

Are there a million ways to describe the world?

Or how to share your authority without losing control

Speaker(s) / Contact

confirmed: Desiree Mayer (NFDI4Culture) / Thorsten Trippel (Text+) / Stefan Buddenbohm (Text+) / Regine Stein (Text+) / Jürgen Kett (DNB) / Barbara Fischer (DNB)

Contact: Barbara Fischer; email: b.k.fischer@dnb.de; Tel: 0341 2271 383

Abstract

What does it mean to share a national infrastructure for research data? In order to be able to share data, it needs to be fair as in the [FAIR DATA Principles](#). Implicitly we need data that is standardised. Standards refer both to the data format meaning technical specification and from the semantic viewpoint which categories we use to sort, structure, and group the data describing our topic. The technical and semantical perspectives are like two sides of a coin. The workshop will focus on the semantic perspective. How do we specify the categories? Are they naming different entities to distinguish groups of objects from others? Are the entity groups described in detail by further properties or attributes? Do they form an ontology? Do they follow rules or frameworks? If so, are these formed by a coordinated body mandated by a specific community? If Yes is the answer to these questions, this describes pretty much the objective of controlled vocabulary and authority files.

On the meta-level, most scientists will agree. Scientists will describe their world or topic in structured categories. These categories are shared by a smaller or bigger community. They are more or less regulated. Classification is the second step in any scientific workflow after the first step of description. The commonly shared entity groups lay the base to construct a network of reference knots. Knots that link diverse information from different sources and allow the flow of knowledge to switch its direction and cross boundaries. Biologists apply the taxonomy for species, Chemists use their periodical system and librarians collaborate by sharing their standardised catalogues in order to facilitate the transfer of books. Within the catalogues, there are several repetitive elements or core entities. Persons appear as authors and subjects of biographies. City names appear repeatedly both as locations for editing houses and as subjects for travelogues. Subject headings group a long list of publications. And there are works, corporate bodies, and conferences too - all together they form for instance in the German-speaking countries the Integrated Authority File (in German: Die Gemeinsame Normdatei, GND). A handy tool used by librarians to precisely define search entries and make their work more efficient by providing controlled vocabulary for items that can be reused infinitely.

First, the workshop will give an introduction to the GND. Addressing both its character as a data hub to a controlled vocabulary and as a governance structure granting the reliability of its data quality. The two following lightning talks will give a short overview of how two NFDI consortia, NFDI4C and Text+, query the needs and requests of their specific communities on the GND, transfer these in workflows and plan to become active contributors to the GND. This means, to find an answer to the question: How to share your authority without losing control? The last part of the workshop invites all participants to share the approaches to the issue in their domains and countries, to identify common patterns and opportunities for further collaboration.