

## The process of record linkage on Roman epigraphical sources Theory, methods and results

Angela Lumezeanu, Rada Varga

**Abstract:** This paper explores the methodological aspects and the results of linking individuals attested in ancient epigraphic sources. The article has three main parts, equally important: the first focuses on the sources and the methodology of the linkage process, the second presents the chosen linkage criteria and the results and the final one showcases a couple of examples, in order to illustrate how a ‘digital’ method can lead to reconstructing ancient people’s lives. The conclusions of our endeavor, which we mainly regarded as a methodological experiment, are multi-folded. First of all, the process we undertook proves that record linkage operated on ancient epigraphic sources has a positive finality. Equally, the necessity of manual verification became very evident throughout the process.

### Motivation: why linkage on Roman epigraphy?

This paper explores the methodological and technical aspects, as well as the results, of linking individuals attested in ancient epigraphic sources. The article has three main parts, equally important: the first focuses on the sources and the methodology of the linkage process, the second presents the chosen linkage criteria and the results (eventually and exhaustively synthesized in Tables 1–6, summing up 32 linked persons) and the final one showcases a couple of examples, in order to illustrate how a ‘digital’ method can lead to reconstructing ancient people’s lives.

We extracted our information from Romans1by1 (R1by1),<sup>1</sup> a population database for persons recorded in ancient epigraphy. Using this method, we will be able to track one individual through multiple attestations, inter-link dissipated sources and through these, gain new information on Roman provincial society, economy and private life.

As far as we know, record linkage has not been applied to people attested by ancient epigraphic sources so far, though the method is being increasingly employed when trying to connect facts and data attested through multiple historical sources. It is a procedure used for matching similar entities across different inscriptions and is, at present, widely utilized in historical demography. Linkage has also begun to be employed for reconstructing ancient individual biographies and social environments,<sup>2</sup> though the methodological difficulties caused by the nature of ancient sources remain a problem (as we will stress in the following pages). Many historical databases contain data entry errors such as misspellings, or parameter values which describe changes to an individual’s features over time (moving to another place, changing occupation, gaining different titles, even changing name through adoption, marriage, etc.). To

---

\* We kindly thank Gabriel Bodard, Vlad Popovici and Annamária Pázsint for reviewing the text and for their valuable suggestions.

1 <http://romans1by1.com/>.

2 Hin et al. (2016).

obtain as successful a linkage as possible, data cleansing and standardization are important. In the end, a good record linkage will provide a basis for analysing social mobility, mortality, migration and reconstructing populations and their behaviour patterns.

In sociological and historical research contexts, record linkage is used for recreating individual life courses and thus making deep level analyses possible. It is an important method, especially when the data sets come from different types of sources. Roman epigraphic monuments, for example, can be funerary, votive, honorific, etc. and record information on the people mentioned on them accordingly. Record linkage can, in this case, create cross-source information networks (and implicitly provide biographical data on one person from multiple sources) and enrich the data available to epigraphists and social historians interested in the Roman provincial world.

In the current paper, our main aims are (a) to investigate whether a linkage algorithm can successfully be applied to the scarce and relatively chaotic (from types of information available) epigraphic sources and (b) to find out whether the method offers additional new finds compared to fully manual linkage.

### Romans1by1: the sources

Romans1by1 is a relational database, built with MySQL and following the best practice models for population databases.<sup>3</sup> It comprises, as already mentioned, people attested through ancient Latin and Greek epigraphy. The database has a research component and a smaller open access subset. At the time of writing, the research component (restricted, more functional but less attractive database interface, available exclusively upon login) had approximately 18500<sup>4</sup> people entries and covers a rather wide range of geographical areas, from the Greek and Latin language provinces of the Roman Empire alike. The open-access, public-facing website<sup>5</sup> had 16277 entries and included all people attested epigraphically in Moesia and Dacia (Fig. 1).

The metadata is grouped into four major tables (*Inscriptions*, *Bibliography*, *Personal data* and *Personal relationships*), giving a total of more than 100 separate attributes.

Most relevant for the present study is the table used for recording data about individual persons (labelled *Personal data*), around which the entire network of relations needed to ensure proper information recording is built. Each new entry represents a singular epigraphic attestation of an individual, and a unique ID is generated which helps link the individual within the various components of the database and other database entries.

---

3 Mandemakers, Dillon (2004).

4 At the time of finalizing the paper (25.06.2019): 20.921.

5 <http://romans1by1.com/rpeople/people>.

To establish the context, the general figures are: 4354 people attested in Dacia, 2356 in Moesia Superior and 6359 coming from Moesia Inferior.<sup>6</sup> The people attested epigraphically in these provinces were all registered<sup>7</sup> in our database and are currently available in open access. Pannonia Superior had, at the moment when the analyses were undertaken, 1990 persons registered, as we have covered entirely the urban centres of Poetovio, Nevioudunum, Siscia, Savaria, Brigetio and Scarbantia. The other entries are the focus of various research projects (occupational inscriptions, associations, etc.) and come from the whole of Latin Europe and the Greek colonies around the Black Sea.

