (LUMC) and since May 2022 also at the Leiden Academic Centre for Drug Research (LACDR) in The Netherlands. 2017-2023 Barend was the elected president of CODATA, and since 2021 he is the Scientific Director of the GO FAIR Foundation.

Marco Roos Marco is assistant professor and group leader of the Biosemantics group, LUMC, Human Genetics department. His research focus is on making state-of-the-art computer science applicable to enhance biomedical research (e-Science), particularly the application of computational knowledge discovery and linked data techniques to address translational research challenges of rare human diseases. Biosemantics group is known for co-founding and advocating the FAIR data principles. At an international level, Marco is focused on the implementation of FAIR principles to create a powerful substrate and world-wide robust infrastructure for knowledge discovery across distributed rare disease data resources.

Annika Jacobsen Annika is a senior researcher at the Biosemantics group, Human Genetics Department, Leiden University Medical Center, The Netherlands. She obtained her Bachelor and Master degrees at the Technical University of Denmark in 2009 and 2012, and her PhD degree at the Vrije Universiteit Amsterdam in 2019. Her research interests are to create and apply interoperable FAIR rare disease data with the aim to learn more about cause, diagnosis and treatment.