

## Von Daten zu Wissen: Workshop zu ausgewählten DH-Tools

veranstaltet durch das DFG-Langfristvorhaben „Kommentierte digitale Edition der Reise- und Sammlungsbeschreibungen Philipp Hainhofers (1578–1647)“ an der Herzog August Bibliothek

Wolfenbüttel, Herzog August Bibliothek, 29.–30. September 2022

Ort: Seminarraum im Meißnerhaus

Organisation: Martin de la Iglesia (HAB), Timo Steyer  
(TU Braunschweig) und Michael Wenzel (HAB)

Der eineinhalbtägige Workshop soll eine Einführung in das Arbeiten mit ausgewählten Technologien bieten, die hochrelevant für eine Vielzahl verschiedener geisteswissenschaftlicher Forschungsfelder sind. In praxisorientierten Sessions erlernen die Teilnehmerinnen und Teilnehmer die Anwendung einiger Tools und wie diese in ihrem wissenschaftlichen Arbeitsalltag eingesetzt werden können. Zugleich soll die kritische Reflexion und Evaluierung dieser Arbeitsmittel angeregt werden. Verbindendes Element dieser Software-Lösungen ist ihre Unterstützung der geisteswissenschaftlich Forschenden in der Kette der Aktivitäten aus Datenerhebung, Datenverarbeitung und Datenauswertung, an deren Ende schließlich die Gewinnung neuen Wissens steht.

Ursprünglich für 2020 vorgesehen, musste die Veranstaltung damals aufgrund der Pandemie ausfallen und kann nun mit leicht verändertem Programm nachgeholt werden.

Die Teilnahme am Workshop ist kostenlos; um Anmeldung per Mail an [iglesia@hab.de](mailto:iglesia@hab.de) oder [t.steyer@tu-braunschweig.de](mailto:t.steyer@tu-braunschweig.de) wird gebeten.

### PROGRAMM

**Donnerstag, 29. September 2022**

17.00 Uhr

Begrüßung

Dr. Christian Heitzmann (Herzog August Bibliothek)

17.15 Uhr

Keynote: Digitale Tools in geisteswissenschaftlicher Forschung –  
Textanalyse, Interpretation und Visualisierung

Jan Horstmann (Münster)

Viele Tools in den Digital Humanities setzen entweder eine umfangreiche Einarbeitung in technische Abläufe und Befehle voraus, sodass sie nur von Personen mit entsprechender IT-Expertise bedient werden können, oder sie bieten eine Bedienoberfläche an, welche einen tieferen Einblick in die zugrundeliegenden Prozesse und Datenstrukturen – und damit eine Kritik derselben – nicht ermöglicht. Zudem orientieren sich Tools häufig nicht an genuin geisteswissenschaftlichen Arbeits- und Erkenntnisprozessen und werfen somit in der Regel neue Fragestellungen auf – und führen damit zu durchaus interessanten, in der Regel methodologisch oder technologisch ausgerichteten Diskussionen –, statt die ursprünglichen (geisteswissenschaftlichen) Fragestellungen zu beantworten. Im Vortrag soll beispielhaft demonstriert werden, wie bereits die Konzeption von DH-Tools vor dem Horizont einer geisteswissenschaftlichen Nutzung und unter Rückbezug auf theoretische Grundlagen geisteswissenschaftlicher Erkenntnisprozesse ablaufen kann, sodass Nutzende in die Lage versetzt werden, sowohl ihre ursprüngliche Fragestellung mithilfe digitaler Verfahren zu bearbeiten, als auch einen tiefergehenden digital-methodologischen Diskurs zu führen.

18.00 Uhr

Keynote: Toolkritik und Rezensionen geisteswissenschaftlicher Forschungssoftware am Beispiel NTEE

Ulrike Henny-Krahmer (Rostock)

DH-Tools ermöglichen, erleichtern und beschleunigen den Umgang mit digitalen Forschungsdaten und können ein wesentliches methodologisches Element im Forschungsprozess sein. Die durch sie ausgelöste zunehmende Formalisierung, Automatisierung und auch Effektivität geisteswissenschaftlicher Forschungsaktivitäten erfordert auf der anderen Seite aber auch eine kritische Einordnung und Auseinandersetzung mit den Werkzeugen, damit die Forschung transparent und nachvollziehbar bleibt und ihre Wissenschaftlichkeit gesichert werden kann. Erfolgt eine solche Kritik digitaler Tools in einem wissenschaftlichen Rahmen, trägt sie außerdem dazu bei, Tool-Entwicklung als Teil des wissenschaftlichen Arbeitens zu verstehen, zu würdigen und in den wissenschaftlichen