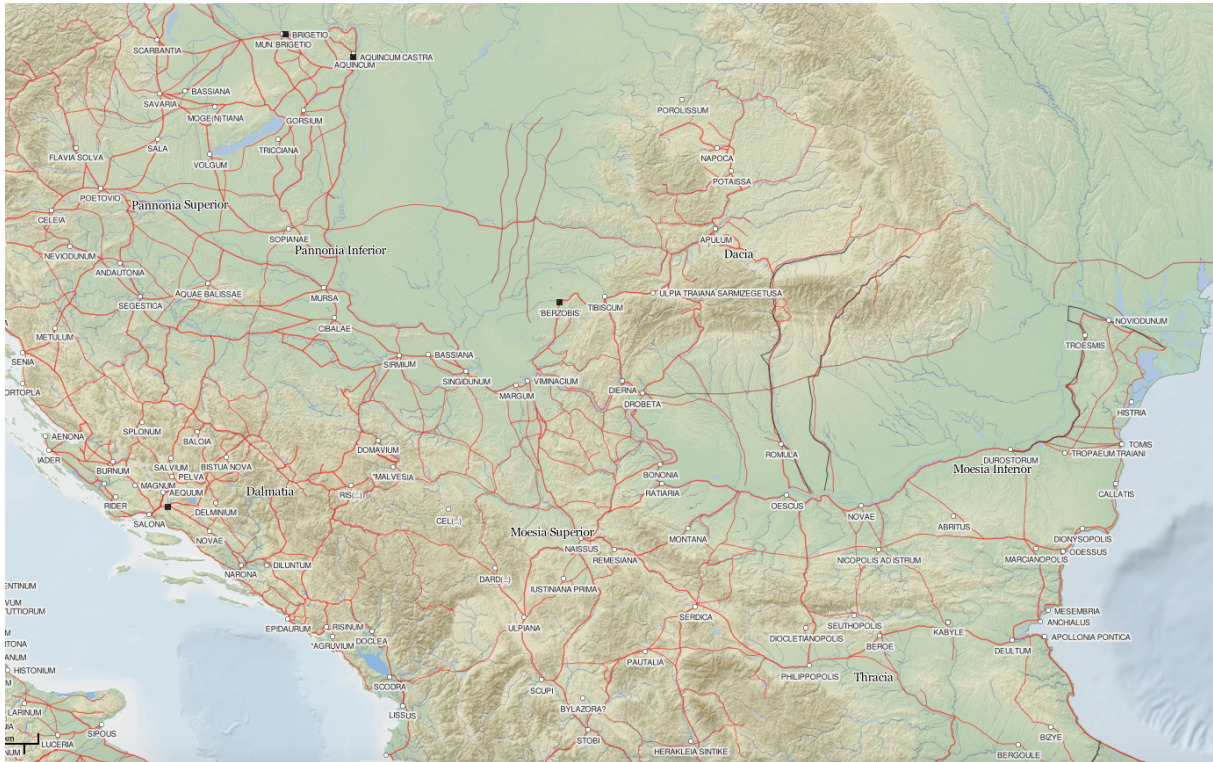


Fig. 1: Map of Moesia Inferior, Moesia Superior and Dacia.

## Manual linkage

One last thing we should explain before describing the linkage process itself is how the database was devised and what it means for our enterprise. Constructing R1by1 was, first of all, a work of historical scholarship; specialized personnel entered the data manually, cross-checking the sources and their bibliography. Thus, we have never dealt with raw data, automatically ingested into a database structure, but rather with controlled data and information provided by responsible specialists. This has influenced the linkage as well, as it meant there was an initial stage of manual linkage.

6 Varga et al. (2018). Our purpose is not to go into the demography of these provinces, but a few facts may be useful at this point: the most recent calculations suggest that, at its peak, the Roman Empire was inhabited by approximately 80 million people. The Lower Danube provinces had around 4 million inhabitants and Dacia up to 1 million.

7 When dealing with such a large amount of sources, sometimes published in local, badly disseminated journals or volumes, there always exists the risk of unknowingly skipping some of them. All this in mind, we tried to collect everything, as the completeness of sources was one of our targets, but we are aware that a small percentage of inscriptions might have eluded us.

As stated before, each new entry represents the epigraphic attestation of an individual, and a unique ID is generated for it. But, in the case of one known person being attested by multiple epigraphic sources, the individual was given single ID from the start and one-person entry may be linked to several epigraphic sources. In this case we have a manual record linkage, performed by the specialists who inputted the information.

