

ROBERT GRAMSCH-STEHFEST / VANESSA KRYPCZYK

Data Mining in der mittelalterlichen Matrikel der Universität Erfurt

Vorgestellt wird ein Pilotprojekt zur mittelalterlichen Erfurter Universitätsgeschichte, in welchem aus der Matrikel die kurzbiographischen Angaben zu 8.912 Studenten, welche zwischen 1392 und 1430 in Erfurt immatrikuliert wurden, semiautomatisiert extrahiert wurden. Der Scan der einschlägigen Edition von J. C. H. Weissenborn wurde mit OCR-Software digitalisiert, die resultierende Textdatei durch die Projektbearbeiterin nachbearbeitet und zuletzt mit einem sog. Parser in eine Datentabelle überführt. Kurz skizziert werden Auswertungsmöglichkeiten und mögliche Erweiterungen der so geschaffenen Datenbasis für eine vertiefte prosopographische Erforschung der mittelalterlichen Erfurter Universität. – *Dr. Robert Gramsch-Stehfest* ist außerplanmäßiger Professor am Historischen Institut der FSU Jena. *Vanessa Krypczyk* absolvierte 2022 ihren Bachelorabschluss Geschichte an der FSU und studiert gegenwärtig im Masterstudiengang Sammlungsbezogene Wissens- und Kulturgeschichte an der Universität Erfurt. – Zur Erfurter Universitätsgeschichte liegen von Robert Gramsch-Stehfest folgende monographische Werke vor: Robert Gramsch, *Erfurter Juristen im Spätmittelalter. Die Karrieremuster und Tätigkeitsfelder einer gelehrten Elite des 14. und 15. Jahrhunderts* (Education and Society in the Middle Ages and Renaissance 17), Leiden – Boston 2003; ders., *Erfurt – die älteste Hochschule Deutschlands. Vom Generalstudium zur Universität* (Schriften des Vereins für die Geschichte und Altertumskunde von Erfurt 9), Erfurt 2012. Zur mittelalterlichen Bildungsgeschichte allgemein: Robert Gramsch-Stehfest, *Bildung, Schule und Universität im Mittelalter*, Berlin – Boston 2019.

1. Aufgaben bei der Digitalisierung prosopographischer Angaben aus Matrikeln

Die Digitalisierung schreitet in allen Bereichen der Wissenschaft, auch in der Geschichtsforschung, schnell voran. Viele Fragen stellen sich, neue Herausforderungen und verlockende Perspektiven tun sich auf. Idealerweise sollte diese Digitalisierung nicht nur in großen, teuren Projekten funktionieren, sondern auch im Kleinen. Genau darauf zielte das Experiment, von dem im Folgenden zu berichten sein wird: Inwieweit kann gängige OCR-Software¹ dazu beitragen, die Digitalisierung und weitere automatisierte Aufarbeitung von Quellen auch in kleinen, eher informellen Projektformaten voranzutreiben? Wie kann der Workflow eines solchen Projektes effektiv gestaltet werden, was ist der Mehrwert für die Forschung? Welche Qualifikationsanforderungen stellen sich insbesondere für die beteiligten Historiker:innen, wie gestaltet sich die Projektarbeit im Spannungsfeld zwischen eigenständiger Forschungsleistung und informatisch-

¹ Zur Erläuterung siehe unten Anm. 11.

historischer Teamarbeit? Eine Bachelorarbeit an der Friedrich-Schiller-Universität in Jena im Bereich der mittelalterlichen Geschichte gab Gelegenheit, diese Fragen in einem konkreten Pilotprojekt zu studieren.²

Zielstellung der Arbeit war die Erstellung eines Datensatzes zu den Studenten der Universität Erfurt, welche sich in den ersten 40 Jahren der Existenz dieser Hohen Schule (1392–1432) in Erfurt immatrikuliert haben. Deren Matrikel ist bereits Ende des 19. Jahrhunderts in einer Edition (von Johann C. H. Weissenborn) erschlossen worden,³ auch gibt es eine Online-Präsenz dieser Publikation, sprich: einen Scan des Buchs, der mittels DFG-Viewer einsehbar ist. Doch kann diese digitale Ressource letztlich nur als online gestelltes Buch benutzt werden, da man in ihr lediglich „blättern“, aber nicht sinnvoll nach Personen oder gezielt nach Ortsnamen bzw. bestimmten Schlagwörtern suchen und erst recht nicht komplexere Anfragen stellen kann.⁴

Es ist hier nicht der Ort, eine ausführliche Einführung in die Quellengattung der Universitätsmatrikeln zu geben; einige Bemerkungen mögen genügen.⁵ Universitätsmatrikeln sind – jedenfalls dem Anspruch nach – vollständige Mitgliederverzeichnisse und mithin das zentrale Dokument einer Korporation, die sich ganz grundsätzlich als Personenverband der Lehrenden und Lernenden definierte (*universitas magistrorum et scholarum*). Wer an die Universität kam, hatte sich beim Rektor vorzustellen, ihm Vor- und Nachname, Herkunftsort und Stand zu nennen, einen Gehorsamseid zu leisten und an den Rektor und die Pedelle eine Immatrikulationsgebühr zu zahlen. Dies wurde in latinisierter Form in das Matrikelbuch eingetragen. Das Verlassen der Universität wurde in aller Regel nicht verzeichnet – man

² Beteiligt waren folgende Personen (alle Jena): Für die Thüringer Universitäts- und Landesbibliothek: Dr. Swantje Dogunke; für das Historische Institut: apl. Prof. Dr. Robert Gramsch-Stehfest sowie Frau stud. phil. Vanessa Krypczyk; für das Institut für Informatik: Johannes Mitschunas M.Sc.

³ Johann C. H. Weissenborn, *Acten der Erfurter Universität (1392–1636)*, 3 Bde. (Geschichtsquellen der Provinz Sachsen 8/1–3), Halle 1881, 1884, 1899.

⁴ Zugänglich unter: <https://www.uni-erfurt.de/amploniana/service/digitalisiert/> (Zugriff: 19.8.2022). Es findet sich hier keine Möglichkeit, im Digitalisat direkt nach Namen zu suchen – was vor einigen Jahren nach meiner (R. G.-St.) Erinnerung möglich war, als der Scan offenbar mit OCR durchsuchbar gemacht worden war. Falls diese Funktionalität noch existiert, ist sie zumindest nicht mehr leicht zu finden. Und auch wenn sie noch vorhanden ist, würde sie nicht das Grundproblem lösen, dass ein durchsuchbares PDF der Edition keineswegs eine nach allen Regeln der Kunst auswertbare Personendatenbank darstellt.

⁵ Vgl. insbesondere Maximilian Schuh, *Matrikeln*, in: Jan-Hendryk de Boer/Marian Füssel/Maximilian Schuh (Hg.), *Universitäre Gelehrtenkultur vom 13.–16. Jahrhundert. Ein interdisziplinäres Quellen- und Methodenhandbuch*, Stuttgart 2018, 103–117. Eine allgemeine Einführung in den thematischen Rahmen der Universitätsgeschichte gibt Robert Gramsch-Stehfest, *Bildung, Schule und Universität im Mittelalter (De Gruyter Studium: Seminar Geschichte)*, Berlin – Boston 2019.