

Von Primärquellen zu Forschungsdaten. Entwicklung digitaler Konzepte zur Integration geschichtswissenschaftlicher Arbeits- und Datenmanagementprozesse

Referentin: Marina Lemaire

Auf Pergament oder Papier überliefertes Schriftgut in Archiven, als Audio-Dateien überlieferte Zeitzeugen-Interviews aus Oral-History-Projekten, News und Posts aus sozialen Medien, usw.: Die Reihe unterschiedlicher Quellenarten lässt sich je nach Epochen und Forschungsgegenstand beliebig fortsetzen. Die Zusammenführung unterschiedlicher Primärquellen in einem Untersuchungskorpus, dessen Erfassung, Erschließung und Auswertung unter Anwendung fachspezifischer Methoden und Standards erfolgt, sind angesichts der für die Geisteswissenschaften typischen Heterogenität der Dokumenttypen, ihrer disparaten Überlieferungslage sowie der oftmals eingeschränkten Zugänglichkeit sehr zeit- und arbeitsintensive Prozesse. Zugleich sind die Quellenkorpora – seien sie in einer Qualifikationsarbeit oder in einem Großprojekt entstanden – ein wertvoller Schatz, der nicht nur die *eine* eigene Forschungsfrage beantwortet, sondern auch disziplinäre als auch interdisziplinäre Perspektiven für nachfolgende Forschungen und Forschergenerationen eröffnet. Diese Option zur Nachnutzung entsteht jedoch nur, wenn Forschungsprojekte allgemeine und fachspezifische Anforderungen eines nachhaltigen Forschungsdatenmanagements berücksichtigen. Dazu müssen parallel zu den verschiedenen Phasen des Forschungsprozesses die für die Nachnutzung relevanten Schritte des Datenmanagementprozesses im Arbeitsprogramm vorgesehen werden.

Bislang liegen Erschließungsinformationen zu den Quellen eher in analoger bzw. schwer bis nichtmaschinenlesbarer Form vor. Sie sind zu großen Teilen nicht einfach auffindbar, da sie sich nur mühevoll und in reduzierter Form aus den Quellenverzeichnissen sowie den Anmerkungen der Publikationen rekonstruieren lassen. Effizienter wäre es, während der Quellenerschließung – unterstützt von einem Softwaresystem – die Erschließungsinformationen (Metadaten) in strukturierter Form aufzunehmen und sie in einer maschinenlesbaren Form zu speichern, um sie mit ForscherInnen auszutauschen oder für künftige Forschungsvorhaben verfügbar zu machen. Denn durch die Bearbeitung einer Quelle in unterschiedlichen Kontexten werden immer tiefergreifende Erschließungsinformationen erhoben und das Wissen zu dieser Quelle wächst. Ein solches Vorgehen hat nicht nur positive Effekte für die Nachnutzenden, sondern auch für die eigene Forschungsarbeit. Diese Ziele können nur erreicht werden, wenn der Forschungsprozess und die Quellenaufbereitung, also die strukturierte Erfassung und Beschreibung, in einem integrativen digitalen Arbeitsprozess zusammengeführt werden. Perspektivisch sollte das Erlernen dieser digitalen Arbeitsweise bereits in der HistorikerInnenausbildung erfolgen.

In der Praxis ist für die Integration des Forschungs- und Aufbereitungsprozesses die Planung eines strukturierten Arbeitsablaufs notwendig, der ausgehend von der Forschungsfrage sowie den dafür notwendigen Quellen und Methoden die Arbeitsschritte zur Quellenbearbeitung inklusive der benötigten Werkzeuge (Software, Hardware) definiert. Der Vortrag wird einen praxisnahen Einblick in die Erstellung eines solchen digitalen Forschungskonzeptes in einem geschichtswissenschaftlichen Projekt geben. Anhand der typischen Arbeitssituationen - Quellenerschließung und -analyse - wird am Beispiel der Virtuellen Forschungsumgebung FuD (fud.uni-trier.de) aufgezeigt, wie man langfristig einen Quellenbestand über mehrere Forschungsprojekte hinweg sukzessive aufbauen kann. Dabei wird erläutert, wie Primärquellen mit Erschließungsinformationen angereichert, untereinander vernetzt, annotiert und schließlich mit Informationen zum Entstehungskontext archiviert werden. Es wird gezeigt, wie aus Primärquellen, dem Rohmaterial der Forschung, komplexe, nach fachspezifischen Standards ausgezeichnete und in einem nachnutzbaren XML-Format (TEI) gespeicherte Forschungsdaten entstehen.