Thus, a total of 477 persons had already been linked manually<sup>8</sup> and registered as attested on multiple sources, at the time we began the automated linkage procedures. Some of them have been identified as the same person attested in one or more documents by the researchers working on the database (and sometimes published as individual case studies),<sup>9</sup> but most of the multiple attestations of the same individual were recorded as such in previous bibliography. In presenting our results, we will exclude these persons, although, of course, they would have constituted linkage hits. We have, however, considered that their identification as unique characters represents either other people's work or previous results we have already highlighted, through 'classical' paper publications or directly through the database.

### The linkage process

Record linkage is the process of identifying data sets that refer to the same entity across multiple sources, to check and find duplicate records of an individual and merge them. Thus every piece of information about the entity is bound together in order to build a complete record of that entity – a person in our case. It is a method of standardizing information identified in sources that do not have a common denominator (ID, national identification number, social number, etc.).

For successful and accurate data linkage, cleansed and high-quality data sets are needed, especially their key fields (the ones which will be directly used as linkage criteria). So, the first step in record linkage is the preliminary process of standardizing and normalizing the data set(s). The next step is to undertake entity resolution using different filters (for example, for modern individuals bearing the same name, one can use the date of birth, parents' names, etc. as disambiguators, thus eliminating some matches and rendering others as likely real identity matches).

For these machine-generated results to have absolute value, a human operator should ideally cross-check and revise them. The fully automated linkage can generate, besides valid results, type I and type II errors, as the statistical sciences call them. A type I error is a false positive, or a false match: the assumption that a relationship exists, when in fact it does not. A type II error is a false negative: a valid match not being registered, a missed linkage hit in our case.

Type I errors can be corrected through verification of the linked entities by a specialist. On the other hand, for linkage applied to historical documents (and not exclusively, of course), type II errors are an accepted risk and due to the insufficiency and/or inconsistency of the data provided by the sources, the historian regards them as unavoidable, especially when working on large and not necessarily standardized data sets. While type I errors can (it is hoped) be corrected by cross-checking all information offered by the linked sources and additional historical background data, type II errors cannot be fixed due to lack of information. Some of the hits we have, throughout the process described below, considered type I errors might not be errors, but the inconsistency and scarcity of the information make them highly risky

---

8 These are available as such in the database. For some examples, see: ID 596 <http://romans1by1.com/rpeople/596> (12 links); ID 718 <http://romans1by1.com/rpeople/718> (15 links); ID 3073 <http://romans1by1.com/rpeople/3073> (26 links).

9 Varga (2013), Varga (2016).

hits. In a world with no identification papers, unique personal numeric codes, social security numbers, etc., identification is considerably more difficult than it is when applied to modern sources. For example, if a Marcus dedicates to Apollo in Tomis during the second half of the 2nd century AD and a Marcus is buried by his wife in Tomis during the second half of the 2nd century AD – these being the only facts recorded by each inscription – we cannot state that we are dealing with the same Marcus. We have no historical grounds for this identification, though theoretically he could be the same person.

Linkage of historical individuals can be manual, semi-automated or automated. To obtain accurate automated linkage, the information offered by sources has to be highly standardized, with powerful disambiguation criteria. In historical modern-era sources the most common problems are the different ways of spelling names (Mary, Mari, Marie), the lack of family names, different types of recorded data or their absence, data recorded in different languages. In our case, the number of empty variables is larger. So, in order to test the linkage possibilities, we used a semi-automated method. Employing a series of procedures written in SQL (Structured Query Language), the computer will compile a table with the probable matches of identical individuals from the R1by1 database. To check if the individuals are correctly linked, we undertook a manual check.

The steps of record linkage, on the data set of R1by1, were:

- preparing the data by creating coherent identifier fields;
- analysing the nature and content of the data set and creating multiple MySQL stored procedures and functions for the linkage (detailed in ‘Linkage criteria and results’);
- manual verification of all the hits and elimination of errors.

As mentioned, preparing the data comes first. The R1by1 database has a table for individuals, already normalized and standardized when we initiated this research, but it has 111 columns, not all of them necessary for linkage. So we created a new table for individuals using MySQL and extracted only the data that we considered important for linkage and which contained sufficient information that could be used for constructing a strong variable. There are 23 columns in total: *praenomen*, *nomen*, *cognomen* (all the individual parts of the Roman citizen’s name – *tria nomina* – compulsory in theory, but not all recorded, all the time, on stone monuments), occupation/profession, juridical status, military status, etc. The number of individuals having these pieces of information registered and extracted from the database is 17643.

The next step was to analyse how we can apply record linkage procedures to the available data. It was impossible to use the same criteria for all the individuals, due to the nature of the sources; for example, 46% of the individuals only have the *cognomen* registered, and only 5% of them have the occupation mentioned. In order not to lose valuable information, we had to break down the data and apply different criteria for linkage. We created several MySQL stored procedures (around 10) and functions and then applied them on the *Individuals* table.

