

The Texting Wilde Project: Thoughts on Tools for a Computer-Assisted Exegesis of a Biographical Corpus

Jason Alexander Boyd

Ryerson University, Canada; jason.boyd@ryerson.ca

The Texting Wilde Project (TWP) is developing computer-assisted methods for the analysis of biographical texts, particularly large corpora of these texts, using TEI. The corpus (currently amounting to just under 1000 items) the TWP will use to develop and test these methods is comprised of early (pre-1945) texts relating to Oscar Wilde (1854-1900), the late nineteenth-century Anglo-Irish writer, wit, personality and convict. Our knowledge of 'Oscar Wilde' is not comprised of a corpus of pure and simple facts that allows us an unmediated apprehension of a real person separated from us only by time, but rather this knowledge is comprised of a densely complex and often contradictory accretion of texts created in that intervening time that constitutes the only Wilde we can know. Many of the 'facts' of Wilde's life originated as narrative elements within particular textual and contextual moments and have acquired factive status through their repetition without attribution or with misattribution in text after text, losing in the process of this continual redeployment their textual provenance, a situation, given the volume of material to be mastered, that has frustrated what might be characterized as an exegetical analysis of Wilde biography. This is what the TWP aims to undertake.

The TWP is encoding texts not primarily as online reading editions for humans, but as texts that can be read (processed) by a computer to facilitate analyses of particular features within and across texts of particular relevance for the study of biographical writing. This is where the TWP faces its most significant challenge, since there are few existing, out-of-the-box or plug-in tools to enable the effective use of the analytic markup it is implementing. This paper will describe and propose ideas for tools and visualizations for the TEI markup being used for the TWP. These include tools for: the identification and tracking of the redeployments of stories concerning Wilde across texts; the extraction and comparison of reports of Wilde's sayings and conversations; the identification and use of sources (including whether accounts of Wilde are presented as first or second hand, attributed or unattributed); the generation of 'messy' timelines to compare disparities between when an event happened and when it was first recorded and to compare linear time vs. narrative order in biography (including dealing with impreciseness in dating); visualizing inclusions and omissions across texts (what periods of Wilde life are focused on/neglected); studying the interplay between "factual" narrative and interpretation; and mapping the emergence of primary documents in the print record through quotation, transcription, and reproduction.