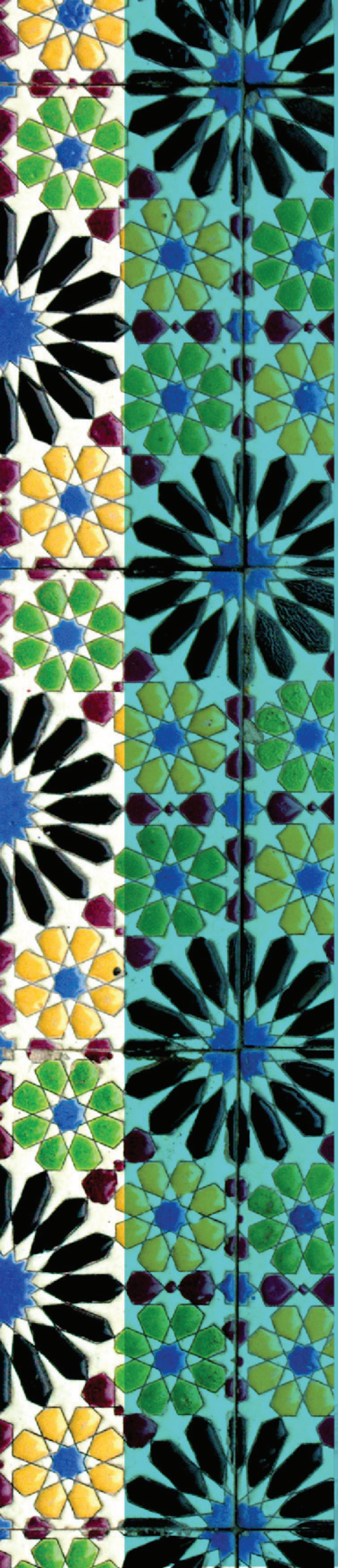


Oswaldo Cruz Foundation

120 ANOS
PATRIMÔNIO DA SOCIEDADE BRASILEIRA

Policy on Management, Sharing and Opening of Research Data

Principles and Guidelines

Rio de Janeiro
2020

Ministério da Saúde

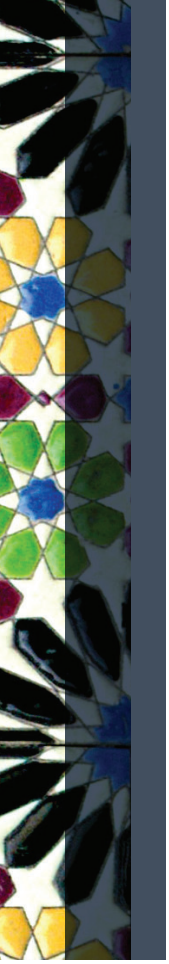
FIOCRUZ
Fundação Oswaldo Cruz

Oswaldo Cruz Foundation

Policy on Management, Sharing and Opening of Research Data

Principles and Guidelines

Rio de Janeiro
2020



OSWALDO CRUZ FOUNDATION

Presidency
Nísia Trindade Lima

Vice-Presidency of Education, Information and Communication
Cristiani Vieira Machado

Vice-Presidency of Environment, Health Care and Health Promotion
Marco Antonio Carneiro Menezes

Vice-Presidency of Management and Institutional Development
Mario Santos Moreira

Vice-Presidency of Research and Biological Collections
Rodrigo Correa de Oliveira

Vice-Presidency of Production and Innovation in Health
Marco Aurelio Krieger

Ministério da Saúde

FIOCRUZ
Fundação Oswaldo Cruz

Oswaldo Cruz Foundation

Policy on Management, Sharing and Opening of Research Data

Principles and Guidelines

Rio de Janeiro
2020

OSWALDO CRUZ FOUNDATION

Technical elaboration - Working Group to review this policy
Alexandro Rodrigues Pinto
Ana Beatriz Aguiar
Anne Clinio
Bethânia de Araújo Almeida
Flávia Tavares Silva Elias
Márcia Luz da Motta,
Maria de Fátima M. Martins
Paula Xavier do Santos (Coordination)
Vanessa de Arruda Jorge

Legal review
Allan Rocha de Souza e Francisco José Tavares do Nascimento

Graphic Design and Publishing
Airton Santos

Final review
Cristiani Vieira Machado e Josué Laguardia

English version
Annabella Blyth

The text of this Policy on Management, Sharing and Opening of Research Data is the outcome of an institutional collective construction. The text was submitted to internal public consultation, to the Working Group on Open Science Fiocruz, and approved at the meeting of the Deliberative Council of Fiocruz held on September 24, 2020.