To find the names with variant spellings, we used the Levenshtein distance. The Levenshtein distance is a method for measuring the similarities between two strings; it calculates how many deletions, substitutions or insertions are used to have a match between the two strings. The result number is called *edit distance*. For a perfect string match there is no need to apply any measure for matching, so it will be 0 editing giving a score of 100% match. For searching we applied the Levenshtein ratio to the *cognomen*. A score of 100 means a perfect match and the score goes down if the strings being compared have different characters. The *edit distance* is normalized, divided by the lengths of the longest string. So, for instance, when comparing the cognomen Apollinaris to Appolinaris the number of insertions/deletions needed for a perfect match is 2 out of 11 characters (the letters composing the name), resulting in a ratio of approximatively 82%.

We created several .csv files (tables) using different matching criteria: occupation, military status, juridical status (citizens, slaves, freedmen), religious position and we proceeded to manual verification.

Employing fully automated linkage, without manual verification, means the possibility of errors is huge when working on unstructured sources and the percentage of type I errors was, in our case, massive (over 90%, on certain samples). Equally, we encountered type II errors due to the original, completely unstandardized, nature of our sources. These ‘anomalies’ are not directly connected to the linkage criteria applied, but to the lack of consistent information in many cases. For example, on an official dedication a man generally wrote all his names and titles, but for more private acts of devotion, he did not have to.

Thus, the linkage decision ultimately has to be made by the historian, based on the possible matches suggested by the computer algorithm. However, the manual linkage is also quite problematic. If we use only the *cognomen* as a criterion, we have 11784 perfect matches. Thus, the personal name alone was not considered a linkage criterion and the variables consisted of combined values.

In addition, trying to apply more sophisticated, but still onomastically based, filters sometimes led us to dead ends, which will be explained below. The onomastic matches, though important for emphasizing or confirming the popularity of certain names in given areas and/or social environments, remain vague and basically useless if not associated with other disambiguating elements.

### Linkage criteria and results

Given these circumstances, we had to apply a combination of filters and operate the linkage following multiple criteria, in order to make manual verifications cost-effective (as man/hours implied by it). Applying each of these criteria generated a different table of results.<sup>10</sup> The potential matches, which would have led to overestimation in the linkage results (type I errors), made us very cautious when applying filtering criteria. While we were keen to discover examples of mobility, in certain cases, as detailed below, we were compelled by the large number of possible matches to apply a geographical disambiguation filter (the province of attestation).

---

10 One criterion that would seem self-evident is the dating of the inscriptions. Unfortunately, most of our monuments are dated inaccurately, with approximations of half a century or even a century. Under the circumstances, the timeframe was used as verification historical criterion (along with palaeography and other stylistic details pertaining to the monument) when performing the checks following the automated linkage.

The personal name was the constant – though painfully insufficient – criterion. Besides this, several variables were used as linking variables and we created more linkage spreadsheets in order to make the manual examination efficient and relevant:

a) One of the most powerful identifiers in Latin epigraphy is *profession* (Table 1). Combined with the individual's name, the profession is a fairly reliable criterion.<sup>11</sup> The automated linkage indicated 32 possible matches; 6 of them were valid person linkages. The rest of the cases were either coincidences or persons of the same family (usually father and son), sharing a name and a profession.

	ID	Name	Occupation	Bibliography	ID	Name	Occupation	Bibliography
1.	16142	Fortunatus	<i>adiutor tabularii</i>	CIL III 4023; <sup>12</sup> CIL III 4062	1188	Fortunatus	<i>adiutor tabularii</i>	IDR III/2, 396
2.	3008	Mucatralis	<i>magister vicus</i>	ISM I 330	3041	Mucatralis	<i>magister vicus</i>	ISM I 342
3.	3057	Iustus	<i>magister vicus</i>	ISM I 349	3053	Iustus	<i>magister vicus</i>	ISM I 347
4.	2453	Lucius Placidus	<i>negotiator</i>	RIB III 3195	1254	Placidus	<i>negotiator</i>	AE 1975, 651
5.	831	Carpion	<i>tabularius</i>	IDR III/2, 277; IDR III/2; 387	8957	Carpion	<i>tabularius</i>	IDR III/5, 10
6.	1210	Caius Aurelius Verus	<i>negotiator</i>	CIL XIII 8164a	1265	Caius Aurelius Verus	<i>negotiator</i>	AE 1983, 722

Table 1. Linked IDs following the profession variable.

b) *Juridical status* is also relevant and can be employed in the linkage, although sometimes it is ambiguous and was not recorded in the database. For citizens, we used full *tria nomina* (Table 2) as disambiguation criterion, though we were very aware this could lead to type II errors. Equally, we used the attestation province as a filter, to keep the linkage results historically realistic – possible. With these filters, we got 56 possible matches, 10 of them being real links. We must stress that this was the table with the second-best success rate.<sup>13</sup>

11 As an experiment, we applied the Levenshtein distance on the cognomen/personal name down to 70; in this context, we realized that below 80 the matches were far too approximate to be taken seriously (the names Ponticus and Ionicus, unmistakably different, have a 75 u match).

