FAIR Principles and Digital Objects: Accelerating Convergence on a Data Infrastructure

Erik Schultes[1(✉)] and Peter Wittenburg[2(✉)]

[1] GO FAIR International Support and Coordination Office, Poortgebouw N-01, Rijnsburgerweg 10, 2333 AA Leiden, The Netherlands
erik.schultes@go-fair.org
[2] Max Planck Computing and Data Facility, Gemeindeweg 55, 47533 Kleve, Germany
peter.wittenburg@mpi.nl

Abstract. As Moore's Law and associated technical advances continue to bulldoze their way through society, both exciting possibilities and severe challenges emerge. The upside is the explosive growth of data and compute resources that promise revolutionary modes of discovery and innovation not only within traditional knowledge disciplines, but especially between them. The challenge, however, is to build the large-scale, widely accessible, persistent and automated infrastructures that will be necessary for navigating and managing the unprecedented complexity of exponentially increasing quantities of distributed and heterogenous data. This will require innovations in both the technical and social domains. Inspired by the successful development of the Internet and leveraging the Digital Object Framework and FAIR Principles (for making data Findable, Accessible, Interoperable and Reusable by machines) the GO FAIR initiative works with voluntary stakeholders to accelerate convergence on minimal standards and working implementations leading to an Internet of FAIR Data and Services (IFDS). In close collaboration with GO FAIR and DONA, the RDA GEDE and C2CAMP initiatives will continue its FAIR DO implementation efforts..

Keywords: Data analytics · Data management · Data intensive science · Digital libraries · Digital objects · FAIR data

1 Introduction

Existing data stewardship practices are highly inefficient. Numerous studies indicate that data scientists both in academia and industry spend 70–80% of their time on mundane, manual procedures to locate, access, and format data for reuse [1, 2]. Methodological legacies inherited from a pre-digital era (e.g., poor capture of metadata, broken links to various research assets) and outdated professional incentives (e.g., only rewarding publication of research articles rather than also datasets and other research outputs) contribute to massive data loss and a well-documented reproducibility crisis [3–5]. Coupled with the exponential increases in data volumes (driven by, among other things, high through-put instrumentation and IoT data streams) the urgency for

automated, commonly usable and persistent data infrastructures (i.e., a datanet for Machines) is increasingly recognised by numerous national and international organisations, science funders and industry [6–11]. Despite the urgent need, building a generalised, ubiquitous, data infrastructure that is widely used by diverse stakeholders is an inherently distributed and difficult process to direct. Knowing this to be the case, members of several RDA groups started the C2CAMP initiative [12] to join results and to build a testbed for a Digital Objects based infrastructure which will help overcoming the huge inefficiencies in data intensive science. In parallel, the GO FAIR initiative was launched to also accelerate data infrastructure development by leveraging general patterns of phased development described in other revolutionary infrastructures, including the Internet and the World Wide Web (WWW) [13].

2   Learning from Previous Revolutionary Infrastructures

Revolutionary Infrastructures (for example, transportation, electrification, telecommunications, and computer networks) follow five phases of development [14, 15]: (1) Vision: New discoveries and technologies lead to the anticipation of broad new application spaces; (2) Creolization: Inspired by the Vision, numerous experimental implementations are created, resulting in an uneven landscape of independently developed prototypes; (3) Attraction: Some solutions prove more viable, and are effectively generalised to achieve a simplified set of 'universal principles' that attract the attention of others working in the field; (4) Convergence: Various Attractors voluntarily decide to bridge otherwise isolated application solutions, and a compelling global infrastructure begins to emerge at the expense of the many other possibilities; (5) Exploitation: As widespread commitment to a particular implementation emerges, economy of scale kicks in, and what was hard and cost-prohibitive, now becomes easy and affordable. Users in the Exploitation phase might not even be aware of the infrastructure systems they routinely use (e.g., most users of the internet are blissfully ignorant of TCP/IP).

In the specific case of the Internet, there had been early *Visions* of interlinked computers throughout the 1950s and 1960s. By 1969, ARPAnet had initiated the phases of *Creolization* (and later *Attraction*) with the co-existence of multiple, specialised solutions, e.g., X25, Ethernet, ARCNET, and others. This work demonstrated the feasibility of computer networks and drew the attention of large investors (e.g., IBM, DEC). But this investment resulted in numerous incompatible standards that first drove insights but later slowed progress. *Convergence* was eventually triggered with TCP/IP protocols (early 1970s) and the 7-layer ISO/OSI reference model (early 1980s). This was because, in particular, the minimal TCP/IP standard allowed various networks to interoperate while at the same time maintaining maximum freedom to engineer solutions at the implementation layer 'below' and application layer 'above' (creating the so-called "hourglass" architecture of the Internet, with TCP/IP at the narrow waist). It was working implementations (however embryonic) and the simplicity of the hourglass approach that motivated influential decision makers "to move towards using TCP/IP as universal for implementing global computer networking". With a stabilized universal in place, *Exploitation* soon followed, with rapid investment in both hardware

and software, that is the now familiar story of the Internet. By 1992, the Internet Society was set up to coordinate further develop TCP/IP approaches to networking.

