Related to other papers in this special issue	25 (p246); 26 (p257); 27 (p264)
Addressing FAIR principles	F, A, I, R

GO FAIR Brazil: A Challenge for Brazilian Data Science

Luana Sales[1], Patrícia Henning[2†], Viviane Veiga[3], Maira Murrieta Costa[4], Luís Fernando Sayão[5], Luiz Olavo Bonino da Silva Santos[6] & Luís Ferreira Pires[7]

[1]Instituto Brasileiro em Informação em Ciência e Tecnologia, Rio de Janeiro - RJ, 22290-160, Brazil

[2]Universidade Federal do Estado do Rio de Janeiro, Rio de Janeiro - RJ, 22290-240, Brazil

[3]Fundação Oswaldo Cruz, Rio de Janeiro - RJ, 21040-900, Brazil

[4]Ministério da Ciência, Tecnologia, Inovações e Comunicações, Esplanada dos Ministérios, Bloco R - Brasília, DF, CEP 70067-900, Brazil

[5]Comissão Nacional de Energia Nuclear, Rua Gal. Severiano, nº 90; Bairro: Botafogo; CEP 22290-901 - Rio de Janeiro, Brazil

[6]GO FAIR International Support & Coordination Office (GFISCO), Leiden 2333 AA, The Netherlands

[7]University of Twente, Enschede 7522 NH, The Netherlands

Keywords: FAIR principles; GO FAIR; GO FAIR Brazil; Open Science; Research data

Citation: L. Sales, P. Henning, V. Veiga, M.M. Costa, L.F. Sayão, L.O. Bonino da Silva Santos, ... & L.F. Pires. GO FAIR Brazil: A challenge for Brazilian data science. Data Intelligence 2(2020), 238–245. doi: 10.1162/dint_a_00046

ABSTRACT

The FAIR principles, an acronym for Findable, Accessible, Interoperable and Reusable, are recognised worldwide as key elements for good practice in all data management processes. To understand how the Brazilian scientific community is adhering to these principles, this article reports Brazilian adherence to the GO FAIR initiative through the creation of the GO FAIR Brazil Office and the manner in which they create their implementation networks. To contextualise this understanding, we provide a brief presentation of open data policies in Brazilian research and government, and finally, we describe a model that has been adopted

† Corresponding author: Patrícia Henning (E-mail: henningpatricia@gmail.com, ORCID: 0000-0003-0739-6442).

for the GO FAIR Brazil implementation networks. The Brazilian Institute of Information in Science and Technology is responsible for the GO FAIR Brazil Office, which operates in all fields of knowledge and supports thematic implementation networks. Today, GO FAIR Brazil-Health is the first active implementation network in operation, which works in all health domains, serving as a model for other fields like agriculture, nuclear energy, and digital humanities, which are in the process of adherence negotiation. This report demonstrates the strong interest and effort from the Brazilian scientific communities in implementing the FAIR principles in their research data management practices.

1. INTRODUCTION

Contemporary science has restored research data to its historical leading role, leaving behind the idea of data as a mere by-product of scientific activities and designating it a primary source of new knowledge. At the same time, the international scenario shows that the advance of science—in all fields of knowledge—is strongly linked to the reuse of research data, which imposes the challenge of managing it in a way that preserves it as long as necessary for reuse in additional research. However, as numerous studies have indicated, 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, procedures for which existing data stewardship practices are highly inefficient [1].

In these terms, the growing interest in data generated by research activities has increased the demand of the researchers, academic, industry and government institutions for a better organisational and technological infrastructure, higher-level human resources, and public policies that support the analysis, management and sustainability of these information resources [2].

In order for data to be interpreted and reused, it is essential to consider the heterogeneous nature and idiosyncrasies of each disciplinary area, as well as the specific technological resources of each field.

It is therefore agreed that dynamic management must go beyond safe storage and access to such data; it must be strongly related to the adoption of best practices throughout the research life cycle.

Effective management takes place through the proper structuring of data, the adoption of quality metadata, the potential for interoperability among them, compliance with legal and ethical issues (e.g. proper licensing) and other practices that are strongly associated with what has been designated by the scientific community as the FAIR principles.

The FAIR principles, an acronym for Findable, Accessible, Interoperable and Reusable, were first discussed in 2014 at the Lorentz workshop "Jointly Designing a Data Fairport" and were later published in 2016. The 15 principles and sub-principles present guiding elements for the process of data management that increase the likelihood of data being more easily findable, accessible, interoperable and reusable. They also give more research transparency and help data sharing possible [3].

Discussions on this subject focus on providing the highest degree of reuse for scientific data through the adoption of persistent identifiers, standards, metadata, controlled vocabularies and ontologies that provide precise meaning to data and other associated objects [4].

