

Critical editing with TXSTEP

Wilhelm Ott, University of Tübingen

In the "Afterword" to his 1984 edition of James Joyce's "Ulysses", Hans Walter Gabler gives a short outline of how he collected the variant readings contained in the different sources, how he used them for establishing, in two steps, the critical text, leaving most of the mechanical work (even the automatic insertion of the diacritic marks for the genetic variants) to the computer, and checking, by subsequent machine collation, the manual work which was carried out interactively at the computer console. According to Gabler, "the systematic and comprehensive reliance on computer aid ... has drastically reformed the editing process... Without it, this edition would neither be as accurate as we hope it is ... nor so rich in recorded facts" (p. 1909).

The TUSTEP tools Gabler had used more than 30 years ago have constantly been adapted in close collaboration with many editorial projects to their respective requirements and to changing technologies like PostScript and PDF or HTML for output or encoding standards like SGML, XML, TEI and Unicode. They have successfully been used for the preparation of many other critical editions; www.tustep.uni-tuebingen.de/ed3.html lists more than 800 volumes of printed editions published between 1972 and 2014 prepared and/or typeset with TUSTEP, the Tuebingen System of Text Processing tools. They include works written in languages using non-latin alphabets, like greek (e.g. the 28th edition of Nestle-Aland, Novum Testamentum Graece, published in 2012), hebrew (e.g. the Mishna edition published by Michael Krupp) and arabic (Kitāb al-A'dād..., ed. 2012 by Gunhild Graf). Current editorial projects relying on these tools include the complete works of Marx-Engels (<http://www.bbaw.de/forschung/mega/>), the letters of Philipp Melanchthon (<http://www.haw.uni-heidelberg.de/forschung/forschungsstellen/melanchthon/edition.de.html>), the works of Christoph Martin Wieland (<http://www.wieland-edition.uni-jena.de/>), of Albertus Magnus ([institut.albertus-magnus-web.de/643_0/editiocoloniensis.html](http://www.institut.albertus-magnus-web.de/643_0/editiocoloniensis.html)), the philosophical works of Gottfried Wilhelm Leibniz (<http://www.uni-muenster.de/Leibniz/seite2.html>) and many others.

The TEI wiki judges the use of TUSTEP for the preparation of critical editions as follows:

- Advantage: does the job
- Drawback: very difficult to learn.

(see http://wiki.tei-c.org/index.php/Publishing_printed_critical_editions_from_TEI#using_TUSTEP)

According to Willard McCarty (Humanities Computing 2005, p. 217), main reasons for these difficulties are the language of documentation and the complexity of the interface.

This mentioned drawback has in the meantime lost much of its impact:

It was in 2010 when, following a proposal of Tobias Ott from pagina GmbH Tübingen, participants from seven universities and two commercial firms met at Stuttgart Media University to discuss the concept of a modern XML-based interface to these tools. Since then, a prototype called TXSTEP has been developed which provides such an interface to most of TUSTEP's program modules. It both

removes the language barrier and provides an user interface which an up-to-date established syntax. It allows the user to take advantage of the typical benefits of working with an XML editor, like content completion, highlighting, showing annotations, and verifying the code. The underlying XML schema contains extensive annotations and documentation on the purpose and syntax of the single functional elements available for building a TXSTEP script. When using a modern XML editor like oXygen, these annotations are shown automatically in a popup window while developing a TXSTEP script, so offering to a considerable degree a self teaching environment.

The paper will demonstrate the power offered by the current version on the example of the support for important steps for the preparation of a critical edition:

- Collating witnesses / collecting variant readings
- Evaluating the collation results
- Constitution of copy text
- Compilation of apparatuses
- Preparation of indexes
- Preparation of printer's copy
- Publishing the text with appartus(ses) in print and/or for the web

Both TXSTEP and TUSTEP are open source under the Revised BSD License and can be downloaded from the TUSTEP homepage. The TXSTEP installation package contains in addition a set of 90 sample scripts for simple and more advanced problems, covering tasks like file transformation (e.g. transforming TUSTEP files to XML), extraction of information, collation of different versions of the same text, evaluation of collation results, index generation and sorting.