It is important to note that the use of TCP/IP has always been voluntary, and at no time was its use ever required. Indeed, top-down enforcement policies would likely have killed its effectiveness as an attractor. Instead, once a 'critical mass' of influential users had adopted TCP/IP, the larger community followed, driving convergence. An analogous pattern of development (voluntary use, attractor effect in the community) occurred soon after with the formation of the WWW, in this case with HTTP playing the role of TCP/IP. The significance of this historical insight can not be understated. It enables some degree of coordination in the development of new infrastructures, because only a relatively few (albeit influential) users need be convinced to invest in a particular technology. Once the 'critical mass' is assembled, the 'long tail' of community stakeholders will likely follow.

Even before the 2000's, visionaries had already anticipated the need for a general-purpose data infrastructure. Digital Object based infrastructures such as the Digital Object Architecture [16], systems supporting Persistent Identifiers (PIDs) and the Semantic Web (a framework for knowledge representation built on top of existing Internet and WWW infrastructures) appeared as an important component, ensuring both data interoperation and machine readability. Since then, difficult problems in this space have been investigated resulting in a plenum of new, co-existing methods, languages, software and specialised hardware, producing by now, a protracted period of Creolization. By 2012 the Attraction phase was underway with public discussions about component specifications, principles and procedures for semantically enabled data infrastructures [17, 18]. RDA was officially started in 2013 as a broad group of data experts including now more than 7000 persons from more than 120 countries and had first results from working groups in 2014. Some of the RDA experts recognised the need to bring the various results together and started first the RDA Data Fabric group [19] to identify Digital Objects as the common ground and to specify additional needs. Then, emerging from RDA, the C2CAMP collaboration was created to not only specify procedures and interfaces, but to start working on a joint testbed in close collaboration with the DONA foundation [20]. Later the GEDE collaboration adopted the DO topic and subsequently organising more than 150 data experts from about 47 European research infrastructures to participate in the discussions on Digital Objects.

By early 2014, in a workshop hosted by the Lorentz Center (Leiden), the above mentioned discussion culminated in the generalised and broadly applicable FAIR Principles for data reuse [21, 22]. In a now widely cited commentary (indicative of the Attraction phase) [23], the FAIR approach had been defined as "Data and services that are findable, accessible, interoperable, and re-usable both for machines and for people" and 15 high-level Principles had been articulated, Fig. 1.
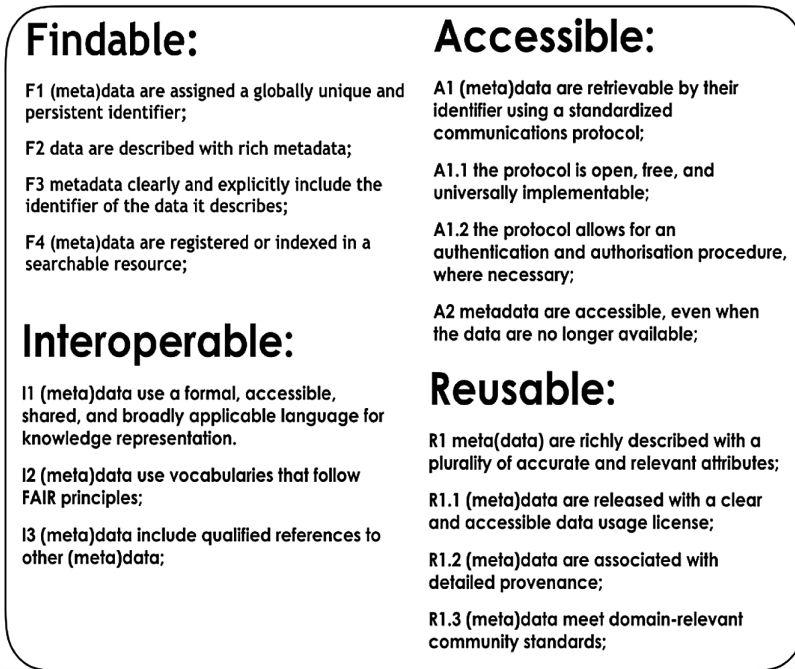
Findable:

F1 (meta)data are assigned a globally unique and persistent identifier;

F2 data are described with rich metadata;

F3 metadata clearly and explicitly include the identifier of the data it describes;

F4 (meta)data are registered or indexed in a searchable resource;

Interoperable:

I1 (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.

