

GO FAIR Brazil: A Challenge for Brazilian Data Science postprint

Authors: Sales, Luana, Henning, Patricia, Veiga, Viviane, Costa, Maira Murieta, Sayao, Luis Fernando, da Silva Santos, Luiz Olavo Bonino, Pires, Luis Ferreira, Henning, Patricia

Date: 2022-11-18T00:00:00+00:00

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 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.

Full Text

Preamble

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 Sales¹, Patrícia Henning^{2†}, Viviane Veiga³, Maira Murrieta Costa⁴, Luís Fernando Sayão⁵, Luiz Olavo Bonino da Silva Santos⁶ & Luís Ferreira Pires⁷

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

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

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

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

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

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

⁷University 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 on Brazilian adherence to the GO FAIR initiative through the creation of the GO FAIR Brazil Office and its approach to establishing implementation networks. To contextualise this understanding, we provide a brief overview of open data policies in Brazilian research and government, and describe a model that has been adopted 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 across all fields of knowledge and supports thematic implementation networks. Currently, GO FAIR Brazil-Health is the first active implementation network in operation, working across all health domains and serving as a model for other fields such as agriculture, nuclear energy, and digital humanities, which are in the process of adherence negotiation. This report demonstrates the strong interest and effort from Brazilian scientific communities in implementing FAIR principles in their research data management practices.

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

1. INTRODUCTION

Contemporary science has restored research data to its historical leading role, leaving behind the notion of data as a mere by-product of scientific activities and designating it as a primary source of new knowledge. Simultaneously, the international landscape shows that scientific advancement across all fields is strongly linked to the reuse of research data, which imposes the challenge of managing data in ways that preserve it for as long as necessary to enable such reuse. However, as numerous studies have indicated, data scientists in both 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].

This growing interest in research-generated data has increased demands from researchers, academic institutions, industry, and government for better organisational and technological infrastructure, higher-level human resources, and public policies that support the analysis, management, and sustainability of these information resources [2]. 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; it must be strongly related to the adoption of best practices throughout the research life cycle.

Effective management occurs through proper data structuring, adoption of quality metadata, ensuring interoperability, compliance with legal and ethical issues (e.g., proper licensing), and other practices strongly associated with what the scientific community has designated as the FAIR principles. The FAIR principles—Findable, Accessible, Interoperable and Reusable—were first discussed in 2014 at the Lorentz workshop “Jointly Designing a Data Fairport” and later published in 2016. The 15 principles and sub-principles provide guiding elements for data management processes that increase the likelihood of data being more easily findable, accessible, interoperable and reusable, while also enhancing research transparency and facilitating data sharing [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]. Beyond the scientific community, government bodies and private initiatives in various countries have joined the group advocating for the dissemination and implementation of FAIR principles. In this context, the Brazilian community has also associated itself with this movement through the GO FAIR initiative, a consortium committed to an international, bottom-up approach for implementing 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 research and government data; address Brazilian participation in the GO FAIR initiative and its role in accelerating convergence towards a FAIR data infrastructure; and present the model used by GO FAIR Brazil for its implementation networks, along with some products 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 are not new issues. Numerous formal initiatives have focused on this direction. Access to information and government data opening were strengthened after the creation of the Law on Access to Information (LAI), no. 12.527, on 18 November 2011 [6]. This law established, among other provisions, that it is an obligation of the State to ensure the right of access to information, that publication is the expected norm and secrecy is the exception, that information must be readily accessible with transparency, and that management must facilitate widespread access. At the international level, Brazil has participated 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.

Regarding research data opening, the National Action Plan on Open Government was also created in 2011, recommending that data—whether processed or not—could be used for knowledge production and transmission in any supported format. Thereafter, in 2012, the National Open Data Infrastructure (INDA) [8] was created to respond to LAI's mandates and comply 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 actions, the publication of data from federal public administration databases, enhancement of a transparent culture, exchange of data between government agencies, development of scientific research, promotion of technological development and innovation in public and private sectors, and fostering of new business and 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,” one of the precursors to guidelines for opening research data in Brazil.

Regarding funding agencies, since 2014, the São Paulo Research Foundation (FAPESP) has been requesting, for research on some special projects, details on how data management will be conducted, including specifications for generated data types, format, preservation modes, and measures regarding privacy and

ethics. Today, FAPESP requires data management plans for thematic projects, with plans to gradually extend this requirement to other project types [11].

The MCTIC created a “Working Group” in 2018 to draft the National Policy for Open Science, which will provide guidelines for the national policy on research data management in Brazil [12]. This group is still working, with an expected final version to be delivered by the end of the first semester of 2019. Concerned with advancing global discussions on data sharing and reuse, on 10 December 2018, the MCTIC promoted an event for the Brazilian scientific community to explain the commitment to Innovation and Open Government in Science and the Fourth OGP-Brazil National Open Data Action Plan. At the same event, with MCTIC support, IBICT officially launched the GO FAIR Brazil Declaration [13], highlighting the emergence of this 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 initially linked to Brazil’s connection to the international GO FAIR initiative. This international initiative adopts a bottom-up approach focused on FAIR principles practices. GO FAIR was launched in late 2017 and, with support from the governments of the Netherlands, Germany, and France, established 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 fragmented and disconnected data 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: see original paper]

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

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 creates a scalable framework used in higher education programs and throughout industry. GO CHANGE supports and coordinates a systemic culture change that transforms existing data management practices into the respected profession of

data stewardship [1]; it also addresses changing and adapting FAIR Guidelines implementation to communities of practice.