Catalog sheet prepared by the Information and Communication Coordination of the Vice-Presidency for Education, Information and Communication (CINCO / VPEIC)

F981p Oswaldo Cruz Foundation. Vice-Presidency of Education, Information and Communication.
Policy on Management, Sharing and Opening of Research Data: principles and guidelines. - Rio de Janeiro: Fiocruz/VPEIC, 2020.
19 p.; 31 cm.
ISBN: 978-65-87063-02-7

1 .Open Science. 2. Open Data. 3. Research data. 4. Data for Research. 5. Sharing Data. 6. Institutional Repository. 7. Open Science Policy.
I. Title.

CDD 020

This work is guided by the Policy on Open Access to Knowledge, which seeks to ensure that the society has free, public and open access to the full contents of every intellectual work produced by Fiocruz.
The contents of this policy may be used for all non-commercial purposes, as long as the authors' moral rights are respected and reserved.

©2020 Fundação Oswaldo Cruz
Av. Brasil, 4365 – Manguinhos
CEP 21045-900 - Rio de Janeiro, RJ, Brasil.
Tel. 55 21 2598-4242
http://www.fiocruz.br

SUMMARY

1. INTRODUCTION ..... pág. 7

2. PRINCIPLES AND GUIDELINES ..... pág. 11

2.1 PRINCIPLE 1 – PUBLIC INTEREST ..... pág. 11

2.1.1 Guidelines ..... pág. 12

2.2 PPRINCIPLE 2 – MANAGEMENT, SHARING AND OPENING OF RESEARCH DATA ..... pág. 13

2.2.1 Diretrizes ..... pág. 14

2.3 PRINCIPLE 3 – REGULATORY FRAMEWORK ..... pág. 15

2.3.1 Guidelines ..... pág. 16

2.4 PRINCIPLE 4 – DEVELOPMENT OF CAPACITIES AND SUSTAINABILITY ..... pág. 17

2.4.1 Guidelines ..... pág. 18

1. INTRODUCTION

Open Science is an international movement that proposes structural changes on how scientific knowledge is produced, organized, shared, and reused, understood as a new way of making science – more collaborative, transparent and sustainable. As its pillar, Open Science recommends the sharing and opening of the process cycle of knowledge production and communication. Therefore, besides the open access to research outcomes, the term can comprise such practices as: sharing of and access to methods and data used in research, publication of research data, and even citizens' participation in the definition of research agenda priorities and in the knowledge production process.

The Oswaldo Cruz Foundation (Fiocruz) is a State institution of science and technology in health, linked to the Ministry of Health, having reached 120 years in existence in 2020. Its mission is to produce, disseminate, and share knowledge and technologies directed to the strengthening and consolidation of the Unified Health System (SUS), contributing to the promotion of health and quality of life of the Brazilian population, the reduction of social inequities, and the national innovation dynamics, having as central values the defense of the right to health and full citizenship.

Fiocruz published in 2014 the guidelines for the Policy on Open Access to Knowledge, whose implementation since then is expressed in the open and free of cost access to its scientific production through the Arca Institutional Repository[1] , in the open access to contents of

[1] Arca: https://www.arca.fiocruz.br
[2] Fiocruz Journals Portal: https://periodicos.fiocruz.br

journals published by Fiocruz, and in Open Education initiatives, such as the delivery of free of cost courses at Fiocruz Virtual Campus[3] . The adoption of other guidelines related to Open Science, supported by the Internal Congress of Fiocruz, represents the expansion of the movement for open access to scientific knowledge, consolidated during the past two decades.

Among these practices, the emphasis of this document is on management, sharing and opening of data gathered or generated during research, as an opportunity to foster the collaboration between researchers, the advancement of knowledge, and the creation of solutions that fulfil the society's needs. Research data management is a strategic activity that enhances the integrality, quality, and reproducibility of research; the institutional memory; and data reuse by the same research group and/or other groups. The process of research data management precedes the decisions related to sharing and opening of data and should consider the technical, legal and ethical aspects. On the other hand, definitions regarding sharing and opening of data should consider, in addition to these aspects, issues concerning institutional interests, national sovereignty, and asymmetries between countries and within countries, reciprocity, among others, in order to avoid that the process increases inequalities in the scientific and technological field and in the access of populations to knowledge.

