Romans 1 by 1. Documenting a population database for the Roman world

Rada Varga[a],[*]

[a] *Babeș-Bolyai University, Cluj-Napoca*

Abstract

The present article briefly documents the epigraphic division of a developing online population database for the Roman Empire, accessible at www.romans1by1.com. The paper presents the motivation of constructing the database and its envisioned architecture, in relation with the various sources, while emphasizing on the steps and procedures required in order to transpose epigraphical information into an ancient population database.

Keywords: Epigraphy, middle classes, provincial society, occupational inscriptions, prosopography

Romans 1 by 1 database and afferent website were created for filling in an existing gap in the study of Roman-era population. The database tries to begin answering to the need of properly cataloguing all attested inhabitants of the Roman Empire. Of course, this is a tremendous task and www.romans1by1.com is only a first step, circumscribed for now to a very specific category of sources.

1. Motivation

But why would an ancient population database be essential? Because a digital resource focused on individuals would reveal linkage possibilities that otherwise elude us, it would finally give us the complete and accurate image of the Roman attested population and, through codifications, it would open the way towards computer-assisted in-depth analyses on all relevant aspects imagined (epigraphic patterns, religiosity, migrations, onomastics, occupations, family data, etc.). A complete,

---

[*] Corresponding author. Email: radavarga@gmail.com

aggregate database will allow a longitudinal (diachronic) view on the attested Roman population from a certain area and ideally from the whole Empire, while also opening transversal (a section in time) perspectives.

The database in itself would have three components (see below Fig. 1), following the best practices in the field (MANDEMAKERS AND DILLON (2004)): the *sources database* (with "facsimile" transcription of the sources' text), the *central database* (the complete, corrected, integrated, standardized and coded form of the sources database) and the *data releases* (destined for on-line usage, allowing easy extraction of data in view of analyses).

Alongside these components, which actually represent various steps in the data preparation process, a population database for the Roman Empire should be built on three pillars, which we might call units or divisions of the database: epigraphic, literary and archaeological, each of them requiring different expertise, different approaches and different standards for the individual recording forms. In the end, of course, all three types of individual records will have to be integrated in a central standardized component, whose structure is to be developed by merging the three aforementioned divisions; consequently, its configuration will ideally take shape only after all three units enjoy a stable architecture at least in the sources component of the database.

The construction of the database will follow a series of steps imposed by best practices: creating a repository of sources; introducing, integrating, standardizing, coding and storing the information; enriching and disseminating the information. The codifications hold an essential part, not only in the individual linkage procedures, but in the analysing process as well. At the point when the database will comprise enough data, properly recorded and with all codifications undertaken, the usage of statistical software in order to identify trends and run comparisons over large scale geographical and administrative units might result in a better than ever understanding of Rome's social history.

2. The sources

Within this theoretical frame, whose amplitude implies a gradual and long-term approach, we have chosen to start by building the epigraphical division of the database. Thus, we began developing a project on the middle classes of the Low Danube provinces and their epigraphic attestations. The database created in this context together with the platform
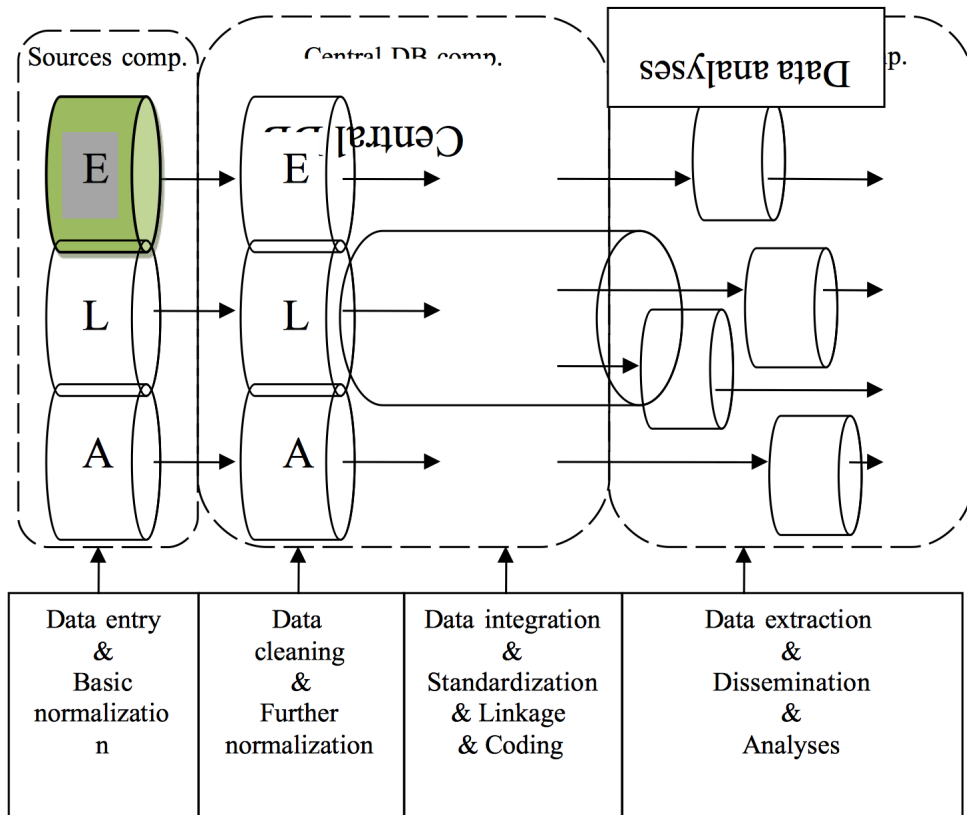
Fig. 1. Schema of the database