I2 (meta)data use vocabularies that follow FAIR principles;

I3 (meta)data include qualified references to other (meta)data;

Accessible:

A1 (meta)data are retrievable by their identifier using a standardized communications protocol;

A1.1 the protocol is open, free, and universally implementable;

A1.2 the protocol allows for an authentication and authorisation procedure, where necessary;

A2 metadata are accessible, even when the data are no longer available;

Reusable:

R1 meta(data) are richly described with a plurality of accurate and relevant attributes;

R1.1 (meta)data are released with a clear and accessible data usage license;

R1.2 (meta)data are associated with detailed provenance;

R1.3 (meta)data meet domain-relevant community standards;

Fig. 1. The 15 FAIR principles ensuring machine findability, accessibility, interoperation and re-use of digital resources.

Immediately following their publication (April 2016), the FAIR Principles (and later, the corresponding FAIR Metrics [24] and FAIR Maturity Indicators [25]) have been acting as a powerful attractor in the emerging data infrastructure.

Following the previous examples, the Convergence phase of the data infrastructure will commence once a 'critical mass' of users commits to particular, minimal specification for automatic routing of FAIR data and services.

In the meantime the strong relationship between the FAIR Principles and FAIR Digital Objects has been observed by the GO FAIR and C2CAMP/GEDE experts [26]. These groups are now working together to harmonise the DO and FAIR approaches into a formally defined "FAIR DO", with the aim to accelerate convergence on a globally distributed data infrastructure. A data infrastructure will likely be substantially more complex than its predecessors in that a FAIR Digital Objects based Internet of FAIR Data and Services (IFDS) necessitates the wide acceptance of the DO Interface Protocol, the use of the potential of the globally available Handle System to solve the binding challenge and to elaborate on semantically enabled metadata descriptions. The 'FAIRification' of digital resources is not trivial, and widespread application will require an ecosystem of methods, tooling, services and training that help communities of diverse stakeholders to create and use FAIR resources. While C2CAMP/GEDE and DONA will showcase a stable DO-based eco-system of infrastructures, GO FAIR will support and coordinate bottom-up community initiatives that aim to 'Make FAIR easy" [27].

3   FAIR Digital Objects

3.1   Digital Objects

Digital Objects were already introduced in an early paper by Kahn & Wilensky in 1995 and then in an updated version in 2006 [28, 29]. As Wittenburg et al. [30] have shown the concept is very much related with computer science concepts such as "object-oriented programming" [31], "abstract data types" [32] and "object stores" [33] which are at the basis of state-of-the-art cloud systems such as Amazon's S3. We can therefore claim that the concept of "objects", is closely related with ideas such as "encapsulation", "virtualization", and "interfacing by defined methods", has shown its great importance to help designing complex systems.
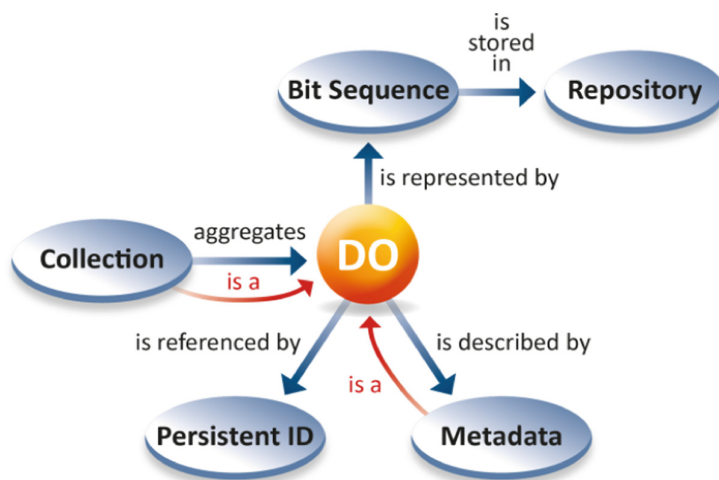
Fig. 2. This figure indicates the core data model as it was worked out within the RDA group Data Foundation and Terminology (DFT).