In addition to the scientific community, government communities and private initiatives in various countries are joining the group that advocates the dissemination and implementation of FAIR principles. In this sense, the Brazilian community is also associated with the group through the GO FAIR initiative, is a consortium committed which has an international, bottom-up approach for the implementation of the global Internet of FAIR Data and Services (IFDS), a shared global environment for data-driven research and innovation [5].

Based on documentary research, we provide a brief analysis of Brazilian policies for sharing of research and government data; we address Brazilian participation in the GO FAIR initiative, with the role of accelerating convergence towards a FAIR data infrastructure; and as a result, we present the model used by GO FAIR Brazil for its implementation networks, along with some products that are already in progress.

2. A BRIEF OVERVIEW OF THE OPENING OF RESEARCH DATA IN BRAZIL

In Brazil, access to information and the opening of government and research data is not a new issue. There are a number of formal initiatives focused in this direction. Access to information and the opening of government data were strengthened after the creation of the Law on Access to Information (LAI), no. 12.527, on 18 November 2011 [6]. This law stated, among other things, that it is an obligation of the State to ensure the right of access to information, that publication is considered the expected precept and secrecy is the exception, that information must be readily accessible with transparency, that management must be done, and that access be widely provided. At an international level, Brazil has been participating in the Open Government Partnership (OGP) since its creation on 20 September 2011[7] as one of the eight founding members that signed the Open Government Declaration and presented their National Action Plans.

With regard to the opening of research data, the National Action Plan on Open Government was also created in 2011, recommending that data, whether processed or not, could be used for the production and transmission of knowledge contained in any supported format. Thereafter, in 2012, the National Open Data Infrastructure (INDA) [8] was created with the objective of answering LAI's statements and complying with Brazil's commitments under the National Action Plan 2018 - 2020 on Open Government [9]. The Open Data Policy of the Federal Executive Branch was established in 2016 and promotes, among several other actions, the publication of data from federal public administration databases, the enhancement of a transparent culture, the exchange of data between government agencies, the development of scientific research, and the promotion of technological development and innovation in the public and private sectors, along with fostering new business and promoting the sharing of information technology resources [10].

In 2016, the Brazilian Institute of Information in Science and Technology (IBICT) – a research branch of the Ministry of Science, Technology, Innovation, and Communications (MCTIC) – launched the "Manifesto

of Open Access to Brazilian Research Data for Citizen Science", this being one of the precursors to the guidelines for the opening of research data in Brazil.

In regard to funding agencies, since 2014, the São Paulo Research Foundation (FAPESP) has been requesting, for research on some special projects, details for how data management will be conducted, including specifications for the generated data types, format, preservation modes, and what care will be taken regarding privacy and ethics issues. Today FAPESP requires data management plans for thematic projects; these plans should be gradually extended to other project types [11].

The MCTIC created a "Working Group" in 2018 to draft the National Policy for Open Science, which will give the guidelines for the national policy for research data management in Brazil [12]. This group is still working, with an expected final version to be delivered at the end of the first semester of 2019. Concerned with the advancement of world discussions on sharing and reusing data, on 10 December 2018, the MCTIC promoted an event for the Brazilian scientific community where an explanation of the commitment to Innovation and Open Government in Science and the Fourth OGP-Brazil National Open Data Action Plan took place. At the same event, with the support of MCTIC, the IBICT officially launched the GO FAIR Brazil Declaration [13], highlighting the emergence of the theme and the country's attention to open science – especially access to scientific and governmental data.

3. AN IMPLEMENTATION MODEL FOR GO FAIR BRAZIL

The understanding of GO FAIR Brazil is linked, initially, to the Brazilian connection to the international GO FAIR initiative. This is an international initiative that adopts a bottom-up approach turned to FAIR principles practices. GO FAIR was launched in late 2017 and – with the support of the governments of the Netherlands, Germany, and France – started the GO FAIR International Support and Coordination Office. More recently, other regional and national offices have emerged outside Europe, including one in Brazil and another in the United States.

GO FAIR prioritises making all sorts of data, which are fragmented and disconnected, more easily findable, accessible, interoperable and reusable (FAIR), thus facilitating their utility for machines and people. GO FAIR operates through voluntary stakeholder participation, attempting to reach a "critical mass" of users committed to a set of absolute minimal technology specifications with unrestricted room to innovate. GO FAIR is stakeholder governed and works with researchers from specialised knowledge domains along with policy bodies, publishers, repositories, and funding agencies.

Figure 1. The three pillars of GO FAIR implementation networks (IN)[①]

GO FAIR implementation networks (INs) are sustained by three pillars: GO BUILD, GO TRAIN and GO CHANGE. GO BUILD focuses on the technological aspects of the Internet for FAIR Data and Services. GO TRAIN is for creating a scalable framework that is used in higher education programs and throughout industry. GO CHANGE is for supporting and coordinating a systemic culture change that transforms existing data management practices into the respected profession of data stewardship [1]; it also looks at changing and adapting the implementation of FAIR Guidelines to communities of practice.

