

Dr. Karsten Stephan und René Reitmann

Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (DZHW), Hannover

Abstract

Lessons Learned: Entwicklung und Betrieb eines Metadatenmanagementsystems für Forschungsdaten aus dem Bereich der Hochschul- und Wissenschaftsforschung

Von März 2015 bis Mai 2017 wurde, gefördert durch das Bundesministerium für Bildung und Forschung, am Deutschen Zentrum für Hochschul- und Wissenschaftsforschung ein Forschungsdatenzentrum (FDZ) für das Forschungsfeld der Hochschul- und Wissenschaftsforschung aufgebaut. Das FDZ nimmt sowohl anonymisierte Daten quantitativer Befragungen als auch anonymisierte Daten qualitativer Interviews auf, dokumentiert diese und stellt sie der wissenschaftlichen Community für Forschungs- und Lehrzwecke zur Verfügung.

Zur Unterstützung der Datenherausgabe, der Datenaufnahme und weiterer DZHW-interner Prozesse des Forschungsdatenmanagements wurde ein IT-technisches „Metadatenmanagementsystem“ benötigt. Aus Sicht der Nutzenden sollte das System eine moderne webgestützte Datensuche für Forschende bereitstellen. Zudem sollten die Prozesse der qualitätsgesicherten Erfassung von Datendokumentationen und Metadaten unterstützt werden. Beide Funktionen können die Beratungsnotwendigkeiten gegenüber Datennutzenden und Datengebenden reduzieren. Das System sollte skalierbar sein und Schnittstellen zu einer Reihe weiterer Anwendungen (z.B. das Onlinebefragungssystem des Hauses) und Dienste (z.B. zur Registrierung von Digital Object Identifiern) bieten. Das Metadatenmanagementsystem soll als zentrales Metadatenrepositorium für die Forschungsdaten des Hauses fungieren und somit auch die Basis für die Erhöhung des Standardisierungs- und Automatisierungsgrades im Rahmen des hauseigenen Forschungsdatenmanagements bereitstellen.

In den Jahren 2015 bis 2017 wurden unter einer Open Source Lizenz in einem offenen, iterativen Entwicklungsmodell (Scrum, DevOps, Continuous Integration) mit modernen informationstechnischen Technologien (AngularJS, MongoDB, Elasticsearch) der zentrale Datenspeicher, die Prozesse zur Erzeugung der zu importierenden (Meta)Daten sowie die Suchfunktionalität implementiert. Diese Komponenten gingen mit Eröffnung des FDZ im Juni 2017 in den Produktivbetrieb. Von Juni 2017 bis Mai 2019 wurden, teilweise mit verringertem Personalaufwand, die Komponenten zur Unterstützung der qualitätsgesicherten Datenaufnahme implementiert. Die entsprechenden Funktionalitäten werden während des laufenden Betriebs schrittweise in das Produktivsystem übernommen.

Der Vortrag geht neben einer kurzen Demonstration der Kernfunktionalitäten des Systems auf die einzelnen Phasen des Entstehungsprozesses und auf die jeweils gewonnenen Erfahrungswerte ein. Thematisiert werden die Konzepterstellung, die Beantragung, der Entwicklungsprozess sowie die Prozesse zum Treffen technologischer und fachlicher Entscheidungen. Im Vortrag wird ebenfalls auf den bisherigen Betrieb, die Wartung und die Weiterentwicklung des Systems eingegangen.

Metadatenmanagementsystem: <https://metadata.fdz.dzhw.eu>

Entwicklungsrepository: <https://github.com/dzhw/metadatamanagement>