Greek in the Shell:
Engineering computational methods for Greek textual criticism

James Brusuelas, University of Oxford

The searchable databases of Classical texts continue to increase and change the way scholars study the languages and literatures of ancient Greece and Rome. However, the functionality of most digital representations of texts is largely dependent upon human guidance. An actual interface between human and artificial intelligence within the plane of Classical methodology is still in its infancy. Specifically, research into how Classical Philological method can merge with computational analysis to create innovative ways of performing textual criticism and improve the reading of ancient manuscripts is wanting. Are we able to successfully build a computational interface that will make use of human and machine intelligence to reconstruct a meaningful, scholarly and potentially publishable text?

The 'Ancient Lives' project has amassed an immense database of crowd-sourced transcribed Greek papyri in order to assist in the classification, cataloguing, and identification of the tens of thousands of unstudied papyrus fragments housed in the Sackler Library of the Ashmolean Museum. Consequently, a database of unedited Greek texts exists for the first time. Thus the aim of this project is to build a computational interface for the critical editing of Greek texts, one initially based on that database. Two of the principal goals are: (1) to refine and implement a consensus algorithm, adapted from those used in the study of DNA sequencing, that collates multiple transcriptions, producing a variety of meaningful statistical data and digitally searchable transcriptions that will be accessible to scholars through a unique graphical user interface (GUI), (2) algorithmic meaning extraction for the purpose of computationally repairing gaps/holes in papyri through an automated projection of characters that are not only dimensionally suitable, but also contextually sound through linguistic parallels elicited from the searchable online databases of Greek texts. Furthermore, based on the extant strings of Greek characters on a given papyrus, automated projection of possible linguistic/word scenarios will assist in the contextualization and identification of unknown literary works.

The purpose of this paper is to display initial project results, to outline in greater detail our methodologies, and, lastly, to introduce a progressive path towards a digital edition that reshapes the critical and philological reading process.