
Datenmanagement in Verbundprojekten

Rollen und Verantwortlichkeiten

12. LUH-interner Erfahrungsaustausch

26. April 2024, 10:00-12:00 Uhr (Videokonferenz)

- Protokoll -

Begrüßung, Vorstellung, Zielsetzung

Das Serviceteam Forschungsdaten begrüßt die Anwesenden und erläutert die Zielsetzung und den geplanten Ablauf. Die Veranstaltungsreihe soll Gelegenheit zum fachübergreifenden Austausch zu verschiedenen Aspekten des Forschungsdatenmanagements geben. Insbesondere sollen typische Herausforderungen thematisiert und bewährte Lösungen bekannt gemacht werden.

Die Organisation des Datenmanagements im Projekt reaCtor (Dr. Michael Steinke)

(Präsentation nicht öffentlich)

Im Projekt wurden Rollen und Gremien definiert, nämlich ein Datenmanager aus dem Kreis der PIs, ein Data Steward als unterstützende studentische Hilfskraft und eine Task Group Datenmanagement, in der alle Partner vertreten sind. Es hat sich jedoch als schwierig herausgestellt, die Rollen konkreten Personen zuzuweisen und ausreichend Feedback zu geplanten FDM-Maßnahmen wie z.B. dem Erstellen einer Projekt-Richtlinie zu erhalten. Obwohl bei allen Projektpartnern grundsätzlich die Einsicht vorhanden ist, dass ein gutes Datenmanagement sinnvoll und notwendig ist, zeigte sich, dass die praktische Umsetzung unterschiedlich gewissenhaft erfolgt und meist eher keine Priorität genießt. Im Nachhinein hat sich gezeigt, dass es hilfreich gewesen wäre, schon bei der Antragstellung Personalressourcen für das Datenmanagement einzuplanen.

Fragen und Anmerkungen

- Wird der Seafile Cloudspeicher des LUIS gebackupt? → Der Cloud-Speicher einer Projektablage wird täglich gesichert, nicht jedoch die persönlichen Seafile-Cloudspeicher
- Erfüllt das LUH-Datenrepositorium die EU-Anforderungen an ein "trusted repository"? Hat es ein entsprechendes Zertifikat? → Institutionelle Repositorien müssen nicht zertifiziert sein, um als "trusted repository" zu gelten. Technisch sollte das LUH-Datenrepositorium die Anforderungen erfüllen. Das LUIS erstellt derzeit eine Übersicht der EU-Anforderungen und in welcher Form das LUH-Repo diese erfüllt.
- Bei größeren Projekten kann ein DMP kaum jedes Detail für jedes Teilprojekt oder jede Partnerinstitution beschreiben: Wäre es nicht sinnvoller, wenn jeder Projektpartner einen eigenen DMP erstellt, ergänzend zum Gesamt-DMP? → Im konkreten Projekt wurde der EU-DMP teils auf Projektpartner heruntergebrochen, aber er muss auch nicht ins kleinste Detail alles für alle Projektpartner erklären. → Auf EU-Ebene ist der DMP eher eine Richtlinie, daher ist die Empfehlung, diesen eher allgemein zu halten und zusätzlich auf Ebene der Projektpartner detaillierte DMP zu erstellen.

- Wie geeignet ist Zenodo als Repositorium? → Grundsätzlich erfüllt das generische Repositorium Zenodo die EU-Anforderungen an ein trusted repository. Dennoch ist es empfehlenswert, zunächst zu schauen, ob es nicht ein geeignetes fachspezifisches Repositorium gibt. Bei großen Verbundprojekten mit internationalen Partnern ist Zenodo auch deshalb durchaus sinnvoll, weil im LUH-Repo nur LUH-Mitarbeitende Daten publizieren können.

Theorie und Praxis des Datenmanagements im EU-Projekt SELINA und in der Abteilung Physische Geographie & Landschaftsökologie (Prof. Dr. Benjamin Burkhard)

(zu Einzelheiten siehe Präsentation)

Der Datenmanagementplan und die Richtlinie zum Datenmanagement wurden vom für Kommunikation und Ergebnisverbreitung im Projekt verantwortlichen Projektpartner entworfen. . Darin werden unter anderem Standards für die Daten- und Metadatenformate festgelegt oder empfohlen. In der Abteilung Physische Geographie und Landschaftsökologie wurde ebenfalls eine FDM-Richtlinie verabschiedet und eine Vorlage für Datenmanagementpläne ausgearbeitet. Obwohl alle Dokumente gut verständlich sind, zeigt sich in der Praxis sowohl im Projekt als auch in der Abteilung, dass es bei der gewissenhaften Umsetzung große Unterschiede gibt, je nach individueller Motivation der Beteiligten.