`www.romans1by1.com` represent the skeleton of the future population database of the Roman Empire.

For now, we are solely focusing on characters attested epigraphically – thus on inscriptions as sources. As constructing a metadata suitable for all social and professional categories of the provincial world is very complex, our database is created for accommodating all information epigraphically provided on members of the provincial middle classes. Terminologically, we consider all those who manifest themselves epigraphically, without being members of the imperial or provincial elite, as member of the middle socio-economical layers. Likewise, we have excluded active militaries (but not their families), as their appurtenance to the army creates their social status.

Regarding the data entry, we are trying to remain faithful to the source and to record, during the first phase, only the minimum of deduced information (gender, juridical status, ethnicity of the name). At the same time, we operate some conventional onomastic transcriptions (AEL will be Aelius from the start, etc.). All these basic normalization procedures are being thoroughly explained and documented in the data entry manual, and the reason why we chose to apply them is to speed up data standardization, run analyses and publish results on small samples, whose standardization procedures would otherwise be overwhelmingly time-consuming in relation with the results.

3. Database architecture

The core of the database's epigraphical division is formed by a table used for recording personal individual data (labelled *Personal data*), around which the entire network of components needed to ensure proper information recording is built. However, when starting data entry, one must begin by recording information about the source in use.

The first table to be filled in is the file of the source – *Inscriptions*. First of all, each inscription gets a source code, formed of 5 digits and a symbol/acronym of the province's name (D for Dacia, MS for Moesia Superior, DAL for Dalmatia, etc.) – so we have, for example 00001MS. The *Inscriptions* table has text fields, as well as value lists. Then, we have text strings for: *Relevant expressions*, *Stylistic details*, *Atypical features*, *Observations*, *Place of discovery*, *Place of provenience*, *Ancient name provenience*, *Timestamp/Timeframe* and *External links*. Although we are aware that some of the data (*Timestamp/Timeframe*) could have benefitted from

a standardized form, at this point we opted for vaster possibilities of expression and adaptation. Other information will be filled using value lists, as standardization is more suited: *Type of inscription*, *Language*, *Material*. For these we will use the Eagle vocabularies,[1] because they offer an already standardized language.

The *Coordinates* table links each inscription with the latitude and longitude of its place of provenience, in order to place it on a map.

The *Inscription bibliography* was conceived so that extracting complete or selective bibliographical lists would be possible. Thus, a normalization table includes all bibliographical titles referred to and from it; through a value list, one can select the *Bibliography abbreviation*. The exact reference is presented and detailed in the *Details* and *Comments* text strings. Of course, all data are linked to the *Inscription code*, selected as well from a value list.

Only after properly documenting the source one can access the main and most complex table, *Personal data*, where basically the individual file of each person is created. At this point, each new entry represents a singular epigraphic attestation of an individual, and a unique ID is generated, which will help link the respective character throughout the various components of the database and with other database entries. The person is also manually linked to the source using a value list of the inscriptions' codes. In the event of one person being attested in multiple epigraphs, each attestation will represent a new entry, and it will be assigned a new unique ID, which will be doubled, during linkage procedures, by a common ID for all instances of the same person.

The *Personal data* table was built to host a large variety of information offered by epigraphs, and its structure has proved, up to this day, rather stable. However, if need be, it can always be extended in order to accommodate any new kind of information. A first set of its variables are name-connected: *Praenomen*, *Nomen*, *Cognomen/Personal name*, *Father/Master* name, *Agnomen*, *Signum*, where each also has a drop-down associated for *Ethnicity* and the *Agnomen* and *Signum* an *Observations* string. While we believe the possibilities of actually identifying *signa* for members of the middle classes, during the Principate period, is rather reduced, we opted for facilitating their correct registration, in case they are discovered. Other data regard *Natione*, *Ethnicity*, *Origo* and *Domus* – if and how they are mentioned in the inscription. As acknowledged

---

[1] http://www.eagle-network.eu/resources/vocabularies

above, some information will be recorded, even if they are deductive and not literally written in the source: *Gender* and *Juridical status* (though the servile one often is literally recorded). For the latter, we have opted for a check box, which, if checked, opens all the available possibilities. The rest of the fields accommodate supplementary information, when and how there is the case: *Occupation*, *Collegium*, *Deities*, <u>Age</u> (at death), *Details of life/death* and *Observations*.