12 We have decided to offer one major bibliographic source for each attestation – generally one of the major, representative ancient epigraphy *corpora*. When more than one references are quoted for an ID it means that the ID was manually linked, pointing to two or more sources.

13 An interesting fact is that we got a double linkage, namely the IDs 1210 and 1265 – Caius Aurelius Verus – who, being a negotiator, was identified according to both criteria presented so far (Tables 1 and 2).

	ID	Name	Bibliography	ID	Name	Bibliography
1.	15486	Publius Atinius Amerimnus	AE 2003, 1549	15542	Publius Atinius Amerimnus	Eck, Pangerl 2009, no. 5
2.	1718	Caius Arrius Antoninus	IDR III/2, 85; IDR III/2, 86	14102	Caius Arrius Antoninus	AE 19998, 1117
3.	13302	Lucius Cadius Celer	AE 1952, 191	3053	Lucius Cadius Celer	ISM I 347
4.	14003	Lucius Viracius Fortis	ISM VI 225	14007	Lucius Viracius Fortis	ISM VI 228
5.	9168	Marcus Aurelius Rogatianus	IDR III/5, 105	9754	Marcus Aurelius Rogatianus	IDR III/5, 257; IDR III/5, 361
6.	15049	Publius Attius Severus	AE 1999, 13131	15266	Publius Attius Severus	Eck, Pangerl 2017
7.	12248	Lucius Iulius Valens	AE 1913, 117	13813	Lucius Iulius Valens	IMS III/2, 85
8.	15485	Lucius Pullius Verecundus	AE 2003, 1549	15496	Lucius Pullius Verecundus	AE 1998, 1169
9.	1210	Caius Aurelius Verus	CIL XIII 8164a	1265	Caius Aurelius Verus	AE 1983, 722
10.	15495	Publius Caulius Vitalis	AE 1998, 1169	15545	Publius Caulius Vitalis	Eck, Pangerl 2009, no. 5

Table 2. Linked IDs using the citizen status & full tria nomina variables.

For freedmen (Table 3), the first option we had was linking by *nomen* and *cognomen*, but in this case, we risked losing those that registered only the personal name. For cross-checking, we extracted both tables, with 2 and 3 variables respectively. In the first case scenario, we obtained 139 automatically indicated links, with only 1 being valid, so a very small success rate; and for the second case scenario, 13 possible matches with the same individual hit. So, in this case, introducing an extra variable did not lead to information loss, but made the manual control much more efficient.

	ID	Name	Bibliography	ID	Name	Bibliography
1.	1847	Publius Tenacius Gemellinus	IDR III/2, 120	11310	Publius Tenacius Gemellinus	IDR III/5, 518; IDR III/2, 582

Table 3. Linked ID using the libertus/liberta status variable.

Slaves were a very complicated category to link, as they only had one name, did not always mention the name of the master (thus the equivalent of the freedmen’s *nomen* in terms of disambiguation) and frequently bore similar names (we have the so-called servile names, Felix, Fortunatus, Hilarus, etc., given to slaves throughout the Empire and which offer no extra personal information on the individual whatsoever). 174 possible matches resulted when using the name and juridical status as sole values. But the mobility of personal slaves was very low, so we decided to make the province of attestation the third linkage criterion. We obtained 18 linkages, all of them, unfortunately, type I errors.

The last juridical category linked using this criterion was that of the peregrines – free men and women, who did not hold Roman state citizenship (Table 4). Their names generally have a structure of personal name – patronymic in genitive; but, of course, the father’s name can be absent from the more modest or



very personal monuments. As their mobility was also low, we again used province as a third variable.<sup>14</sup> Out of 127 matches, 3 were real identity matches.

	ID	Name	Bibliography	ID	Name	Bibliography
1.	7793	Senecio	IDR II 60	7407	Senecio	IDR III/4, 198
2.	7406	Sabina	IDR III/4, 198	7792	Sabina	IDR II 60
3.	6197	Planius	IDR III/3, 423	13476	Planus	IDR I 36

Table 4. Linked IDs using the peregrine status variable.

c) A final way to disambiguate and link effectively was using positions, roles (religious, political, etc.) and ranks as variables. The risk of this linkage procedure producing type II errors is quite high, as not all positions had to be listed on all inscriptions mentioning the same person. Nonetheless, the method yielded some results.