In this perspective, institutional, governmental and development agencies policies, technology infrastructure of data repository, models of data management plan, capacity-building in the scientific community, levels of exclusiveness, reciprocity conditions, licenses, use and cession terms for the legal definition of the degree of data access or confidentiality, in sum, an entire scientific research ecosystem is being implemented in several countries[4]. This tendency is being strongly boosted by several stakeholders in the Science, Technology & Innovation (ST&I) system, who currently require from researchers new practices regarding management, sharing and opening of data. It is noteworthy that, recently, international funders such as the Wellcome Trust, European Commission Horizon 2020 Program, Bill & Melinda Gates Foundation, National Institutes of Health, among others, have established on their calls for proposals the mandatory requirement of Data Management Plan as a condition for grant funding, in a perspective in favor of data opening, as stated on the Horizon 2020 slogan "as open as possible, as closed as necessary". In scientific communication the changes are already perceptible, with the publication of pre-prints and the incorporation of open peer review. Furthermore, several scientific journals now require the availability of data on which papers are based. There are also new communication formats that favor the publication of data, e.g. data papers.

In Brazil, some initiatives have started to emerge, such as the requirement of a Data Management Plan in some funding modalities of Fapesp[5], the open science policy of SciELO[6],

[3] Fiocruz Virtual Campus: https://campusvirtual.fiocruz.br
[4] Livro Verde: https://www.arca.fiocruz.br/handle/icict/24117
[5] Fapesp Agency: http://www.fapesp.br/gestaodedados
[6] Scielo: https://bit.ly/3aZGkp6

the data management model of USP[7], and the inclusion of Open Science as one of the commitments of the 4th National Action Plan of Open Government Partnership[8].

However, it is necessary to have a critical approach to the movement, considering the tensions and risks engendered in the management, sharing and opening of data. The existing asymmetries between countries, such as the installed capacity for research, differences in time limits for regulatory approvals and in intellectual property legislation, among others, may reinforce inequalities in the scientific performance and result in new ways of concentration of knowledge and resources. For this reason it is necessary that a policy on management, sharing and opening of research data considers protective measures that ensure competitive prerogatives capable of valuing research and the national and institutional strategic interests.

Also noteworthy is the emergence of new forms of knowledge production that depend on the access and manipulation of large volumes of data, including the linkage between data from different sources and typologies. The use of administrative data in research, for instance, has a great potential to answer questions and generate evidence in favor of overcoming problems and the improvement of life conditions of the Brazilian society.

In the face of this scenario, since 2017 Fiocruz has been promoting a broad debate aiming at the critical understanding of Open Science with emphasis on management, sharing and opening of research data. The dialogue with the scientific community and all professionals of the institution occurred by means of technical visits to the units, Technical Chambers, Forums of Regional Units, seminars, focal groups, and culminated with a public consultation on the Reference Term that systematized a set of Principles and Guidelines, which provide guidance to management, sharing and opening of research data at Fiocruz. The internal consultation occurred from June 17 to September 10, 2019 and had 286 contributions originated in the 15 units of Fiocruz, followed by a process of consolidation of these contributions, which resulted in the guidelines of the institutional policy for management, sharing and opening of research data to be submitted to the Deliberative Council of Fiocruz.

Fiocruz's policy is based on the principle that data aimed at research, especially when funded by public resources, are goods produced in the name of the public interest. Therefore, as much as possible, such data should be made available openly, ethically, integrally, accessibly, timely and responsibly, considering the strategic policies of national scientific research, institutional interests and regulatory norms in force.

