When Data Crosses Borders – Join Forces!
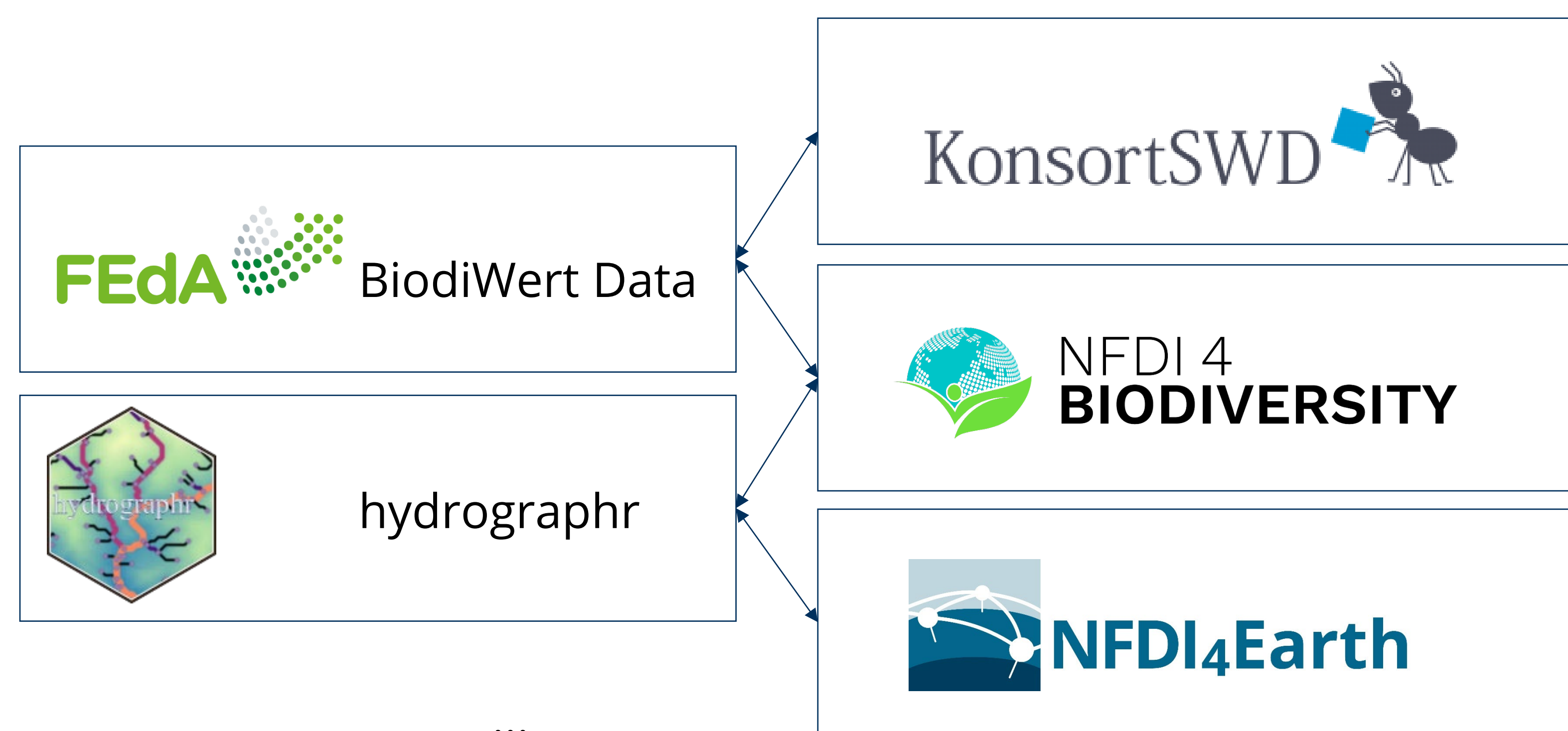## Multidisciplinary Use Cases Within NFDI

Barbara Ebert, Sami Domisch, Christin Henzen, Jimena Linares, Kati Mozygemba, Bernhard Miller, Bernhard Seeger, Jörg Seegert

(Authors in alphabetical order)

Relevance

Multidisciplinary use-cases for integrating research data involve data, software, and methods across research disciplines to address research questions that cannot be efficiently addressed by a single discipline or method alone. These use-cases require collaboration across multiple research disciplines and likewise the integration of heterogeneous data, e.g., experimental, observational, and computational data, and management as well as processing workflows. They also require communication about a common understanding of concepts and methods (e.g., the understanding of "interview" can differ between disciplines and methodological approaches). Multidisciplinary use-cases thus help develop a more comprehensive and nuanced understanding of complex phenomena, by leveraging the complementary strengths across research disciplines and analytic approaches including software and workflows, as well as data types.