One important variable was the religious position – we obtained only 4 previously unknown possible matches and none of them proved to be more than onomastic links. In this case, we made an observation which was to prove valid for decurions<sup>15</sup> and other elite groups as well: there is no point in accepting a linkage lower than Levenshtein ratio of 100%, as the monuments are very carefully carved, the names neatly set out and always correctly spelt.

Another important variable was the quality of decurion.<sup>16</sup> We had in mind primarily the decurions whose attestations come from the same province – we obtained 35 possible matches, with 11 real hits, this being the set with the highest success rate we worked on (Table 5). The decurions with identical names, but attestations in different geographical areas, provided 37 links, all improbable real character matching.

	ID	Name	Bibliography	ID	Name	Bibliography
1.	1844	Marcus Opellius Adiutor	IDR III/2, 116; IDR III/2, 117; IDR III/3, 317	5789	Marcus Opellius Adiutor	IDR III/3, 350; IDR III/3, 377
2.	9701	Aelius Antipater	IDR III/5, 215; <sup>17</sup> IDR III/5, 259	583	Publius Aelius Antipater	IDR III/2, 217; IDR III/1, 65; IDR III/5, 215; IDR III/5, 210; IDR III/5, 439
3.	9222	Flavius Germanus	AE 1998, 1143	8688	Flavius Germanus	ISM IV 40
4.	11089	Publius Aelius Iulianus	IDR III/5, 441	9702	Aelius Iulianus	IDR III/5, 215; IDR III/5, 259
5.	830	Marcus Ulpus Maius	IDR III/2, 293	1393	Marcus Ulpus Maius	IDR III/2, 447
6.	8653	Iulius Marcus	ISM IV 12	8638	Iulius Marcus	ISM IV 11

14 Here, without the geographical disambiguation, we got 1,450 possible links, an unrealistic figure for sure, given the fact that we are, in the end, talking about a poor and dependent category.

15 Member of a city council.

16 At this point, we should mention that we did not use the *ordo senatorius* or *ordo equester* filters because prosopographies of senators and knights exist; they are well-known and generally identified on monuments.

17 In this case, we've corrected to the manual linking: the character was linked to the same source for both IDs, but not identified as one – a normal and frequent mistake when work is done manually.

7.	5545	Aurelius Maximus	IDR III/3, 215	7294	Marcus Aurelius Maximus	IDR III/4, 63
8.	12047	Marcus Titius Maximus	CIL III 6127	12032	Marcus Titius Maximus	AE 1986, 117
9.	1141	Lucius Antonius Rufus	IDR III/2, 104	1139	Antonius Rufus	IDR III/2, 375; IDR III/2, 376; IDR III/2, 440; IDR III/2; 542
10.	5627	Marcus Antonius Saturninus	IDR III/3, 286; IDR III/3, 293; IDR III/3, 322; IDR III/3, 330; IDR III/3, 335	5653	Marcus Antonius Saturninus	IDR III/3, 309; IDR III/3, 308
11.	1818	Caius Iulius Valerius	IDR III/3, 113	9081	Caius Iulius Valerius	IDR III/5, 60

Table 5. Linked IDs using the decurio position & province variables.

A class very difficult to work with was the militaries – numerous, as we were dealing with border provinces; there were 2410 in the database. We tried to link them based on name as well as other details. Again, we decided to link people with identical *tria nomina*, aware that we might lose examples from fairly personal, unofficial monuments. As militaries had higher mobility than other groups, we did not use the province as a linkage variable. Nonetheless, we obtained 108 matches, all of them onomastic. The militaries’ onomastics are notable for their simplicity; they rarely have unusual names. The manual validation process can be very confusing, as the sources themselves offer very few adjacent details and make acceptance or rejection of the suggested link very difficult to decide upon. As an example, 50 of the links are based on the *cognomen* Valens, 10 of them being Caius Valerius Valens. Valens, Victorinus, Maximus, etc. are typical military names which, combined with ‘mainstream’ imperial *praenomina* and *nomina* and lack of other biographical or prosopographic details, make disambiguation and linkage validation impossible. Another linkage possibility we saw for the militaries was using the troop as a criterion (Table 6) as well as the name. Besides sometimes changing troops, soldiers are less often mentioned on monuments than might be imagined, because everybody from the community knew which troop was stationed there and the soldiers implicitly belonged to it. So, we were able to obtain 5 links, one of them being an identity hit.

	ID	Name	Bibliography	ID	Name	Bibliography
1.	3436	Titus Aelius Celer	AE 2006, 1128	3426	Celer	CIL III 789

Table 6. Linked IDs using the military status & military troop variables.

## Life stories

What linkage, be it automated or manual, finally yields are the possibility to reconstruct life courses and events. When it comes to Roman provincial ‘middle classes’,<sup>18</sup> any detail is important, any extra piece of knowledge is welcome and informs us about social and economic realities.

