



READ

READ: Handschriften lesen mit Transkribus

Krakow, September 2017

Die automatisierte Erkennung von handschriftlichen Texten eröffnet den historisch arbeitenden Geisteswissenschaften Zugänge zu bislang unerschlossenen Quellen und unedierte Texten. Dank dem Einsatz von Sprachmodellen und Zeichen- bzw. Worterkennung werden handschriftliche Texte in guter Qualität erkannt und aufbereitet. Das von der europäischen Union geförderte Projekt READ mit der Plattform *Transkribus* vereint neueste Ansätze der Informatik in einem Tool, welches das Arbeiten mit Digitalisaten erlaubt, also geradezu jenen Wunsch von Forschenden zu befriedigen scheint, eines Tages große Textmengen nicht mehr selbst vollständig transkribieren zu müssen.

Um das Projekt in seinen Anwendungsbereichen vorzustellen und die Möglichkeiten für die eigene Forschung orientierten Fachpublikum ausloten zu können, veranstaltet das Staatsarchiv Zürich (beteiligt am READ-Projekt) einen dreistündigen Workshop, der allen Interessierten offensteht.

Im Workshop erfolgt eine Vorstellung der Möglichkeiten (und Beschränkungen) von *Transkribus*. Die Teilnehmenden können anhand von eigenen oder zur Verfügung gestellten Dokumenten das Tool austesten. Dabei wird praxisnah gezeigt, was nötig ist, um die automatisierte Texterkennung sinnvoll zu nutzen. Daneben werden auch weitere Forschungsansätze des Projekts READ vorgestellt.



Die vorgestellte Software *Transkribus* profiliert sich derzeit als funktionsfähigste Plattform, die den Bedürfnissen universitärer Forschung von Historikern und Philologen entgegen kommt. *Transkribus* wurde entwickelt, um alle jene zu unterstützen, die mit der Transkription historischer,

gedruckter oder handschriftlicher Dokumente befasst sind. Bei *Transkribus* handelt es sich nicht nur um eine Software, sondern um eine Reihe von Werkzeugen für die automatisierte Erfassung von Dokumenten zur Verfügung, darunter eine computergestützte Handschriftenerkennung (*Handwritten Text Recognition*), Bilderkennung (*Layout Analysis*) und Strukturerkennung (*Document Understanding*). Diese bedürfen der jeweiligen Einführung und der Adaption für die spezifischen Anwendungsbereich der Forscherinnen und Forscher.



READ

Programm (Vorschlag)

09:00 Begrüssung, Einführung und Vorstellungsrunde

09:15 Vorstellen des Tools *Transkribus* (**Tobias Hodel**, Staatsarchiv Zürich)

10:15 *Pause*

10:45 Austesten von *Transkribus* durch die Teilnehmenden

11:30 Fragen und Rückmeldungen zum Tool, Möglichkeiten zur Nutzung und zukünftige Anwendungsmöglichkeiten

Teilnehmende bringen ihre eigenen Rechner mit. Idealerweise wird bereits vor der Veranstaltung die Software *Transkribus* installiert (<http://transkribus.eu>). *Transkribus* basiert auf JAVA und läuft auf allen gängigen Systemen (Windows, MacOS X, Linux).

Im Workshop kann mit eigenen digitalen Dokumenten oder mit zur Verfügung gestellten Testdokumenten gearbeitet werden.

Eine Registrierung auf transkribus.eu ist notwendig.

Informationen zum Projekt finden sich unter read.transkribus.eu.

Für weitere Informationen steht Tobias Hodel (tobias.hodel@ji.zh.ch) zur Verfügung.



The READ project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 674943.