In 2014, the RDA group "Data Foundation and Terminology" (DFT) published its results on a core data model and the corresponding basic terminology. It summarized the discussions about Digital Objects (DO) as (see Fig. 2):

- DOs are at the core of a proper data organizations in so far as it has the capacity to bind crucial entities which are necessary for a stable and reusable domain of data;
- DOs have a bit sequence (content) which can be stored in various repositories, are referenced by a unique and persistent identifier (PID) issued by a trustworthy globally available resolution system and is described by various types of metadata that can include descriptive, system, access rights, license, contractual, transactional and other kinds of meta information about the DO;
- Metadata itself are DOs;

- DOs can be combined to collections which also are DOs, i.e. have a PID and are described by metadata;
- DOs can include all kinds of digital information such as data, software, configurations, representations of persons, institutions, semantic concepts, etc.

We can also look schematically at DOs from a different point of view, if we extend the above definition by encapsulation principles as being introduced by the RDA group "Data Type Registry". One of the metadata types describing a DO is its "type" which is summarizing several technical metadata attributes. A Data Type Registry allows users to relate data types with operations which are also DOs of a specific type. These defined operations allow users to realize the encapsulation principle as requested by Abstract Data Types. Figure 3 indicates this encapsulation which can be implemented when strong and stable binding is being realized. The usage of PID systems such as the Handle System [34] allows creating such a strong and stable binding, since the PID records allow including pointers (PIDs) to all relevant entities and metadata types associated with a DO.
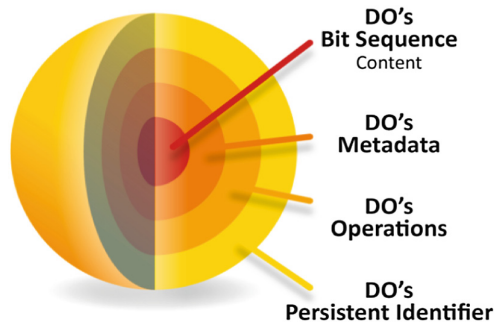

Fig. 3. This figure schematically indicates the types of abstraction, binding and encapsulation that can be implemented with DOs.

Recently, a second version of a protocol to interact with DOs, the DO Interface Protocol (DOIPV2.0), has been opened for broad discussion by DONA. It basically describes how clients interact with DOs where all involved actors are represented by PIDs. DOIP is meant to have a relevance that is comparable to TCP/IP for the Internet, i.e. it should become a fundamental protocol to manage and exchange digital objects.

The definition of the term "Digital Object" in the DOIP document is intended to be restricted in its focus on the minimalistic and operational nature of the protocol, i.e. a DO has a bit sequence, a PID and a type. Although elegant in its simplicity, this minimal definition itself gives no specification for the recording the scientific semantics or other domain knowledge that is equally important in routing and processing research data and services in general included in metadata. Recently, the RDA DFT group started addressing this issue by augmenting the minimal DO definition in the context of the FAIR Principles itself, defining the "FAIR Digital Object", since it includes the

strong binding of different types of metadata which are important for the interpretation and access and reuse of the bit sequences.

The C2CAMP initiative is devoted to implement a FAIR DO based infrastructure including understanding DOs as active entities that have methods associated with them. A broad discussion started in Europe in the realm of GEDE [35] involving 150 experts from about 50 research infrastructures to intensify the discussions not only about the potential of FAIR DOs to build federative data infrastructures, but in particular to also use FAIR DOs to systematically structure the domain of digital entities in scientific disciplines. A recent workshop [36] combining these two roles of FAIR DOs showed globally organised research communities such as biodiversity, climate modeling and language research have far going plans to use their potential to increase trust, to define clear anchors for a complex system of annotation layers, to better utilize automatic workflow frameworks and much more. Moving forward, there is now increasing interest in the fusion of DO and FAIR approaches both at the conceptual and technical levels.

3.2   FAIR Principles

The original publication announcing the FAIR Principles does not discuss implementation choices. Given that many different combinations of technology choices and use of standards could conceivably implement the FAIR Principles, the GO FAIR initiative was launched in late 2017 by the Dutch, German and French governments as a means to pragmatically accelerate community Convergence. The initial vehicle for GO FAIR is the International Support and Coordination Office (GFISCO). Following the examples of the Internet and WWW, the GFISCO operates through voluntary stakeholder participation attempting to reach a 'critical mass' of users committed to a set of absolute minimal technology specifications. Beyond these minimal specifications, there is unrestricted room to innovate.

