Scalable Analysis of Printers of Early Modern Britain by a Semi-automated Curation Process

Iiro Tiihonen. University of Helsinki
Ville Vaara. University of Helsinki
Leo Lahti. University of Turku
Mikko Tolonen. University of Helsinki

The business activity that revolved around books during the early modern period is a widely researched topic. Despite earlier qualitative research in this area, quantitative attempts to analyze printing in the early modern period have been few and limited in scope due to the large amount of tedious manual work that is required for such analysis. Large-scale analysis of bibliographic catalogues is providing new ways to scale up such research.

In order to scale up the possibilities of quantitative research, we introduce and evaluate an analysis process that aims to remarkably reduce the amount of time required in the quantitative research of early modern book trade. The process is based on a semi-automatic analysis of catalogue records, and it is designed to be scalable. Our pilot evaluation is made by re-producing an earlier data set of Gants[1] by using the English Short Title Catalogue (ESTC), and comparing our results against his, taking the latter as a gold standard.

Despite the inherent noise in our semi-automated approach, we can demonstrate a successful replication of the key variables. We were especially successful at replicating the physical estimates, such as the total number of editions produced by a printing house, its total output in edition sheets and the average size of a work produced by a printing house in edition sheets. The average error for these variables was less than 20 percent. These first validations show that the method is amenable for further extension and application, thus providing a new research method in book history.

[1] A Quantitative Analysis of the London Book Trade. Studies in Bibliography, Volume 55. pp. 185-213