



ERC RUTTER project - Internship program 2020

Arabic Text from Manuscript to Contemporary Digital Formats

Supervisor: Juan Acevedo

Profile: The prospective student will have basic Arabic reading skills or an interest in acquiring them, and a general interest in digital humanities.

This internship will:

- provide basic Arabic language skills or improve on them,
- equip the interns with philological skills,
- acquaint them with existing cutting edge technologies to codify Arabic,
- introduce them to the basic principles and techniques of digital humanities, including markup and text encoding formats and conventions.

Tutorials:

Two or three sessions of 90 minutes, depending on previous knowledge: one initial introductory session and one or two additional depending on the intern's interests and progress. Because this is meant to be a hands-on project, direct and frequent consultation will also be encouraged and expected.

Tasks:

- (if needed) Acquiring the basics of Arabic script.
- Familiarise with existing Arabic transliteration and encoding conventions.
- Familiarise with the basics of text encoding and semantic markup.
- Engage head-on in the ongoing transcription and semantic digitisation of an Arabic manuscript.

Basic reference material:

- ALA-LC Arabic romanization table at the Library of Congress site.
- Websites of W3C and TEI initiatives, especially general info on the Semantic Web and TEI-related projects.
- A list of RUTTER-relevant TEI-compliant projects to be browsed attentively.