

# DIGITALER ARBEITSPLATZ

Von der Bibliothek zur Research Infrastructure:  
Mit den Digital Humanities verändern sich in den  
Geisteswissenschaften auch die Forschungsmethoden.

Eine gut sortierte Bibliothek und ein ausreichendes Reisebudget für Archivrecherchen waren in den Buchwissenschaften lange Zeit eine wichtige Grundlage für die Forschungsarbeit. „Mit der zunehmenden Digitalisierung der Forschungsdaten, sprich historischer Dokumente, ändert sich das komplette ‚Forschungsökosystem‘. Wenn sich die Geisteswissenschaften digitale Daten zunutze machen wollen, geht es nicht ohne neue Formen der Forschungsinfrastruktur“, sagt Günter Mühlberger, Leiter der Gruppe Digitalisierung & Elektronische Archivierung am Institut für Germanistik. Er arbeitet mit seinem Team an einer neuen Virtuellen Forschungsumgebung (VFU) zur Texterkennung historischer Dokumente. „Die Digitalisierung des Buchbestandes ist in vielen Bereichen schon weit fortgeschritten. Die Deutsche Forschungsgemeinschaft etwa hat über die sogenannten VD15-VD18-Projekte rund 300 Millionen Seiten digitalisieren lassen. Nun beginnen im internationalen Bereich auch vermehrt Archive damit, die bei ihnen gelagerten Dokumente flächendeckend zu digitalisieren“, erzählt Mühlberger.

Dass die historischen Dokumente damit auch suchbar werden, ist inzwischen ein Gemeinplatz. Doch die Möglichkeiten der Digital Humanities enden nicht beim gescannten Dokument mit einem mehr oder weniger ausgefeilten Suchinterface. „Das ist nur eine schwache Mimikry der alten Bücherwelt“, sagt der Germanist. „In der digitalen Welt können wir Dokumente auf ganz andere Art und Weise verarbeiten als bisher.“ Informationsex-

traktion, Data Mining und Big Data sind hier die Schlagwörter. „Große Datenmengen automatisiert nach sogenannten Entitäten wie Personennamen, Ortsnamen, Datumsangaben, Zahlenkombinationen und Listen ausgewählter Stichworte auszuwerten und diese Daten in menschen- und maschinenlesbarer Weise verfügbar zu machen – da fängt dann eine quantitativ verstandene Geisteswissenschaft erst an.“ Dies zu ermöglichen, darauf zielen neue Plattformen für die Digital Humanities ab. Sie stellen eine ganze Reihe von Werkzeugen zur Verfügung, mit denen die digitalen Daten durchforstet, analysiert und bearbeitet werden können.


## Zukunftssicher forschen

Die von der EU mit acht Millionen Euro geförderte Plattform zur Texterkennung, die Mühlberger mit einem internationalen Team derzeit aufbaut, versteht sich als eine Virtuelle Forschungsumgebung, die es den Geisteswissenschaften ermöglicht, die vorhandene Software für ihre speziellen Zwecke zu trainieren und so etwa Handschriften der frühen Neuzeit automatisiert erkennen zu lassen, oder aber anhand von Schriftproben bisher unbekannte Dokumente von berühmten Persönlichkeiten ausfindig zu machen.

„Für viele Forschende in den Geisteswissenschaften sind Texte die Hauptquellen ihrer Arbeit“, sagt Günter Mühlberger. „Deshalb hoffen wir, dass wir mit unserer Transkribus-Plattform einen Beitrag zu einer effizienteren Arbeit mit diesen Dokumenten leisten können.“ Dies geht einher mit den Leitlinien für die Zukunft der digitalen Geisteswissenschaften in Österreich, die die Österreichische Akademie der Wissenschaften kürzlich formuliert hat und in denen auch die Bereitstellung von stabilen, betreuten Tools und Services



**GÜNTER MÜHLBERGER:** „Im internationalen Bereich beginnen vermehrt auch Archive, die bei ihnen gelagerten Dokumente flächendeckend zu digitalisieren.“

gefordert wird. Die Texterkennungsplattform verwendet etwa eine Schnittstelle, die von der Deutschen Forschungsgemeinschaft definiert wurde. „Dadurch können unsere Nutzer Dokumente unterschiedlicher Archive und Bibliotheken einsammeln und an ihrem digitalen Arbeitsplatz damit arbeiten. Später können die bearbeiteten Dokumente über die gleiche Schnittstelle wieder der Allgemeinheit zur Verfügung gestellt werden. Gleichzeitig sorgen wir dafür, dass auch die eigentlichen Forschungsdaten standardisiert langzeitarchiviert und veröffentlicht werden.“ cf 

Die Software Transkribus entschlüsselt historische Handschriften automatisch. [transkribus.eu/Transkribus/](http://transkribus.eu/Transkribus/)