The number of implementation networks in 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 FAIR principles. GO FAIR Brazil is one of the regional offices of GO FAIR and has the responsibility to disseminate, support, and coordinate activities in Brazil related to the adoption of strategies for implementing FAIR principles as defined by the GO FAIR initiative. GO FAIR Brazil operates across 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], with numerous stakeholders representing Brazilian universities and research institutes present. At this meeting, IBICT was established as the coordinator of GO FAIR Brazil. Its official launch to the scientific community occurred on 10 December 2018 during an event sponsored by MCTIC. Its work is committed to: supporting and coordinating INs according to their specific objectives and adopting FAIR implementation strategies approved by the GO FAIR initiative; systematising existing guidelines—and those established by scientific networks—while recommending alignment with FAIR principles; developing mechanisms to disseminate guidelines defined by INs; and maintaining communication with the GO FAIR International Support and Coordination Office.

The approach used for GO FAIR Brazil creation is based on supporting IN needs. Currently, GO FAIR Brazil includes other INs (e.g., health, farming, nuclear energy, and digital humanities) focusing on specific thematic and community issues. These are outlined by eight actions based on the three pillars of the GO FAIR Implementation Network: build, train, and change. These actions are performed by thematic networks and include: meetings with potential partners, drafting of manifestos, participation in events, development of FAIR Data Management Plans (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, coordinated by 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. This thematic network is responsible for developing strategies to implement FAIR principles in health. The network already has support and participation from several institutions in public health, health surveillance, information and communication in health, history of cultural heritage of science and health, oncology, nursing, violence and health, and professional health education.

Currently, this network has fifteen participating institutions working in an articulated and collaborative manner with its members to comply with FAIR principles in health. The primary focus is promoting health data sharing and reuse. Health research communities have GO FAIR Brazil -Health support in

creating sub-networks to develop interoperability infrastructures, specific data formats, adoption of metadata standards, use of controlled vocabularies and ontologies for health sciences, and to strengthen the semantic Web in health.

To support this, several activities are being conducted: drafting a FAIR DMP that complies with Fiocruz' s research data management requirements; evaluating DMP tools; studying major national and international data repositories to verify their alignment with FAIR principles; surveying Fiocruz researchers' perceptions regarding both sharing 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 members.

Regarding the GO FAIR Brazil -Health network, the GO FAIR Brazil -Health Nursing sub-network is currently in its structuring phase, being conducted by the Federal University of Rio de Janeiro State. For other areas and future projects, discussions are underway to establish the GO FAIR Brazil Nuclear Energy Network under the National Nuclear Energy Commission (CNEN), the GO FAIR Brazil Agricultural Network under the Brazilian Agricultural Research Corporation (EMBRAPA), and GO FAIR Brazil Digital Humanities under the Teaching and Research Coordination (COEPE) of IBICT.

4. FINAL REMARKS

Despite difficulties in interpreting and implementing FAIR principles in data management and 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 data management and sharing. The advantages of adopting FAIR principles through GO FAIR Brazil actions extend beyond the most noticeable benefits, such as developing FAIR data management infrastructures, changing the culture of sharing and reuse of research datasets by Brazilian researchers, creating new approaches for information professionals interacting with researchers, and establishing a more universal perspective in setting sustainability-related policies and improving reward schemes for researchers who publish their datasets. It is also necessary to consider Brazil' s inclusion in international forums for 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, including: participation at the first annual international meeting of GO FAIR Implementation Networks, held on 15 January 2019 at the GO FAIR International Support and Coordination Office in Leiden, the Netherlands; participation at ER 2019 - 38th International Conference on Conceptual Modelling, where GO FAIR Brazil co-organised the Workshop on Conceptual Modelling, Ontologies, and Metadata Management for

FAIR Data; and collaboration of three Brazilian institutions in 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 Brazilian communities in implementing FAIR principles is evident in the numerous activities developed in a short period, indicating a productive future for good research data management, devoted to studies, infrastructure development, and solution proposals for FAIR principles implementation in Brazilian research activities.

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. Ribeiro, 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 Encontro Nacional De Pesquisa Em Ciência Da Informação -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/_ato/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-abertos/inda-infraestrutura-nacional-de-dados-abertos>.
- [9] Brazilian National Action Plan 2018-2020. Available at: <https://www.opengovpartnership.org/documents/brazilian-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}
- [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#.W4QjbeZzZaR>.
- [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>.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv – Machine translation. Verify with original.