## Meeting minutes

Project Background ChemBioSys

- second funding period (4 years, 2018–2022)
- not every working group participates in ChemBioSys
- additional input from website
- What data?
  - Many different data types (working groups on algae, bacteria, etc.), metabolism data from the metabolomics working group
  - data from GC-MS and LC-MS is basically similiar, even though some people might contradit that statement
  - NMR data is diverse
  - Sequencing data can come from two methods: – Genomics (DNA sequencing) – Transcriptomics (more complex)
  - data structure: mostly strings consisting of (A,T,C,G)
  - some image data (images of petri dished, plants, etc.), but these are the minority

Status of data management

- some labs use online repositories for their data management
- up until now no existing data management plan (did i get that right?)
- started in 2018, the data management system shall be rebuild
- infrastructure: one main server, two backup servers in different power networks
- aim of the platform [Seeks]: show the advantages of a data management system, upscaling is planned
- maybe migration to [i didn't get the name of the platform]

Challenges & Trends

- all ChemBioSys data stored in different places: some in local repositories, some in file structure on Uni server, some in data platforms of other institutions
- only 4h resources per week for data management
- no data management mindset: people don't see the benefit in proper data management
  → they rather keep files
  → lot of change management / trainings required
- space limitation of 7TB per project
- data access → all scientists keep their data private
  - master-account at ChemBioSys
  - create account at fair-dom (read only), the data will be the same (did i get that right?)
- new trend: journals for data publishing, for example Journal of Physical and Chemical Reference Data
- what is not considered a challenge yet: consistency/format of datasets and meta data; data governance

How datasets & meta data are stored

- ISA-Principle
  - Investigate
  - Study
  - Assay
- no data format restriction by ChemBioSys
- metadata is not standardized (yet)
- ISAcreator
- up until now metadata is mostly text based (.csv, .txt, .docx) 1
- the usage of ISAcreator is not part of the data management, but the working groups are responsible to implement it in their workflow (did i get that right?)
- eventually, ISAcreator-files will be converted to JSON