The number of implementation networks at the GO FAIR initiative is rapidly growing, encouraging the creation of regional implementation networks that work in a participatory and collaborative way to strategically follow the FAIR principles. GO FAIR Brazil is one of the regional offices of GO FAIR; therefore, it has the responsibility to disseminate, support and coordinate activities in Brazil related to the adoption of strategies for implementing the FAIR principles, as defined by the GO FAIR initiative. GO FAIR Brazil operates in all areas of knowledge and is thematically structured. Its first meeting took place on 25 September 2018 in São Paulo during the 20th anniversary of the Scientific Electronic Library Online (SciELO) Network [14]. A number of stakeholders representing Brazilian universities and research institutes were present. It was at this meeting that was established the IBICT as the coordinator of GO FAIR Brazil. However, their official launch, to the scientific community, happened on 10 December 2018 during an event sponsored by MCTIC. Its work is committed to the following activities:

- supporting and coordinating INs, according to their specific objectives, and adopting the strategies to implement the FAIR principles approved by the GO FAIR initiative
- supporting and coordinating INs that are willing to define strategies for implementing FAIR principles

[①] https://www.go-fair.org/go-fair-initiative/.

- systematising the existing guidelines – and those established by scientific networks – besides recommending they are aligned with FAIR principles
- developing mechanisms to disseminate the guidelines defined by INs
- maintaining communication with the GO FAIR International Support and Coordination Office

The approach used for the creation of GO FAIR Brazil is based on supporting the needs of INs. Currently, in GO FAIR Brazil, there are other INs (e.g. health, farming, nuclear energy, and digital humanities) focusing on specific thematic and community issues. It was outlined by eight actions based on the three pillars of the GO FAIR Implementation Network: build, train and change. These actions were performed by thematic networks. The actions include the following: meetings with potential partners, drafting of manifestos, participation in events, development of the FAIR Data Management Plan (DMP), surveying metadata standards and vocabulary, policy setting, human resources training, and development of FAIR services.

The first IN of GO FAIR Brazil is in the health field under the coordination of the Institute of Communication and Scientific Information and Technology in Health (ICICT) of the Oswaldo Cruz Foundation (Fiocruz), referred to as GO FAIR Brazil – Health. It is a thematic network responsible for developing strategies for implementing FAIR principles in the field of health. This network already has the support and participation of several institutions in the areas of public health, health surveillance, information and communication in health, the history of the cultural heritage of science and health, oncology, nursing, violence and health, and professional health education.

Currently, this network has fifteen participating institutions intent on working in an articulate and collaborative way with its members, in compliance with the FAIR principles, in the field of health. The primary focus of this network is promoting the sharing and reuse of health data. The health research communities have the support of GO FAIR Brazil – Health in creating sub-networks to support the development of interoperability infrastructures, specific data formats, adoption of metadata standards, use of controlled vocabularies, and ontologies for health sciences to strengthen the semantic Web in health.

To support this, several activities are being conducted: drafting a FAIR DMP, which will comply to the management requirements of Fiocruz's domains of research data; evaluating some DMPs' tools; studying some of the main national and international data repositories to verify their alignment with FAIR principles; conducting a survey on the perception of Fiocruz's researchers regarding both the sharing of their research data and the types of data used in their research; developing an ontology of Fiocruz's domain research data; and promoting meetings, courses and workshops to boost and disseminate FAIR principles among its members.

Also, in regard to the GO FAIR Brazil – Health network, the creation of the GO FAIR Brazil – Health Nursing sub-network is in its structuring phase, which is being conducted by the Federal University of Rio de Janeiro State. As for other areas and future projects, there is discussion of establishing the GO FAIR Brazil Nuclear Energy Network under the responsibility of the National Nuclear Energy Commission (CNEN), establishing the GO FAIR Brazil Agricultural Network under the responsibility of the Brazilian

Agricultural Research Corporation (EMBRAPA), and establishing GO FAIR Brazil Digital Humanities under the responsibility of the Teaching and Research Coordination (COEPE) of the IBICT.

4. FINAL REMARKS

Despite difficulties in interpreting and implementing FAIR principles in data management and in its repositories – even in countries with high levels of technological and financial resources [15,16] – Brazil has increased its efforts to adopt FAIR principles as a guiding element in the management and sharing of data.

The advantages of adopting FAIR principles through the actions of GO FAIR Brazil go beyond the most noticeable benefits, such as the development of FAIR data management infrastructures, changes in the culture of sharing and the reuse of research data sets by Brazilian researchers, new approaches by information professionals in interacting with researchers, and, of course, a more universal perspective in setting sustainability-related policies and improving reward schemes for researchers who publish their data sets. It is also necessary to consider Brazil's inclusion in the international forums of standardised treatment of research data, beginning with strategic areas such as health, agriculture, and energy that contribute to shaping behaviour, financing, and research policies in the country.