The *Occupation* field requires some special attention, as it has the *Occupation code* associated with it; as we are trying to propose a codification of Roman occupations/professions, based on and adapting the HISCO classification model[2]. While a raw classification and codification based on HISCO might only be a slight challenge, finding a theoretical model might prove to be a more elaborated task. What we aim at is constructing the "metadata" for the codification of Roman professions attested on stone and analyse how deep the classification can realistically go. Much alike other normalization fields present in the sources component of the database, *Occupation code* was implemented because it helps dealing faster with small and medium size samples (up to hundreds of people) in view of publishing preliminary results which are vital for dissemination and further financial support of the project.

Two important variables, for statistics and working with the data, are *Dedicated by* and *Dedicated for*, which state the relation of the recorded individual with the epigraph and with other persons mentioned by it. Both appear in the simple form of check boxes.

Another particular check box, in need of supplementary explanations, is Later. The database was conceived for attested civilians from the middle classes, but sometimes they are associated in inscriptions with militaries or representatives of the elites. In this case, we have to register a minimum of data on the later as well, in order to build a wider image of the characters that we focus on. When checking *Later*, it opens *Status*, which at its turn opens the following options in the form of check boxes: *Senator*, *Knight*, *Local magistrate*, *Decurion*, *Imperial priest*, *Imperial slave*, *Imperial freedman*, *Military personnel*. If checked, each of these boxes opens a Details text box; additionally, *Local magistrate* and *Decurion* open a *City/Town* value list, while *Military personnel*, opens *Rank* and *Unit* value lists. While elite members and military personnel will at some point be added to the database, for the currently running project purposes, their

---

[2] http://historyofwork.iisg.nl/

social and professional status, together with a minimum of relevant details and relation with the recorded individuals represent enough information.

Field label	Field id	Data type
Praenomen	rperson_praenomen	Text
Ethnicity praenomen	rperson_eth_praen	Value list
Nomen	rperson_nomen	Text
Ethnicity nomen	rperson_eth_nom	Value list
Cognomen/Personal name	rperson_cognomen	Text
Ethnicity cognomen/personal name	rperson_eth_cogn	Value list
Father/Master name	rperson_father	Text
Ethnicity father/master name	rperson_eth_fth	Value list
Agnomen	rperson_agnomen	Text
Ethnicity agnomen	rperson_eth_agn	Value list
Observations agnomen	rperson_remark_agn	Text
Signum	rperson_signum	Text
Ethnicity signum	rperson_eth_sign	Value list
Observations signum	rperson_remark_sign	Text
Natione	rperson_nation	Text
Ethnicity	rperson_nat_ethnicity	Text
Gender	rperson_rgender_id	Value list
Juridical status	rperson_jstatus	Checkbox. Opens checkboxes: Citizen, Libertus, Peregrine, Slave Veteranus (opens Unit/Rank - valuelists)
		Continued on next page

Field label	Field id	Data type
Tribus	rperson_rtribus_id	Text
Origo	rperson_origo	Text
Domus	rperson_domus	Text
Collegium	rperson_collegium	Text
Occupation	rperson_occupation	Text
Occupation code	rperson_occ_code	Text
Deities	rperson_deity	Text
Age	rperson_age	Text
Details of life/death	rperson_death_details	Text
Dedicated for	rperson_dedicated_for	Checkbox
Inscription code	rperson_rinscription_ids	Value list
Later	rperson_now_stat	Checkbox. Opens checkbox: Status
Status	rperson_status	Checkbox. Opens checkboxes: Senator (opens Details – text), Knight (opens Details – text) Local Magistrate (opens Details – text, City – value list), Decurion (opens Details – text, City – value list), Imperial priest (opens Details – text), Imperial slave (opens Details – text), Imperial freedman (opens Details – text), Military personnel (opens Details – text, Unit/Rank – value lists)
		Continued on next page

Field label	Field id	Data type
Observations	rperson_observation	Text

Tab. 1. Metadata of the Personal data table

Based on the personal ID given to each individual, the *Relationship* table will solely name the relationship between individuals (A to B and B to A), choosing from a drop-down menu. The relationships have been encoded from the start, in order to make processing quicker; thus first-degree relationships have 10-codes (101-Husband, 102-Wife), second and third degree relations 20- and respectively 30-codes, non-family relations were given 40- codes and 50-s for the unspecified/unreadable relations. Male relations were given odd numbers and female ones – even numbers.

4. Conclusions

Romans 1 by 1 is a first step towards a comprehensive and exhaustive electronic resource for the attested population of the Roman Empire. The following normal steps for expanding and enriching the database is elaborating the fitted metadata for provincial elites and military personnel epigraphically attested.

Acknowledgement

References

MANDEMAKERS, K., DILLON, L., *Best practices with large databases on historical populations*, Historical Methods: A Journal of Quantitative and Interdisciplinary History, 37 (2004), 34–38, doi:10.3200/HMTS.37.1.34-38, URL http://dx.doi.org/10.3200/HMTS.37.1.34-38.