Thus, Fiocruz's policy on management, sharing and opening of research data complies with the established guidelines of the national policies and regulations in force, including ethics and integrity in research, and the protection of personal and sensitive data, such as the National Policy on Open Data of the Brazilian Executive Branch (Decree 8777/2016) and the

[7] USP: https://bit.ly/2UKwPVf

[8] 4th National Action Plan (2018-2020) of Open Government Partnership (OGP): https://www.gov.br/cgu/pt-br/governo-aberto/a-ogp/planos-de-acao/4o-plano-de-acao-brasileiro/4o-plano-de-acao-nacional_portugues.pdf

General Personal Data Protection Act (LGPD - Law 13709/2018). Copyright, industrial property and/or other rights on data and databases generated in the research process at Fiocruz must be recognized and protected. In principle, data opening should be as broad as possible, complying with legal, ethical and strategic limits concerning confidential, personal and sensitive data. The restrictions to access should be made clear and public within a time limit predefined by researchers, determined on the base of specific legislation and other ethical or other criteria of institutional and national strategic interest.

A pre-condition for this process concerns the strengthening of the process of research data management by the researchers themselves. The adequate data management comprises a series of stages – identification, processing , storage, preservation – that are relevant to allow its reuse by the researchers themselves and their research groups in the future, as well as to subsidize the groups' decision on the type, form, degree, and moment of sharing and opening data to other groups and to the general public.

In parallel with the definition of the Policy's guidelines, ongoing initiatives aim at the construction of infrastructure to enable the management, sharing and opening of data for research, namely: the creation of an institutional model of Data Management Plan; development of an Institutional Data Repository; structuring and delivery of an introductory course on Open Science (Massive Open Online Course); delivery of presential workshops on Data Management; and the formulation of a discipline that is cross-sectional to the post-graduate programs on Open Science.

The implementation of the Policy's set of initiatives will enable the adoption of new practices in the institution that contribute to reaffirm Fiocruz's strategic role in the country's scientific and technological development, the production and dissemination of knowledge aimed at the strengthening of the Unified Health System (SUS), and the improvement of the quality of health and life for the entire Brazilian population.

2. PRINCIPLES AND GUIDELINES

PRINCIPLE 1 – PUBLIC INTEREST

Data aimed at research, especially that funded by public resources, are goods produced in the name of the public interest. They should be made available openly, ethically, integrally, accessibly, timely and responsibly, considering the strategic policies of national scientific research, institutional interests, regulatory norms in force and the reciprocity of access .

Guidelines

D1.1 - Implement management, sharing and opening of research data, with the aim of making accessible datasets generated or reused in scientific research, except those with confidential restrictions justifiable according to the current national legislation.

D1.2 - Foster the opening of administrative data for research use in combination with Fiocruz's Open Data Plan and according to the current national legislation.

D1.3 - Consider as priority the public, national, and institutional interests involved in the production and use of scientific knowledge. The policy of management, sharing and opening of data comprises protective actions to avoid undue appropriation of research data, in order to protect institutional and third parties intellectual property, especially those involving the development and transfer of technology.

D1.4 - Promote the exchange of experiences on management, sharing and opening of data and their potential impacts on and benefits to Public Health, with the aim of ensuring the participation of the civil society and different groups in the field of health, such as: researchers, junior scientists, teachers and students, scientific journalists, managers of science, technology and innovation institutions, editors of scientific journals, among other interested groups.

D1.5 - Make openly available, whenever possible, research data produced at Fiocruz as a complementary strategy for social participation and political articulation performed by social movements and organized groups in the field of public health.

D1.6 - Foster the transparency of the processes of gathering, treatment and analysis of data in accessible language as a means of enhancing their reuse.

D1.7 - Promote the debate with governmental and other funding agencies for the alignment of Fiocruz's policy of management, sharing and opening of research data with the guidelines of the main development agencies.

D1.8 - Establish a model to fulfil the demands of data sharing in the face of emergencies in public health and disasters.

PRINCIPLE 2 – MANAGEMENT, SHARING AND OPENING OF RESEARCH DATA

The management of research data is a strategic activity that enhances the integrity, quality and reproducibility of research, the institutional memory, and the reuse of data. It considers the technical, legal and ethical aspects.

Guidelines

D2.1 - Provide strategies, resources and infrastructure for the management of data used in research during their entire life cycle, with the aim of ensuring the integrity, reliability, security, preservation, and quality of data, for the purpose of scientific transparency, reproducibility of outcomes and reuse in new research.

D2.2 - Provide priority support for research projects subject to requirements of funding agencies regarding research data management and to data requests by scientific journals.

