

Priority Initiative "Digital Information" of the Alliance of German Science Organisations
Position paper "Research data at your fingertips"

By the Working Group research data [February 2015]

See <http://www.allianzinitiative.de/handlungsfelder/forschungsdaten.html>

The following comments and recommendations focus on the topic "research data", from the working group research data of the Priority Initiative "Digital information" of the Alliance of German Science Organisation. Writers and authors: Michael Franke, Stefan Heinzl, Reiner Wall, Dr. Janna Neumann, Dr. Heike Neuroth, Dr. Hans Pfeiffenberger, Henriette Senst, Andrea Victory mountain, Dagmar Sietek, Claus Spiecker, Prof. Dr. Joachim Wambsganss, Dr. Danny Weber, Dr. Stefan Winkler-Nees

I. Vision 2025 "Research data at your fingertips"

Scientists of all disciplines can all research data easy to access quickly and easily, to do research at the highest level and achieve excellent results. You can work with others and keep their research results For-safe. Research data are employed in a form available that both research across disciplinary and across national boundaries-made light and relieved.

The publication of research data and software increases the scientific reputation. Scientists are supported in the search for, collect, capture and manage their data.

Easily usable digital infrastructures and scientific and technical Information specialists support the full research cycle.

II. Status Quo: Preliminary Results

The topic of research data is characterized by a great dynamism and diversity. In a description of the status quo, the following main aspects have to be considered who-to:

1. Specialised and structural observation

All stakeholders are faced with a wide and complex range of tasks that is still unclear and unspecified large extent. Specific responsibilities, such as the long-term preservation of research data, are often not known or even unknown or unassigned. Of particular importance are on the one hand, the distinction between disciplinary and interdisciplinary areas of responsibility as well as the question of whether these duties should be aggregated at national or international level. First successful approaches include interdisciplinary Data Infrastructures discipline-specific and international research Data Infrastructure.(For example, EUDAT, EMBL EBI or ELIXIR; ESFRI ERICs such as CESSDA, CLARIN, DARIAH, SHARE, etc.)

Current organizational structures in Germany are strongly influenced by federal units and partly heterogeneous landscape of science organizations operating in part of a (scientific) competitive situation. This leads to overlapping or not compatible with each other developments.

Nevertheless, important progress has been made in many areas, such as the establishment of offers for the award of persistent identifiers that allow a permanent homing ability and citation to research data and in the organization of university counseling services to assist in research data management. These are worth-full blocks in the further development of research data infrastructures.

2. Awareness of the professional management of research data

Some disciplines, such as the Earth and Environmental Sciences or Social Sciences, have already established a strong awareness of the use of research data and its operation using appropriate tools for professional data management. In most cases, a specialized culture of

sustainable use of data in the sense of establishing good scientific practice has developed. At the same time through the reuse of valuable research data, new results have been achieved in the economic knowledge process. The increasing availability of data allows, in certain areas, the re-use of socially relevant data outside of science. An example in epidemiology: better knowledge about the spread of diseases, their potential impact, and appropriate preventive measures.

Of increasing importance to scientific research are data generated in non-scientific contexts. This includes data arising by law or in the context of administrative processes (Public Sector Information), data from social networks, in the health sector or in citizen science initiatives. In well established institutions, publicly funded (libraries, archives, museums, statistical offices), there is a different out-of-embossed awareness in regard to dealing with research data, but also a particular nascent willingness to deal with it. Many commercial companies have also recognized the economic value of data.

Scientific organizations

Many scientific organizations address the use of research data, but put it in a heterogeneous and still largely uncoordinated way around. Allianz AG research data has the "principles for dealing with research data" therefore developed, which adopts the leadership of all alliance organizations.

However, especially the financing of adequate infrastructure for research data is still not clear. The funding organizations have recognized this need and developed a range of support measures. Common to these measures is that they **only work on temporary projects** and thus do not allow the necessary sustainable operation of infrastructure. At the same time the various promotional offers are not coordinated between the conveyor Organizations.

As part of the research funding it must be recognized by the conveyor that Data management is an indispensable part of research projects and financial sub-support needs.

An important prerequisite for handling research data are corollary work-rule (policies) and formal rules, such as those currently being tested in a pilot project for open research data in Horizon 2020, like the requirement for the use of Data Management Planning, and sharing of research data after project completion.

To successfully implement a professional approach to research data it needs political support. Here are the first positive steps have already been observed, such as the onset of federal / state polls show (GWK6) or the activities of international-Ebene7. From an institutionalization of research data infrastructures, we are still far away. Thus, for example, legal framework in many areas not yet fully elucidated.

Indispensable for scientific data management are also information specialists and service providers. In current practice, these tasks are performed predominantly by career changers; the systematical development of new vocations is lacking.

4. Act in the international arena

Research data are not confined to national borders. Many scientists collaborate internationally and gain new insights precisely and gain reputation. This must be supported by an international dialogue of enabled infrastructures. First positive signs are already visible but face many challenges.

III. Recommendations

The implementation of the following recommendations will help the scientist in Germany in the near future to have an adequate research data infrastructure.

An essential prerequisite for this is the scientific recognition of the publication of research data for future use (citable and accessible) and their loading into account in the evaluation process of individuals and institutions.

1. Nationwide discipline-specific coordination bodies are necessary to establish the respective requirements, identify the necessary infrastructure components that drive the development and expansion as well as putting relevant policies and standards. Ideally wise are represented on these committees science, politics, research funding and infrastructures. As an example of exemplary character, the "Council for Social and host-community data (RatSWD)" can be seen.

2. Similar to libraries and their information literacy, the supply for the sustainable provision of scientific research data needs the **establishment of trust (trustworthy) legally binding organizational structures**. Ideal would be the establishment and development of specialized scientific Research Data Centers.

3. Durable and stable funding models for long-term covering development and investment costs and the costs of the operations are continually to develop and implement. A key component is an adequate base funding of infrastructure.

4. Research requires clear and unambiguous legal framework for the use and publication of research data. Must be regulated, among other legally compliant scientific use of personal data copyright law protected and licensed materials.

5. To enable global and networked research, standardization in the technical area as well as in semantic and syntactic questions are necessary. International coordination, coordination and implementation are in the foreground, such as the successful examples DFN / GEANT11 (network infrastructure), INSPIRE12 (geo-localized data) and DDI13 (Social Sciences) show.

6. The digital transformation process requires additional training of scientific experts. New professions and career paths have to be established by the establishment of additional courses for information specialists.