

Im Rahmen des Projekts tranScriptorium, das sich mit der automatisierten Erkennung historischer Handschriften beschäftigt, entwickelt die Projektgruppe „Digitalisierung und Elektronische Archivierung“ (DEA) am Institut für Germanistik der Universität Innsbruck eine Plattform mit deren Hilfe handschriftliche Dokumente in neuartiger Weise erschlossen werden können.

Die von uns im Rahmen es Workshops vorgestellte Plattform TRANSKRIBUS bietet die Möglichkeit, die herkömmliche Transkription handschriftlicher Texte auf eine neue Ebene zu heben, indem einerseits eine detaillierte Verlinkung zwischen Text und Bild erzielt wird (Block, Zeilen und Wortebene), als auch standardisierte Exportformate erzeugt werden können: TEI (Text Encoding Initiative) sowie PDF (Portable Document Format) zur lokalen Benutzung, aber auch METS (Metadata Encoding and Transmission Standard) für die Integration in Repositorien wie etwa FEDORA.

Gleichzeitig können nun die von Geisteswissenschaftlern produzierten Transkriptionen auch für das Training von HTR Verfahren genützt werden. Mithilfe der automatisierten Erkennung kann nicht nur die Transkription selbst unterstützt werden, sondern können auch noch nicht transkribierte, größere Mengen von Dokumenten automatisiert erkannt und somit durchsucht werden.

Der Workshop richtet sich an:

1. Geisteswissenschaftler (Historiker, Literaturwissenschaftler, Theologen, Juristen, ...), die mit der Transkription und Edition historischer Dokumente befasst sind.
2. Archivare und Bibliothekare (Sondersammlungen), die entweder selbst mit der Transkription handschriftlicher Texte zu tun haben, oder aber daran denken, dies im Rahmen eines Crowd-Sourcing Projekts zu organisieren. Die vorgestellte Plattform verfügt über ein Rechte- und Rollenmanagement, sodass mehrere Personen an einem Dokument arbeiten können, bzw. Dokumente auch für die Öffentlichkeit zur Verfügung gestellt werden können.
3. Studenten der Geistes-, Archiv- oder Bibliothekswissenschaften mit einem besonderen Interesse an der Transkription historischer Handschriften.

Wir erwarten uns von den Teilnehmern eine kurze Beschreibung der von Ihnen bisher durchgeführten Transkriptionsprojekte bzw. der geplanten Vorhaben in diesem Bereich. Alle Teilnehmer sind darüber hinaus aufgefordert, digitalisierte Dokumente zum Workshop mitzunehmen, bzw. diese bereits im Vorfeld auf die Plattform hochzuladen. Entsprechende Zugangsdaten werden von uns nach der Zusage zur Teilnahme zugestellt.

Der Workshop dauert ca. 3 Stunden und wird in folgende Bereiche gegliedert:

1. Einführung in das Thema Handwritten Text Recognition
Vorgestellt werden die grundlegenden Technologien und Tools, die der automatisierten Handschriftenerkennung zugrunde liegen.
2. Vorstellung der TRANSKRIBUS Plattform
Hier wird auf das grundlegende Konzept der TRP eingegangen und die Idee einer digitalen Infrastruktur zur Erkennung von Handschriften im Detail erläutert.
3. Einführung in das Expertentool „Transcribus“
Das für den Geisteswissenschaftler wichtigste Interface zur HTR Technologie im Rahmen der Transcription & Recognition Plattform ist das Tool „Transcribus“. Diese App ist mit JAVA und SWT programmiert und läuft lokal am Arbeitsplatzrechner. Bilder und Daten der Dokumente werden mittels einer Remoteverbindung zum TRP Server geladen und gespeichert, sodass ein Arbeiten in der „Cloud“ möglich ist.

4. Selbständiges Arbeiten mit der Plattform bzw. Transcribus

Auf diesen Punkt liegt der Fokus des Workshops.

Die Teilnehmer sollen die Möglichkeiten und Grenzen der Technologie in allen Einzelheiten am eigenen PC ausprobieren können. Zusätzlich gibt es auch die Möglichkeit diverse andere Anwendungen aus dem tranScriptorium Projekt selbst auszuprobieren.