

Stand-off-Markup

Klug, Helmut W.; helmut.klug@uni-graz.at

Als *Stand-off*- oder externes Markup wird die Annotation von Primärdaten (Text- und Binärdateien) bezeichnet, wenn die Annotationen getrennt von den Primärdaten gespeichert werden. Die Idee dahinter ist es, zum einen binäre Dateien annotieren zu können, in die kein Markup eingefügt werden kann, zum anderen können parallel zur Textebene mehrere Annotationsebenen vorhanden sein, die sich als *Inline*-Markup in XML überlappen würden. Der Primärtext wird auf diese Art nicht oder nur geringfügig durch das eingefügte Markup verändert. Prinzipiell sollten Primärdaten und Annotation auch physisch (in unterschiedlichen Dateien) voneinander getrennt werden. Wenn alle Daten in einer Datei gespeichert werden, sind die Annotationen von den Primärdaten z. B. in der TEI durch die XML-Hierarchie getrennt: `<standOff>` als Geschwisterelement von `<text>`.

Von zentraler Bedeutung für diese Art der Annotation ist die Verknüpfung von Primärdaten mit den Annotationen. Das kann z. B. über *XInclude* und *XPointer* (vgl. TEI P5, Ch. 16.9f.) oder über entsprechende IDs, die vordefinierten Primärdateneinheiten (Indexierung auf Zeichen-, Wortebene, ...) zugewiesen sind, erfolgen. Das Referenzsystem muss jedenfalls stabil und idealerweise auf nachträgliche Änderungen in den Primärdaten ausgerichtet sein. *Stand-off*-Markup kann für unterschiedliche Anwendungsfälle geeignet sein, aktiv eingesetzt wird es z. B. im Rahmen der *Music Encoding Initiative* (MEI) für Codierung von Dynamik, Artikulation, Ornamentation u. dgl. oder in den Bibelwissenschaften als *Open Scripture Information Standard* (OSIS) und *Theological Markup Language* (ThML).

Stand-off-Markup wird seit den 1990ern diskutiert. Burghardt/Wolff fassen die Entwicklungen und die Überlegungen zu einer Standardisierung für literatur- und sprachwissenschaftliche Anwendungen zusammen und formulieren dazu grundlegende Empfehlungen (Burghardt/Wolff 2009, S. 57):

- Der Originaltext sollte in seinem ursprünglichen Zustand im Dateisystem der Annotationsdatei gesichert werden, um die Lesbarkeit und Wiederverwendbarkeit zu gewährleisten.
- Die Indexierung des Originaltextes sollte in einer gesonderten Datei gespeichert werden.
- Die Annotationssoftware, die das *Stand-off*-Format generiert, sollte Synchronisierungsmechanismen enthalten, die es erlauben[,] den Originaltext auch während des laufenden Annotationsprozesses zu ändern.
- Die Software sollte Versionskontrolle und Änderungshistorie der Primärdaten unterstützen.

- Bei der Indexierung sollte der Text am besten zeichenweise erfasst werden, da so später beliebig feinkörnige Annotationen hinzugefügt werden können.
- Die Speicherung von Originaltext und Annotation in unterschiedlichen Dateien erhöht die Lesbarkeit und ermöglicht die Konservierung der Primärdaten.

Literatur:

- Stand-off Markup. URL: <https://beethovens-werkstatt.de/glossary/stand-off-markup/>
- Burghardt, Manuel; Wolff, Christian: Stand off-Annotation für Textdokumente: Vom Konzept zur Implementierung (zur Standardisierung?). In: Von der Form zur Bedeutung: Texte automatisch verarbeiten / From Form to Meaning: Processing Texts Automatically. Narr: 2009.
- Spadini, Elena; Turska, Magdalena: XML-TEI Stand-off Markup: One Step Beyond. In: Digital Philology: A Journal of Medieval Cultures 8: 2019, S. 225–239.
- 16 Linking, Segmentation, and Alignment. URL: <https://tei-c.org/release/doc/tei-p5-doc/en/html/SA.html>
- Vigilante, Raffaele: Why TEI Stand-off Markup Authoring Needs Simplification. In: JTEI 2016–2019.

Verweise:

Datenmodell "MHDBDB", Textformate: GrAF, Annotation (grundsätzlich), Digitale Musikedition, Markup

Software:

XPointer - Test implementation, GATE for ANC

Projekte:

Music Encoding Initiative (MEI) , XML Inclusions (XInclude), XPointer Framework, Open Scripture Information Standard (OSIS), Theological Markup Language (ThML)

Themen:

Annotation und Modellierung

Lexika

- Lexicon of Scholarly Editing

Zitiervorschlag:

Klug, Helmut W.. 2021. Stand-off-Markup. 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.171>