18 This paper does not engage in the ‘controversy’ of the provincial middle classes, their definition, social and economic boundaries – or even their mere nomenclature or the justification for using the anachronistic expression ‘middle class’. Thus, we employed the syntagma as a convention for denominating non-elite categories.

We will select some of the most relevant and coherent examples of life reconstructions resulting from the linkage process.

a) *Longitudinal family reconstructions*. Some of the most interesting cases are those involving more than one generation of a family and reveal previously unknown ties. From Moesia Superior, modern-day Kumavo in Macedonia, come two epitaphs revealing some data on the life and family of Lucius Viracius Fortis. We know nothing about this man's social standing or profession, but we know that he was married to Publicia Callirhoe (AE 1984, 782), who died at the age of 50 and whom he buried. He dedicates the epitaph alone, so we can speculate the couple had no surviving children at the time of the wife's death. Through record linkage, we found another attestation of the same individual, and though neither is very rich in details, we now also know that he was named after his father, Lucius Viracius Fortis, and had a sister named Viracia Ingenua (AE 1984, 785). The siblings buried their father, who died at 70, implying that they were sole heirs. The naming pattern within the family is fairly clear, as both children bear the father's *nomen* and Ingenua might be the name of the mother or a derivative of it. Another thing we might note in this example is the ages of both deceased – it may be a coincidence, but is more likely a case of age rounding;<sup>19</sup> the couple did not know or care about the exact year of birth.

Sometimes, even 'imperfect' links can offer details of family life, even if the subtext remains partly hidden. As an example, we have the case of Marcus Valerius Verus (CIL III 4069) and Lucius Valerius Verus (AIJ 288), attested in two different inscriptions from Poetovio. While they definitely are two different persons, one can safely assume a family connection due to the name, the fact that they both belong to the *ordo decurionum* and the dating during the first half of the 2nd century AD of both dedications.

b) *Life events*. An intriguing and worthwhile emerging methodology comes from the Lived Ancient Religion group<sup>20</sup> and focuses on individualizing texts (and textualizing individuals). This means that, although the text is often a standard one, it becomes part of the individual's biography. The concept was developed in the context of religious dedication/adoration texts, but it fits for the analysis of any monument, as the simple fact of erecting a stone monument clearly becomes an important part of an individual's life course (and somewhat ignored by scholars). This methodological framework is well suited to the results of our linkage and the ways in which they allow us to reconstruct life events.

Thus, we have Marcus Opellius Adiutor, son of Marcus and member of the tribe Papiria (emperor Trajan's tribe, so proof that his family belonged to Dacia's first wave of citizens), decurion and *duumvir* of Sarmizegetusa, patron of the association of craftsmen (*collegium fabrum*) (AE 2003, 1514 = IDR III/2, 116). We know that the association and the *ordo decurionum* honoured him with monuments placed in the city's forum more than once (IDR III/2, 116; IDR III/2, 117), proving his status in local society. But more than this, we find out that he himself dedicated an altar to Jupiter and the other *dii consentes* (the 12 main gods of the Roman pantheon) (CIL III 942 = IDR III/3, 317). This is a very common and conventional form of devotion, which does not say much – but the place where the altar was erected does. The monument comes from Ampelum, modern-day Zlatna, a mining settlement, 100 kilometres north of Sarmizegetusa and part of the *territorium* of the *colonia*. A dedication here most probably indicates that Adiutor had either a property, a *villa*, in the area or business interests in the mines – or both. His connections to Ampelum are reinforced by the existence of another monument (AE 1988, 953 = IDR III/3, 350) and a stamped tile (IDR III/3, 377), previously unlinked to the group of inscriptions described above. The monument is a funerary dedication for a young (30-year-old) decurion from Aequum, a city in Dalmatia, who died here. Worth mentioning is the fact that the gold mines area, situated just 30–40 kilometres from Ampelum, had been heavily colonized with Illyrians, many of them from Dalmatia. As

<sup>19</sup> Cooley (2012), 53.

<sup>20</sup> <https://www.uni-erfurt.de/en/max-weber-centre/projects/cooperation-projects/lived-ancient-religion/>.

