

AAC: Austrian Academy Corpus

“Just in Time” Markup

Stephen Miller

Texts & Files: Using Corpora and Digital
Archives / Editions

Institut für Germanistik, Universität Wien

SS 2009

Authenticated Electronic Editions Project

- Based at the Australian Scholarly Editions Centre (ASEC)
- Housed at the University College, Australian Defence Force Academy

AAC: Austrian Academy Corpus

“The Problem with E-Texts” (2005)

The Problem of the Electronic Scholarly Edition (“The 1990s Vision”)

- “Nevertheless, they were still stuck with the fact that what they had provided users was essentially something that the physical book could have provided anyway: facsimiles and other images, transcriptions and ancillary information. Its digital form lent it new power, it is true, but if you were not a linguist or a computational stylistician then the word search and word frequencies tools helped you very little indeed. You as user still stood essentially outside the archive, just as you had done with the book. This was necessary if the ongoing accuracy of these electronic editions was to be guaranteed.”

The William Blake Archive

- The William Blake Archive
 - Seen (or will be seen...) as *first generation* electronic scholarly edition.

Second Generation Electronic Scholarly Edition

- New paradigm:
 - Authentication
 - Collaborative Interpretation

Authentication

- A guarantee of “textual integrity” of the transcriptions
- Separation of analytical/structural markup from interpretative markup
- Use of checksum values to guarantee data integrity

Collaborative Interpretation

- “By splitting off text from interpretation of it (this is what markup essentially is) we had laid the groundwork for collaborative interpretation of the text that could be recorded separately from it and would not endanger its accuracy.”
- “Could this be one building block in the new era of digital humanities research that is constantly being envisioned but never actually achieved?”

JITM: Just In Time Markup

AAC: Austrian Academy Corpus

**“Are Electronic Editions
Inherently Obsolete?” (2001)**

The Problem with Editions

- “The use of the word “Edition” when dealing with electronic texts is a carryover from the printing paradigm. An edition is defined by the act of publication, at which time it becomes static and part of the written record and cannot be easily corrected or extended. In the case where the scholarly work of an edition is to be extended, this normally involves the publication of a new edition and any necessary corrections are normally incorporated into the new edition.”

“Work in Progress”

- The Edition is now a “Work in Progress” as layers of markup are added to it depending on the interests of scholars
 - Part of this “progress” can be the digital curation of the edition itself
- Keeping an edition in use maintains interest in the asset itself and ensures its digital survival

“Virtual Text” & “Perspective”

- Resulting from philosophy of JITM:
 - Base text + layer(s) of markup leads to a new “Virtual Text”
- This Virtual Text is a “Perspective” of the Text

“Multiple Perspectives”

- Any number of different perspectives are possible and as new layers of markup are made available the number of different perspectives increases factorially

From Edition to “Study”

- From *static* edition to “the idea of a continually evolving electronic work”:
“The term I prefer is, “Study”, where the transcriptions and markup constitute a “study” of the work. I think this term gives the feeling of on-going scholarship, in keeping with an evolving scholarly effort.”

Marcus Clarke: *His Natural Life*

- “A JITM Worksite”
 - http://web.srv.adfa.edu.au/JITM/HNL/Annotation_Viewerr.html

Marcus Clarke: His Natural Life

- The Just In Time Markup (JITM) system generates different customised views ‘Perspectives’ of a transcribed text. As each Perspective is created the system continually authenticates the underlying transcription file. The default Perspective for this worksite corresponds to that as presented in the Academy Edition of *His Natural Life* created by the Australian Scholarly Editions Centre.

Marcus Clarke: His Natural Life

- The user can select a Perspective from among various options. Alternatively, the default Perspective presents the editor's base transcription with page and line breaks and highlighting (italics, underline, superscript etc.) applied to it. The user can remove these if required.

Marcus Clarke: *His Natural Life*

- This JITM worksite is dynamic. Users can contribute new interpretations: email acadedns@adfa.edu.au.