Counting the Bees: A Data-Driven Investigation Into Early Modern British Thought (1605-1776)

Keywords: British philosophy; computational method; axiomatic method

Prevalent narratives in the history of philosophy suggest that the advent of British experimentalism brought about a fundamental change of scientific ideals. The traditional axiomatic ideal of science, deriving from Aristotle and influential throughout the medieval age and the Renaissance, was replaced by new ideals. These new ideals in varying ways replaced strong epistemological and metaphysical principles with weaker ones: probability over certainty, contingency over necessity, instrumentality over truth, and the particular over the universal (e.g. Shapin & Schaffer (1985), Dear (1995), Pasnau (2019)). While such narratives are partially accurate, they remain incomplete for two key reasons.

First, their authors produce generalizations about the nature of science and philosophy in Early Modern Britain on the basis of a limited evidence base, consisting of canonical works by the likes of Locke and Newton, occasionally supplemented with microstudies focusing on a few lesser-celebrated thinkers. However, wide-scope historical claims can only be known to be accurate if they are grounded in a comprehensive collection of sources. Pending histories based on such a collection, large claims about the fate of the axiomatic ideal remain uncertain.

Second, the emphasis on the shifting aspects of the axiomatic ideal obscures potential continuity in other respects. As a result, our understanding of how the composition and hierarchy of scientific concepts, or the axiomatic ordering of propositions, were (re)conceptualized by thinkers outside of the canon is fragmented at best. New discoveries in these respects may expose a greater continuity of scientific ideals than current narratives suggest.

This ongoing research applies a novel methodology from the emerging field of data-driven history of ideas to tackle both issues by systematically retrieving and analyzing evidence concerning the axiomatic ideal from a comprehensive corpus of texts built for this study. Besides being data-driven, the study is also guided by a conceptual model called 'The Classical Model of Science' (CMS) (de Jong & Betti 2010), which codifies the principles of a proper axiomatic science.

The study commences by building a digital bibliographical knowledge graph, called BOOKSHELPhS (Books in the History of English Logic, Philosophy, and Science), that contains as many bibliographical records as possible of logic, philosophy, and science books in English or Latin, published in Britain between 1605 and 1776 and written by authors alive during that period. BOOKSHELPhS is an ongoing project, currently consists of rich metadata of 2.123 editions of 1.272 works, and is available in a Linked Open Data format. All BOOKSHELPhS editions have been painstakingly vetted for inclusion.

BOOKSHELPhS forms the bibliographical backbone for the upcoming phase of the study, which consists in the construction of a high-quality, richly annotated and structured, machine-readable text corpus of BOOKSHELPhS editions. The aim of this corpus is the computational retrieval of paragraphs containing discussions of concepts in CMS on a scale that traditional, manual searching cannot practically match. In the absence of a proper evaluation of language models on philosophy corpora, retrieval will rely on sophisticated string search methods. More precisely, the evidence will be retrieved by first constructing an extensive mapping between CMS concepts and clusters of actor's terms (e.g. axiom to

axiome, *postulatum*, *maxime*, etc.), and then by querying the corpus for paragraphs where these terms (co)occur.

The collected paragraphs are subsequently close-read, annotated using a detailed annotation scheme, and scored for relevance and adherence to conditions in CMS. The result of this process is a dataset of evidence on the topic which, when linked to BOOKSHELPhS metadata, can be used to analyse the Early Modern British philosophical and scientific landscape on an unprecedented scale and in a rich variety of ways. For example, linking the annotations to BOOKSHELPhS metadata facilitates temporal analyses of concept drift, the identification of geographical patterns such as the spread of certain takes on science in specific regions, and publishing trends.

While the primary goal of this study is to expand our understanding of Early Modern axiomatics, the resources produced are not restricted to this aim and hence support a broad range of investigations in Early Modern British philosophy, logic, and science.