

# DigEdTnT – Digital Edition Creation Pipelines: Tools and Transitions

## Pollin, Christopher

christopher.pollin@uni-graz.at  
Zentrum für Informationsmodellierung, Graz

## Strutz, Sabrina

sabrina.strutz@uni-graz.at  
Zentrum für Informationsmodellierung, Graz

## Steiner, Christian

christian.steiner@dhcraft.org  
Zentrum für Informationsmodellierung, Graz

## Klug, Helmut

helmut.klug@uni-graz.at  
Zentrum für Informationsmodellierung, Graz

Digitale Editionen sind ein Kernbereich der Digital Humanities; sie machen historische Quellen zugänglich. Dabei werden computergestützte Methoden zur Umsetzung, Verbreitung und Erforschung von wissenschaftlich fundierten Quellenveröffentlichungen herangezogen. Digitale Editionen umfassen dabei textuelle, visuelle und ggf. auch quantitative Daten und erfordern oft spezielle Benutzeroberflächen, um domänenspezifische Forschungsfragen zu bearbeiten. Obwohl jedes Editionsprojekt seine eigenen spezifischen Anforderungen hat, lassen sich einzelne Schritte identifizieren, die für Editionsprojekte generell notwendig sind. Das ist im weitesten Sinne die Digitalisierung der Quelle mit der Verwaltung von Bildern und Text, die Transkription, die Modellierung relevanter Textphänomene mittels adäquater Auszeichnungssprachen, die Annotation semantischer Informationen und Named Entities, die Erstellung von Indizes, sowie eine den FAIR-Kriterien entsprechende Publikation über das Web. In den letzten Jahren wurde eine Vielzahl an Tools entwickelt, die für all diese Schritte eingesetzt werden.

Editionen bauen in der Regel auf Bilddigitalisaten der Quelle auf, deren Erstellung und Zurverfügungstellung im Aufgabenbereich von Bibliotheken und Archiven liegt. Ein Zugriff darauf ist im Idealfall mittels iiii möglich.

Die Transkription von Texten kann manuell, über Crowdsourcing (z. B. FromThePage) oder automatisiert (z. B. Transkribus) durchgeführt werden. Relevant dabei ist, dass unabhängig vom Werkzeug die Umwandlung des transkribierten Textes in XML/TEI möglich ist. Im bereits modellierten XML/TEI werden schließlich weitere Annotationen durchgeführt. Auch hier gibt es wieder eine breite Auswahl an Werkzeugen, deren Verwendung von projektspezifischen Anforderungen und Benutzergrup-

pen abhängig ist. Einige sind für spezielle Forschungsbereiche konzipiert, wie z. B. LaKomp, andere bieten eine grafische Oberfläche für Editor\*innen ohne tiefgreifende XML/TEI-Kenntnisse (CATMA). Wieder andere kombinieren eine Benutzeroberfläche mit bestimmten Funktionalitäten, wie z. B. einer Registerfunktion (ediarum). In einigen Anwendungsfällen bietet es sich darüber hinaus an, reines XML/TEI im Oxygen XML Editor zu schreiben.

Digitale Editionen produzieren Forschungsdaten und machen diese im Idealfall unter Einhaltung der FAIR-Prinzipien zugänglich. Dies erfordert die Einbindung von Normdaten oder kontrollierten Vokabularen. Werkzeuge hierfür können OpenRefine zur halbautomatischen Verknüpfung von Entitäten sein, oder Tools wie ba[sic]; stärker datenzentrierte Projekte verwenden Tools wie Fast Cat.

Die Veröffentlichung und Langzeitarchivierung kann schlussendlich über Repositories und Tools wie GAMS, ARCHE, teiPublisher oder ediarum.Web erfolgen. Für manchen Editionsprojekten ist es sinnvoll, domänenspezifische Werkzeuge oder APIs, wie z. B. correspSearch für Korrespondenzen, zu verwenden.

Ziel des Projekts *Digital Edition Creation Pipelines: Tools and Transitions (DigEdTnT)* ist es, Best-Practice-Pipelines und Tutorials für ausgewählte Tools und deren Übergänge (=Transitions) zu erstellen, die bei der Wahl der Tools und der Arbeit mit Tools zur Erstellung digitaler Editionen helfen sollen. Denn an den Übergängen ergeben sich mitunter besondere Herausforderungen, wenn beispielsweise Ergebnisse aus Transkribus nach ediarum zur weiteren Annotation überführt werden sollen. Das Projekt setzt dabei insbesondere auf eine Community-basierte Auseinandersetzung mit vorhandenen Tools. Daher sollen in zwei Workshops Nutzer\*innen und Entwickler\*innen zusammengeführt werden, um Anwendungsfälle und Feedback gemeinsam zu diskutieren. Wie die Übergänge zwischen einzelnen Tools abgewickelt werden können, wird letztlich in Tutorials und Guidelines sowie in Code Snippets beschrieben, die wiederum in einschlägigen Kontexten (KONDE Weißbuch, DARIAH Campus, etc.) zugänglich gemacht werden.

Das eingereichte Poster soll zum Projektstart von DigEdTnT die Diskussion zu diesem Thema eröffnen und interessierte Kolleg\*innen sowie auch Toolentwickler\*innen adressieren.

## Bibliographie

**Fafalios, Pavlos and Kostas Petrakis, Georgios Samaritis, et. al.** "FAST CAT: Collaborative Data Entry and Curation for Semantic Interoperability in Digital Humanities". In *Journal on Computing and Cultural Heritage (JOCCH)*. 14, 4, Article 45 (2021), 1-20. <https://doi.org/10.1145/3461460>.

**Fechner, Martin.** "Eine nachhaltige Präsentationsschicht für digitale Editionen." DHd, (2018). urn:nbn:de:kobv:b4-opus4-33277.

**Fritze, Christiane.** "Wohin mit der digitalen Edition?". In *Bibliothek Forschung und Praxis* edited by Achim Bonte et. al., 43(3), (2019), 432-440.

**Holstein, T., Störl, U.** "Towards Supporting Tools for Editors of Digital Scholarly Editions for Correspondences". In *HCI International 2020 – Late Breaking*

*Posters HCII 2020. Communications in Computer and Information Science* edited by C. Stephanidis, M. Antona, S. Ntoa., vol 1293. Springer, Cham. [https://doi.org/10.1007/978-3-030-60700-5\\_25](https://doi.org/10.1007/978-3-030-60700-5_25).

**Horstmann, Jan.** "Undogmatic Literary Annotation with CATMA" In *Annotations in Scholarly Editions and Research: Functions, Differentiation, Systematization* edited by Julia Nantke and Frederik Schlupkothen, 157-176. Berlin, Boston: De Gruyter, 2020. <https://doi.org/10.1515/9783110689112-008>.

**Klug, Helmut W. and Selina Galka and Elisabeth Steiner.** "KONDE Weißbuch im HRSM Projekt 'Kompetenznetzwerk Digitale Edition'". <https://www.digitale-edition.at>.

RIDE – A review journal for digital editions and resources, <https://ride.i-d-e.de>, 20.6.2022.

**Sahle, Patrick.** *Digitale Editionsformen. Zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels. Teil 2: Befunde, Theorie und Methodik.* Schriften des Instituts für Dokumentologie, 2013. <https://kups.ub.uni-koeln.de/5352>.

**Söring, Sibylle.** Technische und infrastrukturelle Lösungen für digitale Editionen: DARIAH-DE und Text-Grid. In *Bibliothek Forschung und Praxis* edited by Achim Bonte et. al. 40(2), (2016), 207-212. <https://doi.org/10.1515/bfp-2016-0040>.