

The opening and sharing of scientific data- The Data Policy of the U.S. National Institutes of Health

The opening and sharing of scientific data- The Data Policy of the U.S. National Institutes of Health

Li-Ting Tsai

Scientific research improves the well-being of all mankind, the data sharing on medical and health promote the overall amount of energy in research field. For promoting the access of scientific data and research findings which was supported by the government, the U.S. government affirmed in principle that the development of science was related to the retention and accesses of data. The disclosure of information should comply with legal restrictions, and the limitation by time as well. For government-sponsored research, the data produced was based on the principle of free access, and government policies should also consider the actual situation of international cooperation^[1] Furthermore, the access of scientific research data would help to promote scientific development, therefore while formulating a sharing policy, the government should also consider the situation of international cooperation, and discuss the strategy of data disclosure based on the principle of free access.

In order to increase the effectiveness of scientific data, the U.S. National Institutes of Health (NIH) set up the Office of Science Policy (OSP) to formulate a policy which included a wide range of issues, such as biosafety (biosecurity), genetic testing, genomic data sharing, human subjects protections, the organization and management of the NIH, and the outputs and value of NIH-funded research. Through extensive analysis and reports, proposed emerging policy recommendations.^[2] At the level of scientific data sharing, NIH focused on "genes and health" and "scientific data management". The progress of biomedical research depended on the access of scientific data; sharing scientific data was helpful to verify research results. Researchers integrated data to strengthen analysis, promoted the reuse of difficult-generated data, and accelerated research progress.^[3] NIH promoted the use of scientific data through data management to verify and share research results.

For assisting data sharing, NIH had issued a data management and sharing policy (DMS Policy), which aimed to promote the sharing of scientific data funded or conducted by NIH.^[4] DMS Policy defines "scientific data." as "The recorded factual material commonly accepted in the scientific community as of sufficient quality to validate and replicate research findings, regardless of whether the data are used to support scholarly publications. Scientific data do not include laboratory notebooks, preliminary analyses, completed case report forms, drafts of scientific papers, plans for future research, peer reviews, communications with colleagues, or physical objects, such as laboratory specimens."^[5] In other words, for determining scientific data, it is not only based on whether the data can support academic publications, but also based on whether the scientific data is a record of facts and whether the research results can be repeatedly verified.

In addition, NIH, NIH research institutes, centers, and offices have had expected sharing of data, such as: scientific data sharing, related standards, database selection, time limitation, applicable and presented in the plan; if not applicable, the researcher should propose the data sharing and management methods in the plan. NIH also recommended that the management and sharing of data should implement the FAIR (Findable, Accessible, Interoperable and Reusable) principles. The types of data to be shared should first in general descriptions and estimates, the second was to list meta-data and other documents that would help to explain scientific data. NIH encouraged the sharing of

scientific data as soon as possible, no later than the publication or implementation period.[6] It was said that even each research project was not suitable for the existing sharing strategy, when planning a proposal, the research team should still develop a suitable method for sharing and management, and follow the FAIR principles.

The scientific research data which was provided by the research team would be stored in a database which was designated by the policy or funder. NIH proposed a list of recommended databases lists[7], and described the characteristics of ideal storage databases as “have unique and persistent identifiers, a long-term and sustainable data management plan, set up metadata, organizing data and quality assurance, free and easy access, broad and measured reuse, clear use guidance, security and integrity, confidentiality, common format, provenance and data retention policy”[8]. That is to say, the design of the database should be easy to search scientific data, and should maintain the security, integrity and confidentiality and so on of the data while accessing them.

In the practical application of NIH shared data, in order to share genetic research data, NIH proposed a Genomic Data Sharing (GDS) Policy in 2014, including NIH funding guidelines and contracts; NIH's GDS policy applied to all NIHs Funded research, the generated large-scale human or non-human genetic data would be used in subsequent research. [9] This can effectively promote genetic research forward.

The GDS policy obliged researchers to provide genomic data; researchers who access genomic data should also abide by the terms that they used the Controlled-Access Data for research.[10] After NIH approved, researchers could use the NIH Controlled-Access Data for secondary research.[11] Reviewed by NIH Data Access Committee, while researchers accessed data must follow the terms which was using Controlled-Access Data for research reason.[12] The Genomic Summary Results (GSR) was belong to NIH policy,[13] and according to the purpose of GDS policy, GSR was defined as summary statistics which was provided by researchers, and non-sensitive data was included to the database that was designated by NIH.[14] Namely. NIH used the application and approval of control access data to strike a balance between the data of limitation access and scientific development.

