
This workshop will be a practical introduction to a set of resources, methodologies and practices, centered around the digital annotation of texts, which in the last decades have come to enrich, and partially change, the field of classical studies (Crane & Terras 2009), especially – but not only – in the areas of papyrology (Depauw & Stolk 2015), epigraphy (Cayless et al. 2009), linguistics (Bamman 2014; Haug 2014) and textual criticism (Dué & Ebbott 2009). There will be two sessions, one on text encoding with EpiDoc and one on treebanking for ancient languages.

The workshop being an introduction, it will firstly be directed to scholars and students with little or no experience in using TEI/EpiDoc or with treebanking, but, of course, also to more advanced participants interested in the use of TEI encoding and treebanking with ancient languages.

Maximum number of participants: 20

Time and Location: 9.00-13.00, Monday 14.03.2016, "Læringsoasen" room, Humanities and Social Sciences Library, University Campus Blindern

Technical equipment: None. The room will be equipped with a desktop computer (with the necessary software already installed) for each participant.

Organizers: Humanities and Social Sciences Library (UiO), Department of Philosophy, Classics, History of Art and Ideas (UiO).

Contact: Federico Aurora, federico.aurora@ub.uio.no

Session 1: EpiDoc

EpiDoc\(^1\) is a set of guidelines for encoding the digital versions of ancient source texts in TEI (Text Encoding Initiative)\(^2\) – originally developed for Greek and Roman epigraphy, but now much more diverse – including a recommended schema and ODD (One Document Does it all) file,\(^3\) a lively community of practice, and an ecosystem of projects, tools and stylesheets for the interchange and exploitation of such texts (Bodard 2010).

The first workshop session will introduce participants to the principles and practices of EpiDoc encoding, which are largely based on the practice of encoding single-source documents and the ancient objects on which they are written. Special attention will be paid to how the community guidelines recommend (a) mapping Leiden Conventions for text transcription (i.e. the standard practice for paper editions, see Cayless et al. 2009: 9) to TEI transcription and critical apparatus elements, and (b) mapping the descriptive and historical metadata features – from projects such as EAGLE (Europeana network of Ancient Greek and Latin Epigraphy)\(^4\) and Papyri.info\(^5\) – to TEI "MSDesc" (manuscript description) elements.

---

1. [epidoc.sourceforge.net](http://epidoc.sourceforge.net)
5. [http://papyri.info](http://papyri.info)
Finally, tools and other methods made available by the community for transforming, publishing, querying, exchanging, and linking encoded materials will also be introduced.

Instructors:

Gabriel Bodard, reader in digital classics at the Institute of Classical Studies, University of London, has a background in classics, with training and experience in both papyrology and epigraphy. He worked for fourteen years in Digital Humanities at King's College London on various digital projects and acquired broad and in-depth knowledge of text encoding and processing, especially the use of TEI XML and XSLT. He is one of the lead authors of the EpiDoc Guidelines (Elliot, Bodard et al. 2007-2014) and regularly organises and teaches training workshops in digital epigraphy and papyrology.

Joanne Stolk, PhD fellow at the Department of Philosophy, Classics, History of Art and Ideas, University of Oslo, has written her PhD thesis on "The Language of the Papyri: Diachronic change in Egyptian Greek (300 B.C. - 800 A.D.)". She has broad experience with encoding papyrus texts in the Papyrological Editor of Papyri.info and has contributed to create a database on irregularities and mistakes in the Greek papyri (Trismegistos Text Irregularities).

Session 2: Treebanking

"A treebank is a large collection of sentences in which the syntactic relation for every word is made explicit – where a human has encoded an interpretation of the sentence in the form of a linguistic annotation." (Bamman 2015). Treebanking means, thus, annotating texts morpho-syntactically (and, more recently, also semantically and for information and discourse structure). This allows to approach the study of ancient languages, and especially of their more abstract structures, in more sophisticated ways, developing empirical analyses and testing theoretical assumptions against large amounts of data in a statistically relevant way. Diachronic treebanks, moreover, especially of languages for which we can count on corpora stretching over more than two thousands years, such as Greek and Latin (including also its Romance continuators), also offer an invaluable opportunity to perform quantitative studies of language change and evolution. And even though in its beginnings the creation of treebanks regarded mostly modern languages, the last years, which have seen a large growth of corpora for ancient languages, (some of them quite general and available online, others created ad hoc and for specific research projects) have also seen the creation of treebanks of historical languages, such as the The Ancient Greek and Latin Dependency Treebank (Bamman & Crane 2007; Bamman et al. 2009) at the Perseus Project and the PROIEL Treebank (Haug & Jøhndal 2008), currently the two largest treebanks of Ancient Greek and Latin.

The second workshop session will give a basic introduction to the theory, methodology and practice of treebanking seen in the light of the particular challenges posed to it by ancient languages (focusing especially on Greek and Latin). Its practical part will be a hands-on demonstration of the PROIEL treebank and its online application. The PROIEL treebank, originally developed for the PROIEL project (Pragmatic Resources in Ancient Indo-European Languages) to contain the original Greek text of the New Testament and, aligned, its translations into early Indo-European languages (Gothic, Armenian, Latin and Old Church Slavic), has later been extended with several other texts in Greek and Latin (e.g. Herodotus,

---

6 http://www.trismegistos.org/textirregularities/
7 http://www.perseus.tufts.edu/hopper/
8 http://foni.uio.no:3000
Caesar, Cicero, the *Peregrinatio Aegeriae*), but also in other historical languages, including the Old Norse *Edda* and many texts in old Romance languages.

**Instructor:**

*Dag Haug*, professor of Latin at the Department of Philosophy, Classics, History of Art and Ideas, University of Oslo, works in historical linguistics, syntax (Lexical Functional Grammar) and formal semantics, having his background in Homeric philology and comparative Indo-European linguistics. He led the PROIEL project and the construction of the PROIEL treebank.

**Person responsible for the workshop:**

*Federico Aurora*, subject specialist in classics and papyrus collection curator at University of Oslo Library and PhD candidate at the Department of Philosophy, Classics, History of Art and Ideas, University of Oslo, is writing a doctoral thesis on the case system of Mycenaean Greek and leads the construction of the annotated corpus of Mycenaean Greek (DAMOS - Database of Mycenaean at Oslo).  

The workshop will be followed in the afternoon by a lecture by Gabriel Bodard with the title: "A Digital Classics research agenda: both Digital Humanities and Classical Studies".

The lecture will be hosted by the Classical Seminar of the Department of Philosophy, Classics, History of Art and Ideas, and will be open to everybody.

**Time and place for the lecture:** Tuesday 14, 2016, 14:15-16:00, room 207, Georg Morgenstiernes building, Blindern University Campus.

---

3 https://www2.hf.uio.no/damos/
References


