Onomastics and Entity-Recognition in Graeco-Egyptian Papyri

London, June 5

Bart Van Beek
bart.vanbeek@arts.kuleuven.be
1. Project history

- 1950: Prosopographia Ptolemaica
- 1990’s: digitization
- 2004-2008: Trismegistos (Cologne-Leuven)
- 2008-2012: Onomastics in Graeco-Roman Egypt
2. Trismegistas Platform (TM)

- Metadata of all published texts from Egypt between 800 BCE and 800 CE
- Physical object information
- Content-related information
- Related database architecture
- Online platform: http://www.trismegistas.org/
2. Trismegistos Platform
3. Trismegistos and onomastics

- combining existing prosopography with text metadata
- expand scope of Prosopographia Ptolemaica
- add onomastic analysis

- Automatization of data entry?
- Duke Databank of Documentary Papyri (DDBDP)
- online: see Papyrological Navigator (http://papyri.info/)
4. Duke databank

- recently: major update
- beta-code > unicode
- xml-encoded
- EPIDOC-compliant
5. Onomastic processing

• Extract all names
• create list of unique name forms
• identify:
  – declined forms (NAMVARCASES)
  – (spelling) variants of a name (NAMVAR)
  – linked to a ‘standard form’ of a name (NAM)
6. Prosopographical processing

- Re-scan xml-files for names, but evaluate the result against set of rules based on onomastic filtering
- color-coded: green - red - purple
- create prosopography entries in TM
7. Future prospects

• combine metadata with full-text corpus?
• web-services interaction with other related projects
• integration example: Google Maps