

Günter Mühlberger

Universität Innsbruck, Digitalisierung und elektronische Archivierung (DEA)

Transkribus. Eine Plattform für die Digitalisierung, Transkription, Erkennung und Suche in historischen Dokumenten

Die Forschungsplattform Transkribus richtet sich an GeisteswissenschaftlerInnen sowie an Archive und Bibliotheken. Mittels Transkribus können Dokumente automatisiert erkannt und durchsucht werden. Dabei ist es unerheblich, ob es sich um gedruckte oder handschriftliche Dokumente, mittelalterliche Codices, frühneuzeitliche Protokolle oder Zeitungen des 18. Jahrhunderts handelt - sofern genügend Trainingsdaten vorliegen, können entsprechende Modelle trainiert und auf einzelne aber auch tausende von Dokumenten angewendet werden. Transkribus versetzt dabei den Benutzer selbst in die Lage, Trainingsdaten zu erstellen, Modelle zu trainieren, diese anzuwenden, die Erfolgsquote zu messen und schließlich auch zu durchsuchen.

Transkribus ist dabei als ein großes Netzwerk im Sinne der „republic of letters“ konzipiert: Je mehr Wissenschaftler, Archive und Bibliotheken die Plattform verwenden, desto größer ist die Wahrscheinlichkeit, dass bereits Daten für die Erkennung eines konkreten Dokumentes vorhanden und nachgenutzt werden können. Im go!digital-Projekt zum Wien[n]erischen Diarium wurde Transkribus mit großem Erfolg für die Erkennung einer typischen Zeitung des 18. Jahrhunderts eingesetzt. Das dabei erzeugte Modell für die Frakturschrift des 18. Jahrhunderts kann nun auch auf andere, ähnliche Dokumente angewendet werden.