For responding the COVID-19 and accelerating the development of treatments and vaccines, NIH's data sharing and management policy alleviated the global scientific community's need for opening and sharing scientific data. This policy established data sharing as a basic component in the research process.[15] In conclusion, internalizing data sharing in the research process will help to update the research process globally and face the scientific challenges of all mankind together.

[1]NATIONAL SCIENCE AND TECHNOLOGY COUNCIL, COMMITTEE ON SCIENCE, SUBCOMMITTEE ON INTERNATIONAL ISSUES, INTERAGENCY WORKING GROUP ON OPEN DATA SHARING POLICY, *Principles For Promoting Access To Federal Government-Supported Scientific Data And Research Findings Through International Scientific Cooperation* (2016), 1, organized from Principles, at 5-8, https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/NSTC/iwgodsp_principles_0.pdf (last visited December 14, 2020).

[2>About Us, Welcome to NIH Office of Science Policy, NIH National Institutes of Health Office of Science Policy, <https://osp.od.nih.gov/about-us/> (last visited December 7, 2020).

[3]NIH Data Management and Sharing Activities Related to Public Access and Open Science, NIH National Institutes of Health Office of Science Policy, <https://osp.od.nih.gov/scientific-sharing/nih-data-management-and-sharing-activities-related-to-public-access-and-open-science/> (last visited December 10, 2020).

[4]Final NIH Policy for Data Management and Sharing, NIH National Institutes of Health Office of Extramural Research, Office of The Director, National Institutes of Health (OD), <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html> (last visited December 11, 2020).

[5]Final NIH Policy for Data Management and Sharing, NIH National Institutes of Health Office of Extramural Research, Office of The Director, National Institutes of Health (OD), <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html> (last visited December 12, 2020).

[6]Supplemental Information to the NIH Policy for Data Management and Sharing: Elements of an NIH Data Management and Sharing Plan, Office of The Director, National Institutes of Health (OD), <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-014.html> (last visited December 13, 2020).

[7]The list of databases in details please see : Open Domain-Specific Data Sharing Repositories, NIH National Library of Medicine, https://www.nlm.nih.gov/NIHbmic/domain_specific_repositories.html (last visited December 24, 2020).

[8]Supplemental Information to the NIH Policy for Data Management and Sharing: Selecting a Repository for Data Resulting from NIH-Supported Research, Office of The Director, National Institutes of Health (OD), <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-016.html> (last visited December 13, 2020).

[9]NIH Genomic Data Sharing, National Institutes of Health Office of Science Policy, <https://osp.od.nih.gov/scientific-sharing/genomic-data-sharing/> (last visited December 15, 2020).

[10]NIH Genomic Data Sharing Policy, National Institutes of Health (NIH), <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-14-124.html> (last visited December 17, 2020).

[11]NIH Genomic Data Sharing Policy, National Institutes of Health (NIH), <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-14-124.html> (last visited December 17, 2020).

[12]*id.*

[13]NIH National Institutes of Health Turning Discovery into Health, *Responsible Use of Human Genomic Data An Informational Resource*, 1, at 6, https://osp.od.nih.gov/wp-content/uploads/Responsible_Use_of_Human_Genomic_Data_Informational_Resource.pdf (last visited December 17, 2020).

[14]Update to NIH Management of Genomic Summary Results Access, National Institutes of Health (NIH), <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-19-023.html> (last visited December 17, 2020).

[15]Francis S. Collins, Statement on Final NIH Policy for Data Management and Sharing, National Institutes of Health Turning Discovery Into

Links

- [About Us](#)
- [NIH Data Management and Sharing Activities Related to Public Access and Open Science](#)
- [Final NIH Policy for Data Management and Sharing](#)
- [Supplemental Information to the NIH Policy for Data Management and Sharing: Elements of an NIH Data Management and Sharing Plan](#)
- [The list of databases in details please see : Open Domain-Specific Data Sharing Repositories](#)
- [Supplemental Information to the NIH Policy for Data Management and Sharing: Selecting a Repository for Data Resulting from NIH-Supported Research](#)
- [NIH Genomic Data Sharing](#)
- [NIH Genomic Data Sharing Policy](#)
- [Update to NIH Management of Genomic Summary Results Access](#)
- [Francis S. Collins, Statement on Final NIH Policy for Data Management and Sharing](#)

Download

- [Principles For Promoting Access To Federal Government-Supported Scientific Data And Research Findings Through International Scientific Cooperation \[pdf \]](#)
- [Responsible Use of Human Genomic Data An Informational Resource \[pdf \]](#)

Tsai, Li-Ting
Legal Researcher

Release : 2021/01