GFISCO is stakeholder governed, and includes researchers from specialized knowledge domains (e.g., earth sciences [37], chemistry [38]) but also policy bodies (e.g., CODATA, RDA, FORCE11), publishers (e.g., Elsevier, Springer-Nature), repositories (e.g., Figshare), and funding agencies (e.g., The American NSF and NIH, the Health Research Board of Ireland, and the Dutch ZonMW). GFISCO brokers among stakeholders, the choice of standards implementing the functions of the FAIR Principles and emerging best practices leading to the Internet of FAIR Data and Services. GFISCO operates via supporting and coordinating Implementation Networks (INs), which are voluntary international consortia that self-organize (and are self-funded) to implement elements of the IFDS. GO FAIR INs belong to 3 broad topical pillars: GO BUILD, GO TRAIN and GO CHANGE.

GO BUILD focuses on the technological aspects of the IFDS, including the design and building of reference implementations for elements composing the IFDS such as FAIR Metrics, FAIR Data Points [39, 40], FAIRification tools and other FAIR-compliant services. Currently, there are 8 INs under the GO BUILD pillar.

Other technology-related activities in GO FAIR include ongoing "Metadata for Machines" workshops and "Community Challenges", aiming to help communities achieve adoption of globally unique and persistent identifiers, agree on common

metadata representation formats, agree on a minimal set of generic metadata content and define domain-relevant community standards.

The overall objective of the GO TRAIN pillar is to create a scalable framework that is used in higher education programs and throughout industry to train large numbers of certified data stewards (estimated to be 500,000 for Europe [41], millions more worldwide). GO TRAIN supports and coordinates two activities: (1) The development of canonical training curricula focused on FAIR Data Stewardship; (2) The development of certification schema for competencies in FAIR Data Stewardship (providing professional career trajectories, which in turn, are intended to drive rapid uptake of FAIR practices among diverse stakeholders). Currently there are two GO TRAIN INs. The first is the Training Frameworks IN which aims to develop schema for FAIR Data Stewardship education (including train-the-trainer curricula and endorsement specifications), with lenses for Managers, Principal Investigators and Data Stewards themselves. Secondly, The FAIR Curriculum IN will re-use the Carpentries Open, community based curriculum development model [42] to develop novel modular lessons for FAIR data stewardship.

The overall purpose of the GO CHANGE pillar to support and coordinate systemic culture change that transforms existing data management practices into the respected profession of data stewardship. This includes the development of new funding schema, sustainability strategies, and business models. GO CHANGE stakeholders range from international policy makers and national governments to organisation managers and front-line data producers and data stewards. A key IN for GO CHANGE is a FAIR resource hub that aggregates multiple resources for FAIR data stewardship planning, compliance, and assessment.

4   GO FAIR, DO FAIR

A preliminary analysis can easily show that there is a close but highly complementary relationship between the FAIR Principles and the concept of FAIR Digital Objects.

4.1   Data to be Findable

The DO model is widely compliant with the F-dimension of the FAIR principles and gives an implementation mechanism. The DO model is explicit in how to do things - in particular the binding of different informational entities associated with the DO to guarantee FAIRness - but does not specify the possible usage of DO's content. It includes certified repositories as active components and care takers of data and does not make statements about the content of metadata, since this is very much purpose dependent and domain specific. Whereas DOs are agnostic about its content and treat all kinds of content (data, metadata, software, semantic assertions, etc.) the same way the FAIR principle F2 requests rich metadata for findability that can be both generic and domain focused.

4.2   Data to be Accessible

Entirely consistent with the Accession-Related FAIR Principles, the DO Core Model enables the building of infrastructure that makes data and metadata accessible since it supports all requirements with respect to open and free to use protocols but also proper authentication and authorization where necessary. Except for the PID infrastructure which is an essential element of DO based infrastructures, the DO model assigns the responsibility to repositories to define policies and implement appropriate mechanisms. As such, authentication and authorization aspects need to be taken care by the inter-acting distributed components on the Internet of FAIR Data and Services. The FAIR Principle A2 stipulates a condition that is only implicit in the DO model, which is that metadata should persist, even if the original data are deleted or in some way no longer available.

4.3   Data to be Interoperable

DOs take care of interoperability at the level of data organisation due to its inherent binding concept and its stable linking based on specific PIDs such as Handles and this in a way that is machine actionable. The DO Interface Protocol is a universal mech-anism to interact with DOs independent of how repositories organise and model their digital entities. Although with respect to other interoperability layers such as structural and semantical encoding of content the DO concept is agonistic, it does facilitate the operational work at these levels by allowing users to use the DO model for all kinds of digital entities and thus guaranteeing stable binding that is necessary for interoperation. However, again we see the complementarity between the FAIR Principles and the DO model, in that the 3 Interoperation-related FAIR Principles are explicit about rich, qualified semantic encoding.

