ANWENDUNG VON ACTIVE MACHINE LEARNING ZUR AUTOMATISCHEN SCHREIBERERKENNUNG IN MANUSKRIPTEN DES 12. JAHRHUNDERTS

Projektträger:
Hochschule für Angewandte Wissenschaften St. Pölten GmbH

Wissenschaftliche Leitung:
Markus Seidl

Weitere beteiligte Einrichtungen:
Stift Klosterneuburg

Forschungsfeld:
Sammlungen Niederösterreich

Förderinstrument: Projekte Grundlagenforschung

Projekt-ID: FTI18-004

Projektbeginn: 01. März 2020

Projektende: folgt

Laufzeit: 36 Monate / laufend

Fördersumme: € 199.700,00

Kurzzusammenfassung:
Das heutige Niederösterreich hat aufgrund der Vielzahl der noch aktiven Klöster, deren Tradition seit ihrer Gründung im Mittelalter ungebrochen ist, eine herausragende internationale Stellung. Im letzten Drittel des 12. Jahrhunderts begannen die Mönche in den Klöstern, systematisch Bücher zu sammeln. Um das Jahr 1200 wurden große Bibliotheken gebaut, von denen die meisten noch heute erhalten sind. Etwa 120 Handschriften aus dieser Zeit befinden sich in Heiligenkreuz, mehr als 130 in Zwettl und fast 120 Codices in der Stiftsbibliothek Klosterneuburg.
Das Hauptziel unserer Arbeit ist es, ein besseres Verständnis der monastischen Schriftlichkeit im hochmittelalterlichen Österreich durch ein breiteres Wissen über die Schreiber zu ermöglichen, die in diesen frühen Wissenszentren gearbeitet haben. Es gibt jedoch keine Hinweise auf die Anzahl der Schreiber und ob diese nur in einem Kloster arbeiteten oder zwischen den Klöstern wechselten. Eine Möglichkeit, diese Faktoren zu bestimmen, besteht darin, die Schreibweisen zu analysieren, mit dem Ziel, verschiedene Schreiber in einer großen Anzahl von Manuskripten durch inhärente stilistische Merkmale ihrer Schriften zu identifizieren. Dies ermöglicht in weiterer Folge, die Aufenthaltsorte der Schreiber und die Organisation der Schreibstuben abzuleiten.
Die klassische Vorgangsweise der manuellen Schreibstilanalyse durch einzelne Experten ist ein langwieriger und zeitaufwendiger Prozess. Es besteht die Gefahr, dass die Ergebnisse dadruch geschwächt werden, dass sie subjektiv durch individuelle Eindrücke erzeugt werden. Es gibt erste Ansätze, um die Identifikation mittelalterlicher Schreiberhände durch maschinelles Lernen zu unterstützen. Diese sind jedoch für große Corpora nicht verwendbar.
Die größte Herausforderung ist das Fehlen einer umfassenden Ground Truth. In einem interdisziplinären Projekt mit Historikern und Informatikern schlagen wir einen Active Machine Learing Ansatz vor, der speziell menschliche Experten einbezieht, um das Machine Learning zu unterstützen und so eine zeiteffiziente Kennzeichnung von Schreibern in großen Corpora ermöglichen wird.