

TEI (Text Encoding Initiative)

Eibinger, Julia; julia.eibinger@uni-graz.at

Die *Text Encoding Initiative* hat sich die laufende Weiterentwicklung von Standards für digitale Textkodierung und deren Aufrechterhaltung zur Aufgabe gemacht. Ihren Ursprung hat die TEI in einer Konferenz von 1987, aus der im Folgejahr ein Projekt und schließlich das heute bestehende Konsortium entstand.

Seit ihrer Entstehung hat die TEI insgesamt fünf Versionen ihrer Richtlinien zum Umgang mit digitalen Textressourcen veröffentlicht. Ende der 1990er-Jahre sprach die TEI eine explizite Empfehlung für XML aus und widmete die vierte Fassung ihrer Guidelines auch ganz den nötigen Anpassungen für diese Markup-Sprache. Heute wird die TEI als Quasi-Standard für digitale Texte verstanden, die in XML kodiert sind. Diese enge Verbundenheit resultiert nicht zuletzt aus den personellen Überlappungen, da Mitglieder des *Technical Council* der TEI stark an der Entwicklung von XML beteiligt waren. (TEI P5 2020, S. xxiii) Allerdings ist die TEI als eigenständige Markup-Sprache nicht abhängig von XML und könnte in Zukunft für andere Markup-Sprachen ebenso adaptiert werden, wie das Ende der 1990er-Jahre in Bezug auf XML der Fall war. (TEI P5 2020, S. xvi)

Der große Vorteil einer solchen standardisierten Diskussionsgrundlage für Markup-Sprachen ist das Potenzial für Austausch zwischen einzelnen Wissenschaftlerinnen und Wissenschaftlern, aber auch ganzen Forschungsprojekten und -teams. Dabei kann unabhängig von System oder Applikation gearbeitet werden und Daten können problemlos transferiert werden. Zudem bietet die TEI ein Framework für unterschiedliche Textsorten und Anwendungsfälle. Anwenderinnen und Anwender müssen also nicht selbst ein vollständiges Inventar an Tags für ihre Projekte aufstellen. Es besteht allerdings die Möglichkeit, grundlegend den TEI-XML-Dialekt zu verwenden und in der Praxis um eigene notwendige Elemente zu ergänzen. Dabei muss allerdings geregelt vorgegangen werden, um die intendierte Austauschbarkeit der Daten zu gewährleisten. (TEI P5 2020, S. xvi)

Die TEI bietet eine Reihe von Modulen, die XML-Elemente bestimmten Klassen zuweisen und ihre Attribute definieren. Manche fungieren als Basismodule bzw. werden für jede Art der Textkodierung empfohlen, während Elemente in anderen Modulen für die Kodierung ganz bestimmter Textsorten angelegt sind (z. B. poetische Verstexte). Für ein TEI-konformes XML-Dokument braucht es ein hinterlegtes TEI-Schema, gegen das validiert werden kann. Für dieses Schema können einzelne Module ausgewählt und auch Elemente entsprechend den eigenen Anforderungen adaptiert werden. (TEI P5 2020, Kap. 1)

Die Verwendung von TEI-XML bringt auch Vorteile für die Transformation von Dokumenten, da die TEI hierfür ein Set an XSLT-Stylesheets zur Verfügung

stellt. Damit können einfache TEI-Dokumente in z. B. *Word*, *LaTeX*, *ePub* und andere Formate konvertiert werden. Das frei zugängliche Konvertierungstool *OxGarage* ermöglicht zusätzlich die Transformation nach TEI für verschiedene Dokumenttypen wie Word, Spreadsheets oder Präsentationen. Ebenso gibt es online frei zugängliche Werkzeuge für die Publikation und Präsentation von TEI-Dokumenten im Web. Eine verhältnismäßig unkomplizierte Variante ist *TEI-Boilerplate*. Dabei wird ein TEI-Dokument mittels XSLT in einer HTML5-Shell eingebettet, die Gestaltung wird mittels CSS realisiert.

Es erscheint wenig überraschend, dass sich die TEI aufgrund der vorgeannten Vorteile als go-to-Lösung für Digitale Editionen etabliert hat. Die gemeinsame Basis gibt der DH-Community die Möglichkeit, sich über generelle Problemstellungen auszutauschen, Herangehensweisen weiterzuentwickeln und gegebenenfalls auch Daten nachzunutzen. Auf der Suche nach Vorbildern und Beispielen gibt es verschiedene Möglichkeiten zur Orientierung; die TEI führt z. B. selbst Projekte an, die TEI-XML verwenden.

Literatur:

- TEI: P5 Guidelines TEI Guidelines. URL: <http://www.tei-c.org/Guidelines/P5/>

Software:

OxGarage, TEI Boilerplate

Verweise:

XML, XSLT, Was ist XML-TEI?, Markup, Annotation, Modellierung

Projekte:

Text Encoding Initiative, Projects using the TEI

Themen:

Einführung, Annotation und Modellierung

Lexika

- Edlex: Editionslexikon
- Lexicon of Scholarly Editing

Zitervorschlag:

Eibinger, Julia. 2021. TEI (Text Encoding Initiative). In: KONDE Weißbuch. Hrsg. v. Helmut W. Klug unter Mitarbeit von Selina Galka und Elisabeth Steiner im HRSM Projekt "Kompetenznetzwerk Digitale Edition". URL: <https://gams.uni-graz.at/o:konde.178>