4.4   Data to be Reusable

The DFT Core Model explicitly mentions the role of key properties of DO's content being part of the PID record or being referred to by stable and persistent links. Due to strong typing as suggested by the RDA Kernel Information group of all these attributes machine actionability is given a great advantage. The binding concept of the DFT Core Model enables the linking of various aspects closely related with the DO such as provenance, smart contracts (actionable licenses), transaction records and even more that go beyond the FAIR principles. The DO Core Model is agnostic with respect to the concrete specifications, since it respects (indeed, expects) that other groups such as W3C (PROV), the blockchain community, etc. are providing mechanisms and defi-nitions which will be used to implement special wishes. Thus again we can say that the DO concept facilitates the implementation of the FAIR requirements, although the FAIR components have the capability and mandate to express rich and nuanced semantics.

4.5   Summary

Due to their complementarity we see GO FAIR and DO activities as a giant step towards improving data practices and it was a logical step for the C2CAMP/GEDE initiatives to become an Implementation Network in GO FAIR and to also align discussions with the GEDE DO Topic Group as well. GO FAIR distinguishes three major areas of work (see Fig. 4) to build FAIR compliant infrastructures: data, tools and compute resources, which in the DO domain are Digital Objects of different types. All three areas share one central infrastructure, the turbine's driving axis. To expand this metaphor one could imagine the DOs to be the driving axis that combines all three areas and the DO Interface Protocol and the protocol to resolve persistent identifiers as the underlying basic protocols all areas are using. While DOs implement the F and A dimensions of FAIR more or less directly, they facilitate the I and R dimensions.
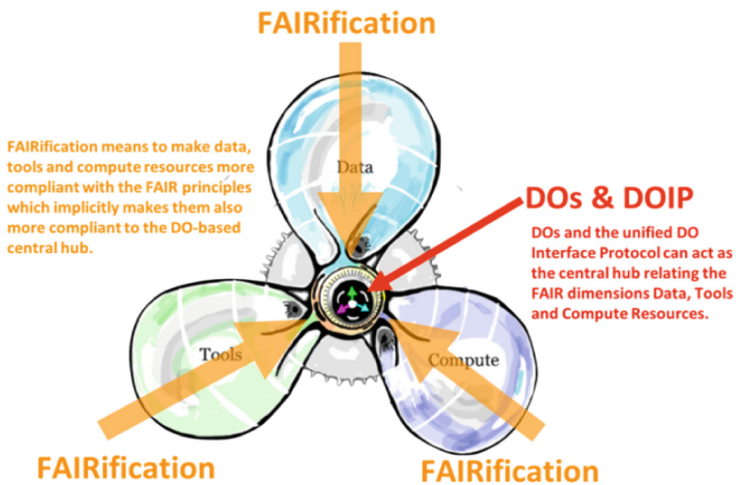

Fig. 4. This figure indicates the major dimensions of the GO FAIR work and the interpretation of Digital Objects being the driving wheel combining all three dimensions.

The FAIR Digital Object approach provides technical solutions needed to implement FAIR principles. In particular, federated systems such as intended, for example, by the European Open Science Cloud will need such a basic interoperability layer to achieve the required scalability, stability and FAIRness. Building such a comprehensive and expensive infrastructure eco-system will need to be based on solid fundaments as offered by the FAIR principles and FAIR Digital Objects to overcome major hurdles in making data more reusable.

5   Participating

5.1   RDA GEDE DO Topic Group

GEDE, the Group of European Data Experts, is organised within RDA and defines so-called topic groups to allow interested experts to work on specific thematic topics. One of these topics are the FAIR Digital Objects where 150 distinguished data experts from about 50 European research infrastructures and some international colleagues are discussing intensively about how to improve data work by adopting FAIR DOs. Currently, a set of more than 30 use cases has been presented by different communities which will lead to a new paper on FAIR DOs driven by scientific interests. Participation in GEDE DO is open to anyone interested.

5.2   C2CAMP

C2CAMP is a global collaboration of experts who want to build DO-based infrastructures and tools that emerged from the work in RDA groups and that closely collaborates with the GEDE DO topic group. C2CAMP participation is open for anyone who wants to actively contribute to the FAIR DO testbed.

In 2018 C2CAMP joined GO FAIR as an implementation network to foster the interaction with other implementation networks.

