TM4DH (Topic Maps for Digital Humanities): Examples and an Open Source Toolkit

John Walsh (*jawalsh@indiana.edu*)
Indiana University

My paper will discuss the use of Topic Maps to explore literary topics and navigate TEI-encoded literary texts. The presentation will include a brief introduction to Topic Maps and the XML Topic Map (XTM) syntax; an argument for the benefits of Topic Maps and similar metadata and ontology formats in electronic text and textual analysis applications; examples of fully-developed topic maps describing poetic genres and verse forms as well as the life and work of an individual author, Victorian poet Algernon Charles Swinburne. The presentation will conclude with a demonstration an open source Topic Map toolkit, developed by the author and consisting of a Java-based web application and XSLT stylesheets for the presentation and navigation of XML-based Topic Maps.

Topics Maps have been the subject of recent attention in the digital humanities community. They are discussed briefly in John Bradley's "A Model for Text Analysis Tools" in a recent issue of *Literary and Linguistic Computing*. And at Digital Resources for the Humanities (DRH) 2003 and the 2003 TEI Members Meeting, Stuart Brown gave presentations on Topic Maps as a means for navigating the TEI Guidelines and comparing different sets of local TEI extensions. At DRH 2004, the author presented a paper on "Topic Maps and TEI-Encoded Literary Texts". The current paper will build on this previous work by exploring more fully developed and generally applicable topic maps in conjunction with the open source toolkit that can be used by others in the digital humanities community to deliver their own Topic Map applications.

Topic Maps are a powerful XML metadata format that may be used to create multi-dimensional indices and interfaces to humanities resources and TEI-encoded data. The basic building blocks of Topic Maps are topics. A topic may represent any subject or concept. A Topic Map about a poetry collection, for instance, may contain topics representing "poetry," "sonnet," "lyric," "poet," "Wordsworth, William," and "The Solitary Reaper." Topics may have multiple names and may be connected via associations. For instance, one poet may "influence" another, a poet may "author" a poem, or a poet may "mourn" a historical figure in an elegy. Topics may also be typed and may have multiple types. So "Hamlet" may be an instance of the topic "play." Or "Rossetti, Dante Gabriel" may be an instance of both the "poet" and "painter" topics. Topic Maps allow one to build lists of concepts important to a collection or to a particular area of research and link those concepts to electronic resources, including TEI documents and individual elements within TEI documents. Topic Maps are powerful research and pedagogical tools that facilitate the organization, presentation, navigation, and visualization of conceptual and factual data.

To illustrate the discussion, I will use Topic Maps and TEI texts being developed as part of the *The Swinburne Project*, a digital collection of works by Victorian poet Algernon Charles Swinburne. Swinburne was an important cultural figure whose impact was felt beyond the domains of literature and poetry. He is an ideal central figure for the study of a wide range of nineteenth-century cultural and historical topics. Swinburne was an incredibly learned poet, and his range of form and allusion is extensive. His works include numerous and often obscure allusions to the bible, classical mythology, and Arthurian legend. He wrote a number of political poems addressing contemporary events. He wrote parodies of other contemporary poets, including Tennyson, Browning, and Rossetti. As Jerome McGann has noted, "[n]o English poet has composed more elegies than Swinburne." To address the breadth and range of form and allusion in Swinburne's work, the Swinburne Topic Maps include extensive lists of genre forms, people, biblical figures, mythological figures, Arthurian figures, as well as events and works by other artists and poets. These lists may then be used to build elaborate indices, navigation mechanisms, and data visualizations. The combination of TEI-encoded texts and XML Topic Maps allows the construction of a complex database of nineteenth-century British culture with Swinburne at its center. Supplementing the Swinburne-specific Topic Maps will be a more generally applicable Topic Map on poetic genres, meters, and verse forms. Swinburne was a versatile versifier who wrote in a great variety of forms. This "genre map," begun as a description of forms used by Swinburne, and since extended to include additional forms, may be used in conjunction with other digital collections.

The TM4DH (Topic Maps for Digital Humanities) open source topic map toolkit consists of a configurable Java-based Web application and bundled XSLT stylesheets for the presentation and navigation of topic maps conforming to the XTM 1.0 specification (<http://www.topicmaps.org/xtm/1.0/>), which describes an XML grammar for interchanging Web-based topic maps. TM4DH generates a homepage for the topic map. This homepage includes general metadata for the topic map (author, title, date, etc.) and a listing of major topic categories. From this homepage users can navigate to pages for individual topics, including detailed information about the topic along with its occurrences (internal to the Topic Map or external on the Web) and associated topics.

Bibliography

[No source references provided. Eds.]