Fragen und Anmerkungen

- Mit welchen Systemen werden die Metadaten am Institut verwaltet? → Derzeit wird die strukturierte Metadatenerfassung in Word-Dokumenten entworfen. Wenn die finale Version steht, soll ein Datenbank-Backend mit einer grafischen Oberfläche zur Eingabe und Abfrage der Metadaten entwickelt werden.
- Wie können Mitarbeitende zum Datenmanagement motiviert werden? → Je weniger manuelle Arbeit notwendig ist, desto besser. Wo immer etwas automatisiert oder vorausgefüllt werden kann, sollte diese Möglichkeit genutzt werden. → Oft sind Metadaten allerdings sehr heterogene Daten, sodass es kaum Möglichkeiten zur Automatisierung gibt.

Wie Messdaten und ihre Metadaten standardisiert erfasst und aufbewahrt werden können - Ein Praxisbeispiel (Dipl. Sozialwiss. Stefan Warnken)

(zu Einzelheiten siehe Präsentation)

In einem Institut fallen viele Terabyte Messdaten plus Metadaten in unterschiedlichen Formaten an, deren Speicherung und ggf. Zugänglichmachung eine Herausforderung darstellen. Nach einer Testphase konnten geeignete Speicherformate identifiziert und Standards für Daten und Metadaten festgelegt werden. Dennoch ist es schwierig, die Anwendung dieser Standards allgemein durchzusetzen. Wichtigster Hebel ist daher das Hervorheben der vielen Vorteile, die mit deren Einhaltung verbunden sind.

Fragen und Anmerkungen

- Gibt es Bemühungen, die Umsetzung des FDM zukünftig offiziell verpflichtend zu machen und dafür Personal einzusetzen? → Eine Verpflichtung zum Datenmanagement und zur Erhebung der Metadaten ist am Institut derzeit aufgrund der Struktur der Fakultät und der weitgehenden Unabhängigkeit der Institute nicht umsetzbar.
- Braucht es ein fachspezifisches Datenrepositorium für Bauingenieure oder reicht das generische LUH-Datenrepositorium? → Grundsätzlich reicht das LUH-Repositorium, aber für große Datenmengen ist eine S3-Speicherlösung (kommerzieller Cloudspeicher von Amazon) oder Vergleichbares notwendig, was auch für die NFDI interessant sein könnte.

Fragen und Diskussion

- Wer hat in seinen Projekten/Instituten bereits FDM-Richtlinien oder klar definierte Rollen und Verantwortlichkeiten? Welche Erfahrungen gibt es damit?
→ An einem Institut gibt es derzeit die Bemühung, alle Analyseskripte auf einem gemeinsamen

GitLab zu sammeln, um gemeinsame Routinen zu etablieren. Diese Bemühungen hängen jedoch vom Engagement einzelner Personen ab, und auch die Umsetzung des FDM ist sehr personenabhängig. Ein wirksamer Hebel, alle vom Mitmachen zu überzeugen, ist noch nicht gefunden.

→ In zwei SFBs mit INF-Projekten wird die Erfüllung von in internen Richtlinien vereinbarten FDM-Mindeststandards von der Geschäftsleitung unterstützt. Die intrinsische Motivation der Forschenden ist aber unterschiedlich ausgeprägt. Einige investieren viel Zeit und Energie, andere machen nur das Nötigste. Für den Betrieb der technischen Systeme zur Daten- und Metadatenerfassung gibt es spezielles Personal im Projekt und Unterstützung durch die TIB.

- Es fehlt ein zentrales Git-Repository, in dem das Rechteverwaltung vollumfänglich genutzt werden kann und das auch direkt eine Veröffentlichung ermöglicht. Dazu ist derzeit ein Datenexport nach GitHub nötig.

Termin- und Themenplanung für Folgeveranstaltungen

Die kommende Veranstaltung wird am 24. Oktober 2024, 14:00-18:00 Uhr in einem neuen Format und in Präsenz stattfinden. Thema: Wie schützt man Daten durch Verschlüsselung? (in Kooperation mit dem Chaos Computer Club)