5.3   GO FAIR Implementation Network

GO FAIR INs foster a collaborative community of harmonized practice which leads to Convergence and allows members to 'speak with one voice' on critical issues regarding FAIR data infrastructures. Anyone (i.e., a person, an institution or a network organisation) can join an existing or create a new GO FAIR IN [43]. The list of current GO FAIR INs can be found at the GO FAIR website [44]. The requirements to become an IN are minimal: (1) have a plan to implement an element of the IFDS (including adequate resourcing to accomplish the proposed goals); (2) comply with the GO FAIR Rules of Engagement (essentially, commitment to the FAIR Principles and 'no vendor lock-in[1]'); (3) have sufficient critical mass to be regarded as thought leaders in the field of expertise.

6   Conclusions

As described by Wittenburg and Strawn [14] we see trends to convergence finding in the data domain. Two major action lines have been kicked off almost in parallel: on the one hand by the RDA groups that worked together in the RDA Data Fabric group and later started the C2CAMP and GEDE collaboration; on the other hand by the group working on the FAIR principles and provided the background in which the GO FAIR initiative was launched. Both initiatives saw the need to turn specifications into active

---

[1] https://www.go-fair.org/implementation-networks/rules-of-engagement/.

implementation work and thus contributing to the emerging practical eco-system of data infrastructures. In addition, they understood that FAIR principles and FAIR Digital Objects are complementary.

A new wave of investments in large research and data infrastructures can be observed including the European Open Science Cloud and national science clouds in most of the European member states. The relevant actors sense that what they are aiming at is finally a complex enterprise with many open questions - their undertaking is a huge experiment that will lead to a transformation of science. Two of these open questions are: how complexity can be broken down and how a stable fundament for the coming decades can be achieved that will not hamper the needed progress in science. It should be noted that the severity of obstacles to data reuse is driven ultimately by Big Data (Moore's Law) and in this sense, the problems extend far beyond the research domain. Industry is confronted with similar challenges and thus may need to find similar solutions if it will completely be locked in proprietary platforms.

We recommend therefore following the trend to FAIR data and doing this by implementing the FAIR DO concept that has as core elements globally resolved persistent identifiers and the Digital Object Interface Protocol - all being open specifications governed by the non-profit Swiss DONA Foundation.

References

1. Stehouwer, H., Wittenburg, P.: RDA Europe: Data Practices Analysis (2018). http://hdl.handle.net/11304/6e1424cc-8927-11e4-ac7e-860aa0063d1f
2. Data Scientist Report. Crowdflower (2017). https://visit.crowdflower.com/WC-2017-Data-Science-Report_LP.html
3. Schloss, P.D.: Identifying and overcoming threats to reproducibility, replicability, robustness, and generalizability in microbiome research. mBio **9**(3), e00525–18 (2018). https://doi.org/10.1128/mBio.00525-18
4. Gorgolewski, K.J., Poldrack, R.A.: A practical guide for improving transparency and reproducibility in neuroimaging research. PLoS Biol. **14**(7), e1002506 (2016). https://doi.org/10.1371/journal.pbio.1002506
5. Mons, D.: Data Stewardship for Open Science: Implementing FAIR Principles, 1st edn. Chapman and Hall, Boca Raton (2018)
6. Research Data Alliance. https://www.rd-alliance.org
7. Implementation Roadmap for the European Open Science Cloud (2018) http://www.esfri.eu/ri-world-news/implementation-roadmap-european-open-science-cloud
8. New Models of Data Stewardship. NIH Data Commons. https://commonfund.nih.gov/commons
9. How expensive is FAIR compliance and how expensive is it to not be FAIR compliant. RDA 11th Plenary BoF meeting (2018). https://rd-alliance.org/how-expensive-fair-compliance-and-how-expensive-it-not-be-fair-compliant-rda-11th-plenary-bof
10. G7 Science Ministers' Communique (2017). http://www.g7.utoronto.ca/science/2017-G7-Science-Communique.pdf

11. Progress Towards the European Open Science Cloud: GO FAIR Office Established, Global ActionPlatform (2017). http://globalactionplatform.org/post/progress-towards-the-european-open-science-cloud-go-fair-office-established

12. Lannom, L.: Managing digital objects in an expanding science ecosystem (2017). https://www.rd-alliance.org/sites/default/files/CENDI-15.Nov_.17-Lannom-Final-2.pdf

13. Hughes, T.P.: Networks of Power: Electrification in Western Society 1880–1930. Johns Hopkins University Press, Baltimore (1983)

