

Position Statement on Open Data by the Young Academies of Europe and the Global Young Academy

We define Open Data as all data that is freely available to external parties to use, link and redistribute as they wish, without restrictions from mechanisms of control and regardless of the data scope (academic, non-academic). Open Data needs to be supported and widely implemented in all fields of inquiry. The effective adoption of Open Data requires substantial investments, specialized infrastructure, field-specific evaluation, and regulatory efforts. This is to a degree more urgent than in the case of Open Access, where the infrastructure and associated business models are often more mature. Rather than implementing Open Data as a blanket policy, we recommend a **staged adoption**, where benefits and investments required to achieve those benefits are carefully evaluated and balanced against each other at every step. In particular, the following points need to be taken into consideration:

1. NOT ALL DATA NEED TO BE OPEN

Careful consideration needs to be given to what data are disseminated, why, how and when (with which priority and timeline). Not all data can be shared, and there is not always an added value for data that can be shared. Consider for instance the notes of a researcher doing qualitative analysis, personal data used in medicine and data generated by easily reproducible computer simulations.

2. ONE SIZE DOES NOT FIT ALL

We recommend the articulation of best practices on a field-by-field basis to ensure that all critical aspects of Open Data are taken into account prior to making Open Data mandatory. At the same time, dialogue across disciplines should be encouraged to compare strategies and share experiences. Depending on their ethos, history and publication culture, different research areas currently attach varying degrees of importance and recognition to data sharing. Additionally, clear guidelines are needed for the protection of sensitive personal, security-related, or commercially relevant data. Applicable licensing models, intellectual property concerns and their implications for relations with sponsors need to be clearly articulated as well. The type of data that should be shared, in what formats, with what meta-data should be clarified, and legal aspects regarding data privacy and safety issues in relation to animal activism should be resolved at a regional level (e.g. European level).

3. LONG-TERM SUSTAINABILITY IS CRUCIAL

It is critical to have suitable infrastructure in place to protect the data in the years to come. Data can have a longer lifespan than specific interpretations, and can be used in ways that differ from those intended when they are first generated. Any initiative that supports Open Data needs to consider modalities, technology standards, incentives and infrastructures for sustainable data sharing, publishing and re-use. Without those in place, making data open is not productive.

4. OPEN DATA REQUIRES OPEN SCIENCE

For Open Data to function, researchers need access to related research components, such as software, code, and models.

5. KEEP BUREAUCRACY DOWN

Open Data requirements should not become a burden for already overstretched researchers. Bureaucracy surrounding Open Data should be kept to an absolute minimum, and the logistics involved in making data publicly available should not be the sole responsibility of researchers. The provision, financing, maintenance of secure, user-friendly, easily citable and sustainable facilities for data curation and long-term storage should be the responsibility of e.g., funding agencies and/or (inter)national data protection authorities. We recommend leveraging economies of scale by establishing, for example, Europe-wide open source data hosting solutions that are available to researchers globally at minimal costs.

6. CREDIT WHERE IT IS DUE: GENERATING INCENTIVES

Adequate mechanisms are needed to ensure that researchers at all stages of their careers get due credit for their work and efforts in producing, sharing and re-using relevant research data. This would encourage authors not only to share their data, but also to document and prepare these in a way that is accessible, and that maximizes the utility of these data for other researchers. Best practices need to be urgently established on a field-by-field basis that assure that credit will be properly allocated to the effort of individual researchers and researcher teams.

7. TRAIN AND SUPPORT RESEARCHERS IN DATA MANAGEMENT AND RE-USE

There is currently a large gap between data creation and use, and data management. More resources are needed to train researchers in how to share and document their data most effectively, and to support their efforts to share and re-analyze existing data with appropriate experts (data scientists and archivists). This will help overcome both practical and psychological impediments of data-sharing, and secure the quality of the disseminated data.

Ratified by Global Young Academy, Young Academy of Europe, Young Academy of Norway, De Jonge Akademie (NL), Die Junge Akademie (Germany), Young Academy of Scotland, Young Academy of Sweden, Polish Young Academy, The Israel Young Academy, Young Academy of Belgium.

Compiled by Sabina Leonelli (GYA), Lennart Martens (BYA) and Steven Latre (BYA) based on input provided by

Katrien De Moor, on behalf of the Young Academy of Norway

Caroline Lynn Kamerlin and Monica Brinzei, on behalf of the Young Academy of Europe

Steven Latre, on behalf of the Young Academy of Belgium

Sabina Leonelli and Moritz Riede, on behalf of the working group Open Science of the Global Young Academy

Jakob Macke, on behalf of working group Big Data of Die Junge Akademie

Martijn Wieling, on behalf of De Jonge Akademie

Sara Strandberg, on behalf of the Young Academy of Sweden

Contact: <http://www.eya.eu/> (European Young Academies), <http://globalyoungacademy.net/> (Global Young Academy)



Akademiet for yngre forskere
The Young Academy of Norway

