

Digital Zusammenwachsen: Forschungsdaten- management im Forschungsverbund MWW

,
steyer@hab.de
Forschungsverbund MWW

,
lydia.koglin@klassik-stiftung.de
Forschungsverbund MWW

,
fritz@dla-marbach.de
Forschungsverbund MWW

Das Poster stellt ein Konzept für das digitale Zusammenwachsen heterogener Datenbestände dreier Gedächtnisinstitutionen zur Diskussion vor. Im Zentrum steht das Forschungsdatenmanagement, welches neben Aspekten der Datenmodellierung und Beschreibung von heterogenen Datenbeständen über Metadaten, auch die Präsentation, Benutzung und Langzeitarchivierung der digitalen Bestände umfasst.

Der durch das BMBF geförderte Forschungsverbund Marbach Weimar Wolfenbüttel, bestehend aus dem Deutschen Literatur Archiv Marbach, der Klassik Stiftung Weimar und der Herzog August Bibliothek Wolfenbüttel, entwickelt seit 2014 einen gemeinsamen virtuellen Forschungsraum. Auf diesem sollen die einzigartigen wissenschaftlichen Bestände der drei Einrichtungen gemeinsam präsentiert und der Forschung zur Verfügung gestellt werden. Durch Methoden der Digital Humanities werden neue Zugänge zu den digitalen Forschungsdaten des Verbundes geschaffen und die Implementierung innovativer Tools wird neue Forschungsfragen an das Material ermöglichen. Dabei wird die Sicherung der wissenschaftlichen Erträge durch die Etablierung eines verlässlichen Speichers garantiert, der überdies die Langzeitarchivierung digitaler Objekte übernehmen wird. Langfristig soll der Forschungsraum zu einem universellen Rechercheinstrument, einem virtuellen Arbeitsplatz und einem Repositorium für Forschungsergebnisse für die Verbundeinrichtungen ausgebaut werden.

Die MWW-Forschungsumgebung ist dabei als Teil der digitalen Infrastrukturen der drei Einrichtungen zu sehen, über den die verteilten Ressourcen aggregiert werden. Der Forschungsraum soll dezidiert so gestaltet werden, dass er mit anderen Forschungsumgebungen kompatibel ist, aber trotzdem die Spezifika der Bestände

der Verbundeinrichtungen berücksichtigt und den digitalen Geisteswissenschaften zur Verfügung stellt.

Funktionsumfang und Schnittstellen des Forschungsraums orientieren sich dabei an den historischen Sammlungsschwerpunkten der Verbundeinrichtungen. Im Gegensatz zu universell ausgelegten Forschungsinfrastrukturen handelt es sich daher bei dem MWW-Forschungsraum um eine bestandsbezogene Infrastruktur. Dies hat zur Folge, dass nicht die Produktion neuer Forschungsdaten und Workflows im Vordergrund steht, sondern die Integration von bereits vorhandenen Daten und die Verknüpfung bereits bestehender Arbeitsprozesse. Für die Verwirklichung dieser Ziele bedarf es eines komplexen Forschungsdatenmanagements, um den gemeinsamen Zugang zu ermöglichen. Die Herausforderung besteht in der Bildung eines interoperablen Datenpools, welcher die technische Heterogenität der Daten und die unterschiedlichen Workflows bei der Produktion, Präsentation und Sicherung der jeweils eigenen Forschungsdaten für den Forschungsraum meistert. Die Datenbasis bildet die Modellierung und Beschreibung der digitalen Informationseinheiten über eine Metadatengrammatik. Neben den zentralen Nachweissystemen mit standardisierten Metadaten gibt es in den Einrichtungen diverse Fach- und Spezialdatenbanken, die technisch und inhaltlich häufig Insellösungen darstellen: Daten wurden und werden kontinuierlich erhoben oder stellen Ergebnisse von abgeschlossenen Projekten dar. Das Ziel ist es, aus diesen inhaltlich hochwertigen aber technisch nicht aktuellen Standards entsprechenden digitalen Daten der Verbundeinrichtungen, Forschungsdaten zu generieren, also gut strukturierte, über etablierte Metadatenstandards erschlossene und somit auch leicht zu prozessierende Daten.

In der Regel ist aber bei der Erhebung dieser Forschungsdaten nicht ein Data-Reuse-Szenario konzipiert worden ist. Zu diesem Szenario gehört aber nicht nur die Aufbereitung der Forschungsdaten für den MWW-Forschungsumgebung, sondern auch deren Bereitstellung für externe Dienste über standardisierte Schnittstellen. Des Weiteren wird ein aktiver Austausch mit anderen Forschungsumgebungen angestrebt.

So sind auch die langfristige Sicherung und Nachhaltigkeit dieser Forschungsdaten ein Bestandteil der digitalen Infrastruktur. Dabei werden nicht nur die Daten des Forschungsraums, sondern alle langzeitarchivwürdigen Daten der Verbundeinrichtungen in einem verlässlichen Speicher archiviert. Das Modell eines verteilten Speichers von drei Gedächtniseinrichtungen und die sich dadurch ergebenen Synergieeffekte sollen hier verdeutlicht werden. Damit begegnen die Verbundeinrichtungen auch der Herausforderung der stetig wachsenden Zahl an Forschungsdaten und dem Wunsch der Wissenschaftler nach langfristiger Verfügbarkeit ihrer Daten.

Ergänzt wird der virtuelle Forschungsraum durch eine „Workbench“ zur Bearbeitung und Analyse der

Forschungsdaten. Dabei wird jedoch nicht primär eine Neuentwicklung von Tools das Ziel sein. Vielmehr wurden durch Umfragen mit Digital Humanists und „traditionellen“ Geisteswissenschaftlern der Bedarf an möglichen DH-Tools und -Funktionalitäten ermittelt. Die Auswertung zeigte, dass bereits eine Vielzahl an Tools existiert, welche die Wissenschaftler für ihrer Arbeit unterstützen. Allerdings scheint es, dass diese Tools entweder nicht bekannt sind oder ihre Nutzung als zu kompliziert erachtet wird, um diese nebenbei zu erlernen. Daher konzentriert sich der MWW-Forschungsraum auf die Integration bereits vorhandener und etablierter freier DH-Dienste bzw. Funktionalitäten, welche zusammen mit umfangreichen Tutorials und Best-Practice-Beispielen im Forschungsraum angeboten werden sollen. Gerade im Bereich der Tools ergeben sich umfangreiche Kooperationspotentiale mit anderen Forschungsumgebungen, da der Forschungsverbund über eine große Menge an wissenschaftlich hochwertigen Datenmaterial verfügt, dafür über bisher über nur sehr wenige Tools für die Analyse und Visualisierung der Daten.

Im Zentrum der Präsentation steht das Zusammenspiel der soeben skizzierten drei Komponenten Metadatengrammatik, virtueller Forschungsraum und verlässlicher Speicher. Die Projekte haben die Konzeptionsphase hinter sich und würden die DHd-Tagung dazu nutzen, das entwickelte Konzept für die erste Förderphase, die Mitte 2018 endet, zu präsentieren und zur Diskussion zu stellen.

MMW Forschungsverbund (2013-*):
Forschungsverbund Marbach Weimar Wolfenbüttel <http://www.mww-forschung.de> [letzter Zugriff 15. Februar 2016].