Digital Index Apuleianus

## Description and Impact

The *Index Apuleianus* by William Abbott Oldfather et. al. was published in 1934 by the American Philological Association ([review](https://www.journals.uchicago.edu/doi/abs/10.1086/361903)). With the permission of the SCS as copyright holder, we plan to digitize this work and, using a program developed by Michael Skalak, convert it into a fully lemmatized text and database. Further processing will tie the data to the set of Latin lemmas used in the Morpheus code, which is used by Perseus and other datasets of parsed Latin. We will use this data in [The Bridge](http://bridge.haverford.edu/), an application for the creation of correct vocabulary lists for Greek and Latin authors. This will **allow students and teachers to create accurate custom vocabulary lists for all works Apuleius, and substantially increase the readability of his texts**. (This is already the case for the *Aeneid*, Caesar’s *Gallic War*, et al., thanks to the Bridge.) Another benefit for scholars will be ready access to the frequency and morphology information contained in the *Index*, which is currently hard to access. The scholarly effort that went into the creation of the *Index* can thus be liberated for re-use in at least three ways:

* Vocabulary lists for easier reading (as in Dickinson College Commentaries editions)
* Data for word frequency and stylistic analyses
* The development and better training of automatic parsers based on a larger corpus of hand parsed Latin

Indirectly and eventually, this project will make automatic parsers better so students of all Latin authors, not just Apuleius, can read them more easily. We imagine there will be other potential uses for this data we have not thought of. We plan to release the data on Github under a creative commons share alike license, in at least three formats:

* a .txt file of the professionally digitized book
* “unscrambled” in a spreadsheet that constitutes a lemmatized text of the works of Apuleius, each text form associated with the correct Morpheus lemma
* In a spreadsheet with dictionary forms and definitions provided by The Bridge, in a way that will constitute a full lexicon of Apuleius

## Costs

Professional digitization of the book by NewGen Software will cost approximately $800. We would need approximately $500 for post-processing and checking. This job would go to a qualified scholar, such as a graduate student or non-traditional classicist, who would be able to perform this exacting work on a contracted hourly basis.

## Other Potential Sources of Funding

This is part of a larger project to digitize classical concordances. We have applied for research funding from Haverford and Dickinson Colleges to support the larger project.

## CVs attached

## Professional Reference:

Helma Dik, Associate Professor in the Department of Classics, University of Chicago, 1115 E. 58th Street, Chicago, IL 60637, (773) 834-2170, helmadik@uchicago.edu