

Von der Käfighaltung zu glücklichen Daten

Wie Wissen zukunftsfit wird

Wie können Daten von Forschungsinstitutionen Forscher*innen erreichen?

Als 2011 Handkeonline als Rechercheplattform zum Werk des österreichischen Schriftstellers Peter Handke konzipiert wurde, lag der Fokus auf der Präsentation des biografischen, literaturwissenschaftlichen und archivalischen Wissens. Die Nachhaltigkeit der Modellierung, der Verfügbarkeit und der Wiederverwendbarkeit der Forschungsdaten spielten damals keine Rolle. Die Datenbank wird bis heute gepflegt; das darin enthaltene Wissen war aber lange weder mit anderen Institutionen teilbar noch maschinell verwertbar. Eine Verwendung in anderen Forschungskontexten war also ausgeschlossen.

Notwendigkeit der Migration

Ziel ist die Transformation sämtlicher für Handkeonline generierter Beschreibungsdaten zum Werk Handkes ins Format TEI/XML und die Bereitstellung eines werkgenetischen Katalogs als Linked Open Data-Datensatz. Dabei werden die Richtlinien zur Modellierung der FRBRoo angewendet. Die Daten werden aus diesem Grund in die von der Österreichischen Nationalbibliothek errichtete Nachhaltige Infrastruktur für digitale Editionen migriert und ein Re-Design der Datenstruktur wurde vorgenommen.

Handkeonline

Die Forschungsplattform Handkeonline wurde im Zuge eines FWF-Forschungsprojekts am Literaturarchiv der Österreichischen Nationalbibliothek Wien 2011 – 2015 entwickelt. Sie versteht sich als virtuelles Archiv, das alle Materialien des Vorlasses, der verstreut in öffentlichen und privaten Archiven in Österreich, Deutschland und der Schweiz vorliegt, mit dem veröffentlichten Werk Peter Handkes zusammenführt. Der Vorlass umfasst Notizbücher, annotierte Bücher, Fotos, Landkarten oder sogar Wanderstöcke, Bleistiftmanuskripte und Typoskripte unterschiedlicher Textfassungen, aber auch Druckfahnen, Aushänger und vieles mehr. Die Materialien werden tabellarisch und in Form von Texten beschrieben und in eine entstehungschronologische Beziehung gesetzt, die in Paratexten kommentiert wird. Als Ergänzung fungiert eine Bibliografie des umfangreichen Primärwerk Handkes, das sich in über 140 Büchern über sämtliche Genres erstreckt, sowie der relevanten Forschungsliteratur dazu.

Da die Daten sehr grob strukturiert vorlagen, wurden sie maschinell für eine weitere feingranulare Verarbeitung ausgewertet und aufbereitet.

Die Daten wurden mittels XSLT und Python in das Zielformat gebracht.

Mit dem Wechsel der Infrastruktur werden wiederverwendbare, Software-Komponenten der Plattform für die Online-Präsentation genutzt. Dabei wird mit aktuellen Bibliotheken und Frameworks gearbeitet, um ein responsive, intuitiv nutzbare Benutzeroberfläche anbieten zu können.

Export

Preprocessing

Modellierung

Transformation

Anreicherung

Integration

Der gesamte Datensatz, bestehend aus Handkes Werken und seinen Notizbüchern, Fotos und Scans der dazugehörigen Materialien und einer umfassenden Bibliografie, wurde aus der relationalen Datenbank exportiert.

Es wurde eine TEI-Datenmodell entwickelt, das die Daten in ihrer Heterogenität repräsentieren kann und kompatibel mit der Zielinfrastruktur ist.

Personen, Orte und Werke wurden mit Normdaten versehen und sind deshalb eindeutig zuordenbar. Außerdem wurden Bezüge zwischen Werken und Materialien im Datenmodell hergestellt.