All work being carried out by GO FAIR Brazil in adopting FAIR principles has been highlighted in international forums. For instance, the participation at the first-annual, international meeting of GO FAIR Implementation Networks, held on January 15, 2019, at the GO FAIR International Support and Coordination Office location in Leiden, the Netherlands; the participation at the upcoming ER 2019 - 38th International Conference on Conceptual Modelling, where the GO FAIR Brazil is one of the co-organisers of the Workshop on Conceptual modelling, Ontologies, and Metadata Management for FAIR Data; and the collaboration of three Brazilian institutions in the effort that has been developing "The FAIR Funder: A pilot programme to make it easy for funders to require and for grantees to produce FAIR Data" [17].

The interest of the Brazilian communities in the implementation of the FAIR principles is evident in the several activities developed in a short period of time. This indicates a productive future on behalf of good research data management, devoted to studies, development of infrastructures, and proposition of solutions for the FAIR principles implementation in research activities in Brazil.

AUTHOR CONTRIBUTIONS

All authors contributed significantly and equally to this article.

REFERENCES

[1] E. Schultes, G. Strawn, B. Mons. Ready, set, GO FAIR: Accelerating convergence to an Internet of FAIR Data and Services. In: Proceedings of the XX International Conference "Data Analytics and Management in Data Intensive Domains" (DAMDID/ RCDL'2018), Moscow, Russia, 9-12 October 2018. Available at: http://ceur-ws.org/Vol-2277/paper07.pdf.

[2] L.F. SayÃo O papel dos repositórios digitais na curadoria de dados. In: Vechiato, Fernando Luiz et al. (Org.). Repositórios digitais: teoria e prática. Curitiba: Editora da UTFPR, 2017, v. 1, pp. 143-166.

[3] M.D. Wilkinson, M. Dumontier, I. Jan Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg … & B. Mons. The FAIR guiding principles for scientific data management and stewardship. Nature 3 (2016), 160018. doi:10.1038/sdata.2016.18.

[4] P. Hening, C.J.S. Ribeuro, L. Sales, J. Moreira & L.O. Bonino da Silva Santos. Desmistificando os Princípios FAIR: Conceitos, métricas, tecnologias e aplicações inseridas no ecossistema dos dados FAIR. XIX Eencontro Nacional De Pesquisa Em CiÊncia Da InformacÃo – Enancib Nancib 2018. Available at: http://enancib.marilia.unesp.br/index.php/XIXENANCIB/xixenancib/paper/viewFile/1475/1717.

[5] Internet of FAIR Data & Services (IFDS). Available at: https://www.go-fair.org/resources/internet-fair-data-services.

[6] Brazilian Law on Access to Information (LAI). Available at: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm.

[7] Open Government Partnership (OGP). Available at: https://www.opengovpartnership.org/members/brazil.

[8] National Open Data Infrastructure (INDA). Available at: https://www.governodigital.gov.br/transformacao/cidadania/dados-abertos/inda-infraestrutura-nacional-de-dados-abertos.

[9] Brazilian National Action Plan 2018-2020. Available at: https://www.opengovpartnership.org/documents/brazil-national-action-plan-2018-2020/.

[10] The Open Data Policy of the Federal Executive Branch. Available at: https://www.cgu.gov.br/noticias/2018/12/cgu-publica-novo-plano-de-dados-abertos-e-preve-abertura-de-11-bases-proprias.

[11] SÃO Paulo Research Foundation (FAPESP). Available at: http://www.fapesp.br/gestaodedados/.

[12] Ministry of Science, Technology, Innovation and Communications (MSTIC) - Working Group. Available at: http://www.mctic.gov.br/mctic/opencms/legislacao/portarias/Portaria_MCTIC_n_6072_de_23112018.html.

[13] GO FAIR Brazil Declaration. Available at: https://www.go-fair.org/wp-content/uploads/2019/06/Declaration_GO-FAIR-Brazil_Jun2019.pdf.

[14] SCIELO. 20 Years of SCIELO. Available at: https://www.scielo20.org/.

[15] A. Dunning, M. Smaele & J. BÖhmer. Are the FAIR data principles fair? In: 12th International Digital Curation Conference. 2017, Edinburgh. Available at: https://zenodo.org/record/321423#.W4QjbegzZaR.

[16] FAIR data advanced use cases: From principles to practice in the Netherlands. 2018. Available at: https://zenodo.org/record/1250535#.WyI3RkiFNaQ.

[17] FAIR funder collaboration: A pilot programme to make it easy for funders to require and for grantees to produce FAIR Data. Available at: https://arxiv.org/ftp/arxiv/papers/1902/1902.11162.pdf.