

Die Odyssee zum richtigen Standard - Herausforderungen einer konsistenten Datenmigration von Ulysses: A Critical and Synoptic Edition (1984)

Schäuble, Joshua

joshua.schaeuble@uni-passau.de
Universität Passau, Deutschland; Reichsuniversität
Groningen, Niederlande

Crowley, Ronan

crowle01@gw.uni-passau.de
Universität Passau, Deutschland

Mit „Ulysses: A Critical and Synoptic Edition“ erschien 1984 eine der ersten Forschungseditionen, die auf Basis der systematischen Verwendung von Kollationsierungssoftware digital erzeugt wurde. Das Münchner Team um Hans Walter Gabler verwendete hierzu TUSTEP sowohl zur Validierung der Transkripte einzelner Zeugen als auch zur Erschließung der zeugenübergreifenden Synopse. Für die gedruckte Edition wurden die halbautomatisch erzeugten Kollationsergebnisse mit einem eigens entwickelten System komplexer Diakritika ausgezeichnet, die es dem geübten Leser ermöglichen sollten, die Textentstehung über stellenweise mehr als zwanzig inter- und intradokumentarische Textstufen hinweg in einer synoptisch integrierten Textfassung nachzuvollziehen. Während die Konzeption und Umsetzung dieser Arbeit bis heute als bahnbrechend im Bereich der Computerphilologie zu bezeichnen ist, konnte das Potenzial der resultierenden Druckausgabe für die Joyce-Forschung nicht annähernd ausgeschöpft werden. Zu komplex war das Markup, dem es gelingen sollte, zu verknüpfen, was zuvor getrennt war und zu hoch war der Aufwand, sich in diese Systematik einzuarbeiten.

Im Digitalen hingegen führten die Daten jene Odyssee fort, die die Druckedition beenden sollte. Auf der Suche nach einem Markup-Standard, der es vermag, die Inhalte der Druckedition digital zu repräsentieren, wurden die TUSTEP Ergebnisse zunächst von Tobias Rischer im Rahmen seiner Diplomarbeit (1997) in SGML/TEI transformiert und anschließend in mehreren Überarbeitungen über TEI P4 bis hin zur aktuellen Version der TEI P5v3 (2016) migriert. Dieser Beitrag vollzieht die Evolution dieser „Legacy Data“ nach, bis hin zu ihrer jüngsten Station - der noch andauernden Bemühung einer Migration nach TEI P5v3, welche im Rahmen des DFG-

und NEH-geförderten Kooperationsprojektes „Diachronic Markup and Presentation Practices for Text Editions in Digital Research Environments“ am Lehrstuhl für Digital Humanities der Universität Passau durchgeführt wird.

Erstmals seit der zweiten, überarbeiteten Ausgabe der synoptisch-kritischen Gabler Edition 1986 gelang es, aus den TEI-Daten die synoptische Visualisierung der Druckedition zu rekonstruieren und somit eine Konsistenzprüfung gegen die ursprünglichen Daten zu ermöglichen. Erst durch diese visuelle Rückführung offenbarten sich migrationsbedingte Fehler und Provisorien, welche zuvor, wenn überhaupt, nur in Fußnoten und privaten Aufzeichnungen vergangener Beteiligter dokumentiert wurden. Neben dem allgemeinen Versuch, die vollzogenen Änderungen aus den Aufzeichnungen und Migrationsergebnissen früherer Projekte zu rekonstruieren, hat es sich das Passauer Team zur Aufgabe gemacht, Strategien zur Entdeckung, Typisierung und Korrektur derartiger „Migrationsverluste“ zu entwickeln. Ein wesentlicher Bestandteil dieser Arbeit ist die Abschätzung der Leistungsfähigkeit und Wirtschaftlichkeit von automatisierten Batch-Konvertierungen mittels XSLT und Python im Vergleich zur manuellen Intervention und Korrektur der Kodierung.

Neben der Identifikation und Korrektur von „Migrationsfehlern“, steht die Rekonstruktion der textgenetischen Perspektive, durch welche sich die Druckedition auszeichnete, im Vordergrund. Während Gabler die textuelle Entwicklung, welche er mittels der Kollation chronologisch aufeinander folgender Textzeugen erschlossen hatte, im Druck synoptisch darstellen konnte, beinhalteten die TEI Guidelines bis zur Version P5v2 kein Modell zur Auszeichnung textgenetischer Prozesse. Es fehlte schlicht die Möglichkeit zur formalisierten Dokumentation einer stufenweisen, zeugenübergreifenden Chronologie der Textentwicklung. In der Druckedition wurde jeder auktorialen Textänderung *genau eine* Textstufe aus der heuristisch erschlossenen Chronologie zugeordnet. Diese lineare Textentwicklung über intra- und interdokumentarische Textstufen, in Gablers Terminologie auch Overlay und Level genannt, musste im Digitalen in eine Auszeichnung überführt werden, welche die Genese in den Hintergrund rückt und zu jeder auktorialen Modifikation anstelle einer Textstufe eine Liste sämtlicher Zeugen verzeichnet, auf welcher die spezifische Änderung Bestand hat. Diese Art der dokumentenorientierten Kodierung von Textgenese entspricht zwar bis heute der gängigen Auszeichnungspraxis historisch-kritischer Editionsprojekte, repräsentierte aber zu keinem Zeitpunkt die textgenetische Intension der 84er *Ulysses* Edition. Erst mit der Integration eines textgenetischen Modells in die TEI Guidelines, kann die ursprüngliche Intension erstmals auch in TEI kodiert werden. Hierzu bedarf es einer weiteren Episode der Datenmigration auf der Odyssee zum richtigen Standard.

Bibliographie

Bru#ning, Gerrit / Henzel, Katrin / Pravida, Dietmar (2014): „Multiple Encoding in Genetic Editions: The Case of Faust“, in: *Journal of the Text Encoding Initiative* 4. Available from: jtei.revues.org.

Burnard, Lou / O'Brien O'Keeffe, Katherine / Unsworth, John (2006): *Electronic Textual Editing*. New York: Modern Language Association of America.

Burnard, Lou / Jannidis, Fotis / Pierazzo, Elena / Midell, Gregor / Rehbein, Malte (2010): „An Encoding Model for Genetic Editions“, in: *TEI: Text Encoding Initiative*. Retrieved from .

Joyce, James / Gabler, Hans Walter (eds.) (1984): *Ulysses: A Critical and Synoptic Edition*. New York: Garland.

Joyce, James (1922): *Ulysses*. Paris: Shakespeare and Company.

Fordham, Finn (2010): *I do, I undo, I redo: The Textual Genesis of Modernist Selves in Hopkins, Yeats, Conrad, Forster, Joyce, and Woolf*. Oxford / New York: Oxford University Press.

Rischer, Tobias (1997): *Eine TEI/SGML-Edition der textkritischen Ausgabe von James Joyces Ulysses*. Diplomarbeit, LMU München.

TEI Consortium (eds.) (2016): *TEI P5: Guidelines for Electronic Text Encoding and Interchange. P5v3*. Available from: <http://www.tei-c.org/Guidelines/P5/> .