Diskurs einzubinden. Der Vortrag greift Rezensionen von geisteswissenschaftlicher Forschungssoftware, wie sie z.B. in der Zeitschrift RIDE und künftig in dem Journal CKIT publiziert werden, als eine solche mögliche Form der Toolkritik auf. Am Beispiel des Tools *NEISS TEI Entity Enricher (NTEE)*, einem Werkzeug zur automatischen Erkennung von Entitäten in TEI-Dokumenten mithilfe von Machine Learning-Verfahren, wird illustriert, in welcher Weise Software-Rezensionen zum wissenschaftlichen Einsatz von DH-Tools beitragen können.

## Freitag, 30. September 2022

09.30–11.00 Uhr

Einführung in OpenRefine mit praktischer Übung

Hanna Varachkina (Göttingen)

Teilnehmende bekommen einen Einblick in OpenRefine, ein leistungsstarkes Werkzeug für die Arbeit mit unordentlichen Daten. Sie lernen einige Funktionen von OpenRefine kennen, die sie zur Bearbeitung eines Datensatzes verwenden werden. Die Lektion basiert auf dem Lehrmaterial von *Library Carpentry*.

11.00–11.30 Uhr: Kaffeepause

11.30–13.00 Uhr

ResearchSpace und CIDOC-CRM: Knowledge Graph als Datenbank

Robert Casties (Berlin)

ResearchSpace ist ein Baukasten mit dem sich flexible Oberflächen für die Bearbeitung von semantisch strukturierten Daten schaffen lassen. In der Präsentation wird eine sehr kurze Einführung in RDF und CIDOC-CRM als Basistechnologien von Knowledge-Graphs und Linked-Data im Geisteswissenschaftlichen Bereich gegeben und dann anhand des Beispiels der ISMI Manuskript-Datenbank einen Überblick über die Vorteile und Herausforderungen dieser Technik. Für die Teilnehmer\*innen wird ein online-Zugang zu einer Test-Datenbank zum Ausprobieren bereitgestellt.

13.00–14.00 Uhr: Mittagspause

14.00–15.30 Uhr

Einführung in Structure from Motion

Astrid Schmörlzer (Bamberg)

Structure from Motion (SfM) ist eine Technik der photogrammetrischen Entfernungsbildgebung. Aus zweidimensionalen Bildsequenzen werden dreidimensionale Modelle erstellt. Diese 3D-Modelle erlauben besonders in der Epigraphik das genaue Studieren eines Objektes, dessen Verfügbarkeit stark eingeschränkt ist. Als virtuelles Modell kann es verschiedenen Personen zugänglich gemacht werden und erlaubt anhand der existierenden Möglichkeiten eine eingehende Analyse der Inschrift aber auch der übrigen Oberfläche. So können z.B. Neulesungen und -interpretationen von Inschriften festgestellt werden. In der kurzen Einführung soll ein einfacher Arbeitsablauf unter der Verwendung des Programmes Agisoft Photoscan/Metashape vorgestellt werden. Diese Software ist kostenpflichtig; eine kostenlose 30-Tage-Testversion ist unter <https://www.agisoft.com/downloads/installer/> erhältlich. Diese sollte vor Beginn des Workshops bereits installiert werden. Testdaten werden zum Download bereitgestellt.

15.30–16.00 Uhr: Kaffeepause

16.00–17.30 Uhr

Klassifikationsschema vom Entwurf bis zum Einsatz

Andreas Wagner (Frankfurt)

Die Session thematisiert, wie Klassifikationsschemata als Linked Data im Projektkontext entwickelt und eingesetzt werden können. Um beim Data Cleaning und in Annotationsprozessen ein dem Projektbedarf entsprechendes kontrolliertes Vokabular verwenden zu können, muss dieses nicht nur konzeptuell entwickelt werden, sondern es muss auch formal kodiert, in einer technischen Infrastruktur vorgehalten und in Arbeitswerkzeuge integriert werden. Mit SKOS, der Reconciliation API, GitHub, SkoHub, TEI Publisher und OpenRefine werden Standards, Plattformen, Tools und ein Workflow vorgestellt, in denen diese Anforderungen umgesetzt und das Vokabular zugleich als Ressource eigenen Rechts publiziert und nutzbar gemacht werden kann.