

Umgang mit Forschungsdaten in Verbundprojekten

5. *LUH-interner Erfahrungsaustausch*

Zeit: 29. Oktober 2020, 10:00-12:00 Uhr

Videokonferenz

- Protokoll -

Begrüßung, Vorstellung, Zielsetzung

Das Serviceteam Forschungsdaten begrüßt die Anwesenden. Anschließend stellen die Teilnehmenden sich und ihre Arbeitsbereiche vor. Das Serviceteam Forschungsdaten erläutert die Zielsetzung und den geplanten Ablauf. Die Veranstaltung soll Gelegenheit geben:

- FDM-Maßnahmen, die sich in bestimmten Projekten bewährt haben, bekannt zu machen
- auf typische Probleme hinzuweisen und vor Ansätzen zu warnen, die sich in der Praxis nicht bewährt haben
- Lücken im derzeitigen Unterstützungsangebot der LUH zu identifizieren, damit diese geschlossen werden können

Data Management in the BRIWECS Project (Tsu Wei Chen)

(zu Einzelheiten siehe Präsentation)

In dem Konsortium BRIWECS werden an sieben über Nord- und Mitteldeutschland verteilten Standorten neue Zuchtsorten von Winterweizen erforscht. Für standortübergreifende Auswertungen und Vergleiche ist es wichtig, dass Daten nach einheitlich Standards erhoben, gespeichert und dokumentiert werden. Ein für das ganze Konsortium gültiges Datenmanagementkonzept wurde jedoch erst im Laufe des Projektes erarbeitet und wird noch nicht an allen Standorten gleichermaßen konsequent umgesetzt. Das macht aufwendige manuelle oder teilautomatische Nachbearbeitungen und Qualitätsprüfungen notwendig.

Die Ablage der Daten erfolgt nun in dem Seafile-Cloudspeicher einer [LUIS-Projektablage](#), auf der alle am Konsortium Beteiligten Zugriff haben. Es gibt sowohl allgemeine Ordner für das ganze Konsortium als auch Ordner für jeden Standort. Die darin befindlichen Unterordner sind nach einem einheitlichen Schema strukturiert. Problematisch ist, dass Lese- und Schreibrechte nicht auf Ordnerbene differenziert werden können, so dass schlimmstenfalls versehentlich Daten im „falschen“ Ordner gelöscht oder geändert werden könnten.

Nachtrag des Service Teams Forschungsdaten zur Seafile-Nutzung: Im persönlichen [Cloud-Seafile](#) ist es möglich, ordnerweise Lese- und Schreibrechte zu differenzieren. Warum diese Funktion im [Projekt-Seafile](#) deaktiviert ist, können nur die LUIS beantworten. Ggf. kann der dortige [support](#) dieses Feature auf Anfrage für eine bestimmte Projektablage wieder freischalten.

Forschungsdatenmanagement im Exzellenzcluster PhoenixD (Reinhard Caspary)

(zu Einzelheiten siehe Präsentation)

Das interdisziplinäre Cluster zu optischen Technologien umfasst Teilprojekte aus der Chemie, der Physik und den Ingenieurwissenschaften. Die Forschungsmethoden - vor allem Laborversuche und Computersimulationen - und die anfallenden Forschungsdaten sind entsprechend heterogen. Die beteiligten Partner nutzten bisher individuelle Lösungen für das Datenmanagement. Ein Datenaustausch innerhalb des Clusters findet vor allem informell über persönliche Kontakte einzelner Beteiligter statt. Auch die Publikation von Daten wird unterschiedlich gehandhabt.

Für die Zukunft wird angestrebt, einheitliche Mindeststandards für Datenformate und Metadaten zu etablieren, die dann um fach- oder teilprojektspezifische Angaben ergänzt werden können. Außerdem sollen Daten so abgelegt werden, dass andere Cluster-Mitglieder darauf zugreifen können. Insbesondere soll verhindert werden, dass Daten oder die dazugehörigen Kontextinformationen verlorengehen, wenn Beteiligte ausscheiden. Das Cluster vernetzt sich mit verschiedenen NFDI-Konsortien und möchte sich an den dort erarbeiteten Standards und Best Practice-Beispielen orientieren.

Fragen / allgemeine Diskussion

- Bei selbstgeschriebenem Code besteht die Gefahr, dass er mit neueren Versionen der eingebundenen Bibliotheken nicht mehr korrekt funktioniert. Lösung: Mittels Git werden nicht nur Änderungen am Code mitgeloggt, sondern auch dokumentiert, welche Bibliotheken in welcher Version für die Ausführung benötigt werden.
- Metadatenerfassung: im Bereich Simulation werden viele Metadaten automatisiert erhoben und im Fileheader der Ergebnis-Dateien gespeichert. Mit einem Skript können diese Dateien dann durchsucht werden.
- Fachspezifische Vokabulare: In interdisziplinären Projekten kann es vorkommen, dass bestimmte Begriffe in den einzelnen Fächern unterschiedliche Bedeutungen haben. Das kann zu Missverständnissen beim fachübergreifenden Austausch führen. Lösung: Einen für alle Beteiligten zugänglichen Thesaurus oder ein Glossar, also ein Verzeichnis der Begriffe und ihrer (ggf. unterschiedlichen) Definitionen, anlegen.
- Wie und wann sollte in einem Verbundprojekt das einheitliche Datenmanagement verbindlich geregelt werden? → Grundsätzlich: je früher desto besser. Es kann aber sein, dass erst einmal ausreichend Erfahrung gesammelt werden muss, um einschätzen zu können, welche Regelungen sinnvoll und praxistauglich sind. Alle Beteiligten sollten möglichst hinter den Verabredungen stehen, damit diese dann auch zuverlässig eingehalten werden. Ein regelmäßiger Austausch zum Thema Datenmanagement bei Projekttreffen ist empfehlenswert.

Termin- und Themenplanung für Folgeveranstaltungen

Das Thema der kommenden Veranstaltung soll „Zugriff auf, Eigentum an und Verwertung von Forschungsdaten“ sein. Der Termin wird in der zweiten Aprilhälfte 2021 liegen. In der darauffolgenden Veranstaltung Ende 2021 könnte es dann um Externe Fachdienste und Initiativen (NFDI etc.) gehen.