14. Wittenburg, P., Strawn, G.: Common Patterns in Revolutionary Infrastructures and Data. US National Academy of Sciences (2018). https://www.rd-alliance.org/sites/default/files/Common_Patterns_in_Revolutionising_Infrastructures-final.pdf

15. International DAITF Workshop at the ICRI 2012 Conference (2012). http://www.icri2012.dk/www.ereg.me/ehome/index06e1.html

16. Digital Object Architecture. https://www.dona.net/digitalobjectarchitecture

17. Berg-Cross, G., Ritz, R., Wittenburg, P.: Research data alliance, data foundation & terminology group core terms and model (2016). http://hdl.handle.net/11304/5d760a3e-991d-11e5-9bb4-2b0aad496318

18. The FAIR Data Principles. FORCE11 (2016). https://www.force11.org/group/fairgroup/fairprinciples

19. RDA Data Fabric. https://www.rd-alliance.org/group/data-fabric-ig.html

20. DONA Foundation. https://www.dona.net/

21. Jointly designing a data FAIRPORT. Lorentz Center faculty of Science of Leiden University, Leiden The Netherlands (2014). https://www.lorentzcenter.nl/lc/web/2014/602/info.php3?wsid=602

22. FAIR Principles Explained. GO FAIR. https://www.go-fair.org/fair-principles/

23. Wilkinson, M.D., et al.: The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data **3** (2016). https://doi.org/10.1038/sdata.2016.18

24. Wilkinson, M.D., et al.: A design framework and exemplar metrics for FAIRness. Sci. Data **5**, 180118 (2018). https://doi.org/10.1038/sdata.2018.118

25. Wilkinson, M.D., et al.: Evaluating FAIR-compliance through an objective, automated, community-governed framework. bioRxiv 418376 (2018) https://doi.org/10.1101/418376

26. Data Type Registries Recommendations (Endorsed). Research Data Alliance. https://www.rd-alliance.org/group/data-type-registries-wg/outcomes/data-type-registries

27. GO FAIR International Support and Coordination Office (GFISCO). http://go-fair.org

28. Kahn, R., Wilensky, R.: A framework for distributed digital object services (1995). http://www.cnri.reston.va.us/k-w.html

29. Kahn, R., Wilensky, R.: A framework for distributed digital object services. Int. J. Digit. Libr. **6**(2), 115–123 (2006). https://doi.org/10.1007/s00799-005-0128-x. https://www.doi.org/topics/2006_05_02_Kahn_Framework.pdf

30. Wittenburg, P., Strawn, G., Mons, B. et al.: Digital objects as drivers towards convergence in data infrastructures (2019). http://doi.org/10.23728/b2share.b605d85809ca45679b110719b6c6cb11

31. Object Oriented Programming. https://de.wikipedia.org/wiki/Objektorientierte_Programmierung

32. Liskov, B., Zilles, S.N.: Programming with abstract data types. In: ACM SIGPLAN Notices, vol. 9, no. 4, pp. 50–59. ACM, New York (1974)

33. Objet Storage. https://en.wikipedia.org/wiki/Object_storage

34. Handle System. https://en.wikipedia.org/wiki/Handle_System

35. GEDE Digital Object Topic Group. https://rd-alliance.org/group/gede-group-european-data-experts-rda/wiki/gede-digital-object-topic-group

36. GEDE Workshop on Digital Objects. https://rd-alliance.org/group/gede-group-european-data-experts-rda/wiki/first-gede-do-workshop-september-18

37. American Geophysical Union's Enabling FAIR Data Project. http://www.copdess.org/enabling-fair-data-project/
38. Supporting FAIR Exchange of Chemical Data Through Standards Development. GO FAIR Chemistry Implementation Network (ChIN). https://iupac.org/event/supporting-fair-exchange-chemical-data-standards-development/
39. Wilkinson, M.D., et al.: Interoperability and FAIRness through a novel combination of Web technologies. PeerJ Comput. Sci. **3**, e110 (2017). https://doi.org/10.7717/peerj-cs.110
40. FAIR Data Point Specification. https://github.com/DTL-FAIRData/FAIRDataPoint/wiki/FAIR-Data-Point-Specification
41. 500,000 data scientists needed in European open research data. JoinUp Platform, European Commission (2016). https://joinup.ec.europa.eu/news/500000-data-scientists-need
42. The Carpentries. https://carpentries.org
43. GO FAIR Implementation Networks. https://www.go-fair.org/implementation-networks/
44. GO FAIR Current Implementation Networks. https://www.go-fair.org/implementation-networks/overview/