



Big Data and Medieval Studies: the Present and Future of Medieval Text Archives

Trinity College Dublin, 27-28 June 2017

The last thirty years have seen the production of numerous large archives of medieval English texts, including the Dictionary of Old English Corpus (c. 3 million words), the York-Toronto-Helsinki Parsed Corpus of Old English Prose (c. 1.4 million words), the Manchester Eleventh Century Spellings Database (c. 300,000 words), the Linguistic Atlas of Early Middle English (c. 650,000 words) and the Corpus of Middle English Prose and Verse (c. 5 million words). Since each of these freestanding corpora was built for a different purpose, there is minimal interoperability, and the user must learn separate user interfaces and search protocols for each. Their extraordinary collective power as a tool for cultural, historical, literary and linguistic analyses thus remains to be exploited. Early publications using the materials produced by the Early English Books Online Text Creation Partnership (EEBO-TCP) have shown the revolutionary power of big data to reconfigure our understanding of the early modern, print past. This colloquium seeks to catalyse a similarly radical transformation in the possible methodologies for the study of the medieval period, by encouraging collaboration to increase the use and utility of existing text archives and setting a blueprint for their future development.

The colloquium will feature presentations from all the major text corpora of medieval English. A small number of places have been reserved for other contributors, and abstracts are now sought for 15-20 minute papers describing methodologically innovative, current research using these or other medieval text archives. Types of research particularly relevant to the aims of the colloquium include:

- Research that spans multiple corpora that are non-congruent (e. g. parsed and unparsed corpora, manuscript-focused and text-focused corpora, corpora of texts in different languages)
- The use of text archives for purposes beyond which they were designed
- The use of text archives to address broader cultural, literary or historical research questions

Abstracts of 250 words should be submitted to Mark Faulkner (faulknem@tcd.ie) by 26 February 2017.



Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin



IRISH RESEARCH COUNCIL
An Chomhairle um Thaighde in Éirinn

Trinity Long Room Hub
Arts & Humanities Research Institute