BioDATEN – Provider’s perspective in the NFDI

3rd NFDI Community Workshop in Munich (10/02/20)
Services for Research Data Management in Neuroscience

Dirk von Suchodoletz, Speaker DataPLANT consortium
Universities of Freiburg, Kaiserslautern, Tübingen, ...
FZ Jülich, ...
Outline

• Background: The BioDATEN science data center and the DataPLANT consortium
• Computer centers as service providers
• The changing role of the (university) computer center as a service provider
• Structural and legal challenges of cooperation
• ... in the context of data repository operators
• Actually offered (provider) resources in DataPLANT (in-kind contribution)
Background: About BioDATEN

• Science Data Center focusing on Bioinformatics in the state of Baden Württemberg
  - Standardization in the Bioinformatics community
  - Qualification and education
  - Research and development
  - Services and infrastructure

• Four year project started in July 2019, funded by the Ministry of Science, Research and the Arts to support the BW NFDI activities
**About DataPLANT**

- Building on existing, working infrastructure
- Preserving research context
- Comparably smaller, but focused community
- Configuration: 4 applicant institutions, ~30 participants
- 11.5 Mio € budget total planned

Wide representation of the fundamental plant research community and of infrastructure+services

⇒ strength; rooted in fundamental plant hotspots

⇒ **complementary expertise** of scientific community, developers and service providers
Contradicting requirements

- Often impossible to define the exactly “right” resource (sizing challenge)
- Clouds and aggregated HPC resources
  - IT industry and science driven
- Increased efforts for market analysis, system selection, proper procurement and operation of (large scale) IT infrastructures
- Contradiction of typical project run times vs. delay for equipment provisioning and write down time spans of that equipment
Ongoing cultural change

• More and more cloud-like operation of workflows
  - Data and compute moving into centralized resources
• Shared infrastructures allow for a fast start of projects (as necessary overhead capacity is much easier to maintain and justify in centralized resources)
  - Shared overhead much less than individual ones
• Grant providers start to understand the changed technological landscape and to adopt their requirements
Cooperation

• Significantly widened scope of IT services

• Cooperation a chance to keep up with the demand and offer a relevantly wide service, e.g.
  – bwLehrpool
  – bwHPC + bwDATA, bwCloud
  – de.NBI cloud

• Allows for specialization and community focus

• Unsolved issues still remain
Motivation for providers

• Massive changes in the IT landscape and user expectations require re-orientation of university (scientific) computer centers

• Define future strategies and operation models like find a new meaning in research data management

• Recalibrate the role of (de-)central IT service provisioning (for basic and high level services)

• Provide efficient infrastructures and consultation to the various scientific communities

• Participate in infrastructure calls
Motivation for users

• Offloading of non-domain specific tasks and services (organization, procurement, operation)
• Focus on domain specific tasks
Governance

• Ensure a persistent relevance of future computer centers through user participation and feedback loops
  - Close cooperation and consultation (like we do in NEMO and for bwSFS) helps all stakeholders to have suitable, up to date infrastructures tailored to the users’ needs

• Such structures in its infancy, but required for NFDI (grant application)

• NFDI wide coordination should advance this topic

• Core for a successful NFDI operation
Re-Financing

• Often grant driven (and inherently not sustainable)
• Changed flow of money from simple project driven and organization centered practice to fragmented streams to different providers
• Large infrastructure initiatives (de.NBI, NFDI, ...) need
  – To solve role of personnel (permanent vs. project based)
  – Define suitable business and operation models
  – Untangle the federal / state requirements of cash flows in mixed consortia (honor the VAT requirements)
Infrastructure4NFDI

• Not part of the ongoing NFDI call yet, but nevertheless required

• Many consortia bring in established infrastructures (which are often quite un-coordinated)

• Cross-cutting topic and will be a task of the NFDI directorate to be defined / solved

• de.NBI cloud might get established as the federal cloud infrastructure for the NFDI (many consortia asked for letters of support)
Resources offered (in DataPLANT)

• Commitment for
  - 3.156+ cores
  - 9.900+ Tbyte

• Enables hosting/execution of
  - The DataPLANT Hub
  - Workflows
  - Data
  - Metadata
  - and much more ...
Thank you for your attention!