Topic and Data-driven Perspectives

There are **two perspectives to identifying multidisciplinary use-cases**: one based on research topics, the other data-driven.

The **topic-based approach** is driven by the desire to tackle **complex questions with scientific rigor**. This underscores the need to link data from **different methodological approaches** and disciplines to solve complex research questions. Examples include estimating the socio-economic impacts of climate change and biodiversity loss or assessment of conservation efforts. Climate change's overarching relevance as well as its complexity have brought, among others, researchers from the Earth System Sciences (ESS), the life sciences and the social sciences to **join forces and approach the same questions from different angles**. To do so, it is necessary to create ways for the various types of data to be combined and used for different forms of analysis.
Likewise, disaster prevention and risk management and esp. its aspects of vulnerability, preparedness and resilience also require a multidisciplinary approach, e.g., by linking georeferenced survey data with existing crisis-relevant contextual information of various disciplines.

The **data-driven approach** is not motivated by known substantive research questions but seeks to enable the **identification of relevant patterns** through enhancing the data. The approach will thus emphasize the machine-actionability of data from different sources and this facilitates their interoperability. Researchers – and artificial intelligence – will then, for example, be able to look for patterns – such as health effects (measured e.g., through treatment data) in brightly lit areas (measured using satellite data).
To be sure, the first approach also benefits from **richly annotated, machine-actionable metadata**.

Benefits and First Identified Examples

Data drives a better understanding of complex issues. Yet multidisciplinary use- cases generate particular benefits:

1. data can be **contextualized in the knowledge** of each participating field, which
2. given the likely relatively unconnected nature of research in the area offers **new hypotheses** that
3. can be tested by **methods from each** of the participating disciplines and thus
4. will **increase knowledge** in likely more than one field.

Multidisciplinary use cases promise non-linear benefits as compared to single-discipline use-cases. The examples below illustrate that, in order to reap these benefits, some effort is required

The hydrographr R-package
Freshwater ecosystems are relevant to both Earth Systems Sciences and especially to Hydro Sciences and - as a habitat - to biodiversity research. The hydrographr R-package was developed within a use case in NFDI4Biodiversity - on the basis of an NFDI4Earth pilot - that aims to lower the burden for potential users by offering easy-to-use functions within R. The package is tailored towards high-resolution stream network data and capitalizes on fast and RAM-efficient open-source command-line GIS software. Users can hence create their workflows in R on their local computers employing high-resolution network data across large spatial extents, since the actual data processing is taking place outside R.

FEdA BiodiWert Data
FEdA funds interdisciplinary projects on the impact of nature conservation efforts. A collaboration on RDM with NFDI4Biodiversity began with FEdA's "BiodiWert" projects. These are mostly multi-method research.
A support workflow involved experts and curators from GFBio's Data Centers and KonsortSWD partners that assisted in the creation of DMPs and preparation of specific data types for archiving. A data policy, was developed together with the FEdA coordination office as a guideline for archiving and sustainable publication of research data. This collaboration is a pilot for consulting on multi-disciplinary data management under NFDI's umbrella.

The Way Ahead

Overall, the value of multidisciplinary use-cases for data sharing lies in their ability to facilitate collaboration between different scientific disciplines and methodological approaches. By leveraging diverse data sources, researchers can gain new insights and make more accurate predictions about complex phenomena. Additionally, by promoting collaboration across different consortia, the NFDI can help break down barriers between research fields and promote a more holistic approach to scientific inquiry.

In a first summary, we identified the following important questions **for successfully detecting and processing multidisciplinary use cases across research data infrastructures**.

1. How can we find meaningful multidisciplinary use-cases?
2. How can the needs of researchers, the various discipline-specific issues, and scientific methods be taken into account and integrated?
3. What are the ways and processes to systematically explore and exploit them?
4. How can we encourage researchers working on multidisciplinary use-cases in the future?
5. What are the specific requirements in multidisciplinary use cases? Which obstacles have to be overcome for a good scientific service?
6. What challenges and opportunities arise from the consideration of discipline-specific or multidisciplinary use cases for the development of (national) research data infrastructures?
7. How can multidisciplinary use cases be stimulated or initiated?
8. How do we consider legal and ethical problems that occur by combining different data types?

Are you interested in future activities?

Come talk to us! Bring your ideas!
Joerg.Seegert@tu-dresden.de or
Christin.Henzen@tu-dresden.de