D2.3 - Stimulate data management as one of the Best Practices in Scientific Research at the institutional ambit, especially in articulation with the actions of Fiocruz's Quality Management System.

D2.4 - Conduct diagnoses on data gathered, generated, and used in research developed at Fiocruz with the purpose of identifying the specificities of each field of knowledge regarding data management.

D2.5 - Use Fiocruz's Data Management Plan (PGD) as a strategic tool for the activity of research in the institution. The research areas can adapt this instrument to their specificities or to the demands of development agencies.

D2.6 – Implement the Institutional Repository of Research Data articulated with national initiatives whenever possible.

D2.7 - Recommend that data are deposited on Fiocruz's institutional repository or, upon external requirements, also on repositories and platforms recognized by Fiocruz.

D2.8 - Preserve data, for a period defined together with the researchers, in a way that they remain accessible and reusable for future research.

D2.9 - Stimulate the adoption of the FAIR Data Principles in data management and evaluate its degree of application.

D2.10 - Adopt principles and protocols internationally recognized and necessary to promote the integration and interoperability of databases that support the development of public interest research, on the frontier of knowledge, drawing on the analysis of a large volume of data from multiple sources.

D2.11 - Adopt guidelines in the internal programs of research development for the formulation and implementation of data management processes, predicting the availability of resources for the feasibility of this action.

PRINCIPLE 3 – REGULATORY FRAMEWORK

Copyright, industrial and/or other rights over data and databases generated in the research process at Fiocruz will be recognized and protected. In principle, the opening should be the as broad as possible, complying with legal, ethical and strategic limits concerning confidential, personal and sensitive data. Restrictions to access should be made clear and public within specific time limits based on specific legislation.

Guidelines

D3.1 - Promote responsible and transparent data management, providing broad access, availability, authenticity, and reliability, aiming at research integrity. However, exceptions should be applied when established by legislation and internal operational regulations to be formulated by Fiocruz, especially those concerning the protection of personal data and intellectual property.

D3.2 - Establish institutional framework to guide the application of regulations and legal instruments that rule the protection and ownership of rights over research data, in consonance with other institutional policies of Fiocruz.

D3.3 - Formulate normative instruments to regulate the authorization for the use of research data and transfer of rights to/from third parties, protecting the rights of the holder of personal data and, in particular, sensitive data, in accordance with the four levels of access and confidentiality described on the data management plan:

a) Public – when the researcher authorizes access, reuse, and distribution of research data and there are no legal, ethical, and strategic restrictions;

b) Restricted – when the research data can be shared under specific conditions that will be established in terms and agreements of sharing;

c) Confidential – when the access is not allowed due to legal, ethical, and strategic issues;

d) Blocked – when the researcher authorizes data access only after an established period of time.

D3.4 - Reinforce the importance of the adoption of parameters established by the Ethics Committees on the use of scientific research data.

D3.5 - Define the creation of an instance or the extension of functions of an existing instance to advise on the management of confidential, personal, and sensitive data.

PRINCIPLE 4 – DEVELOPMENT OF CAPACITIES AND SUSTAINABILITY

Data management for the purpose of research aiming at their access, sharing, and opening requires capacity development of existing professional profiles and the broadening of the human resources framework involved in these activities.

Guidelines

D4.1 - Ensure the sustainability of the necessary actions for the management, sharing and opening of research data.

D4.2 - Provide the necessary organizational, physical and economic stability for the sustainable development of digital platforms that are secure, reliable and adequate for gathering, depositing, processing, integration, access, sharing, analysis, and preservation of data that comply with ethical standards of research and regulatory frameworks.

D4.3 - Maintain investments for the increasing adoption of open access to data and the technical-scientific production of Fiocruz.

D4.4 - Internalize at Fiocruz the competences, attributions, and capacities necessary for the development of management, sharing and opening of data, by means of the expansion of new professional profiles and capacity-building of existing profiles.

D4.5 - Implement forms of assessment that value the performance of researchers and research groups on management, sharing and opening of research data, including the creation of collaborative networks.

D4.6 - Conduct and publicize qualitative and quantitative studies on the impact of sharing and opening of data regarding the collaboration between groups, scientific production and the policy's social relevance.

Ministério da Saúde

FIOCRUZ
Fundação Oswaldo Cruz