Adiutor buries the man as heir, we can safely assume the inheritance was the deceased's share a joint business. The last attestation of Adiutor seems more modest – a stamped tile – but is equally important. The tile, bearing his initials, indicates that either he owned an *officina* producing bricks and tiles, or a very large stock was delivered to him – for a large construction, presumably. What we do not have is the final part of this interesting life story – a funerary dedication, which would have presumably indicated if the man had a family, children, slaves, heirs, etc. So, in this case, linking the character throughout all these inscriptions led to us finding out important details on his social standing and on the Dacian mining business, but nothing allowing family reconstruction.

c) *Mobility*. Traders are, of course, one of the groups with the highest mobility. A case we linked is that of Placidus Viduci f., *negotiator Britannicianus* and *cives Veliocassinius* (AE 1975, 651 = AE 1982, 724). He makes a dedication to the goddess at Ganventa around 200 AD, without mentioning anything else about his trade or himself on the monument. If this were the only source on Placidus, it would still have been most interesting to see a man from Gallia Lugdunensis, with a local, probably native, citizenship operating on the Germania Inferior – Britannia trade route. Nonetheless, the trader is also attested on a second dedication from 221, from Eburacum (AE 1977, 512 = RIB 3195). This is a construction inscription, dedicated to the *genius loci* and *numina* of the emperors, for a vault and passageway erected by the merchant. While the citizenship and the quality of trader of Britannic routes are expressed in a more or less similar fashion as in the previous inscription, the name of the dedicator is Lucius Viducius Placidus, with the possible presence of the patronymic between *nomen* and *cognomen*.<sup>21</sup> There are several possible reasons for the name mismatching. One is that he became a citizen through the Constitutio Antoniniana,<sup>22</sup> or in other circumstances, between the dedications of the two monuments and chose to follow the Germanic tradition of assuming the father's name as a *cognomen*. Another hypothesis is that he had citizenship at the time of the Dea Nehalennia dedication as well, but as it was a more private act of devotion did not feel the need to write his whole, official name on the monument. The father-son hypothesis, though attractive at first sight, was rightly discarded by Wim Broekaert. Another matter raised by Placidus's inscriptions is that of his real residence. The euergetism of a stranger was rare in Britannia – in the whole Western Empire, in fact – and, along with the indication of *negotiator Britannicianus*, might suggest that the *cives* from Lugdunensis had moved to, lived and worked in Eburacum. As a 'newcomer', it was important to establish status, hence the self-financed construction work for use by the community.

## Conclusions

Epigraphic linkage sheds light on particular cases and increases knowledge of Roman society by allowing us to understand slices of real lives – not from books written by the elites, but from inscriptions dedicated by the 'middle class' itself.

Before we started the process of record linkage on our sample, we had 476 cases of linked individuals. These matches were done by manual linking when entering the data, as explained above. After the semi-automated process, we found 32 more individual matchings, so we increased the linked persons sample by 7.5%. If we take into account that the database covers a large geographical area and also a very extensive timeframe (more than 400 years), to which we can add the substantial lack of recorded data, the linkage percentage of 2.5% – 3% is fairly normal. In Dacia and Moesia, which were covered

<sup>21</sup> Broekaert (2013), 220–222, for commentaries.

<sup>22</sup> Constitutio Antoniniana was an edict issued by Emperor Caracalla in 212 AD. Through this constitutive act, the emperor granted citizenship to all the free inhabitants of the Roman Empire (Ulp. D. 1,5,17; Cass. Dio LXXVII 9, 5).

extensively (14323 individuals), we have the majority of identified links found (87% of linked records), which gives us an idea of the impact of record linkage on regions that are thoroughly documented. These figures and percentages might seem small for a historian dealing with other historical periods, but for classical antiquity and especially for a provincial environment they are satisfactory.<sup>23</sup>

The conclusions of our endeavour, which we mainly regarded as a methodological experiment, are multifolded. First of all, the process we undertook proves that record linkage operated on ancient epigraphic sources has a high value and it is something worth investing effort in. Equally, the necessity of manual verification became very evident throughout the process and this is closely connected to our second conclusion: working on unstandardized primary sources changes the rules and requires resourcefulness in dealing with the variables, as well as the ‘touch’ of a historian specialized in the considered period and region. The third set of conclusions is of scientific nature and regards the fact that record linkage operated on Roman-era epigraphy does bring forth scientifically important data; by reconstructing life courses and family groups from provincial milieus of the Roman Empire we gain new knowledge on society, economy, commemoration patterns and ultimately the socio-psychology of these areas.

---

23 In this context, one must pay particular attention to the relation between epigraphic representativeness and the relevance of quantitative data. To this end, J. Herman’s work (Herman [1983], 1047) is especially interesting, dealing with the use of Latin in the Galli. Starting from some demographic approximations, he formulated the hypothesis that the people epigraphically attested in Gallia during the entire period of Roman domination represent 0.02% of all the inhabitants of the three Roman provinces. At the same time, Herman supposed that the percent of the population actually expressing itself through epigraphy did not exceed 0.5%. Following the same argumentative line, but insisting less on statistics, G. Alföldy (1965, 19 sqq.) noted that 4400 funerary inscriptions have been preserved from Dalmatia during the 1st–3rd centuries AD, compared to an estimated population of the province of 600.000 – 700.000 inhabitants during a single generation (one could presume that a sum of ten successive generations lived in the province during the Principate).

## References

- Alföldy (1965): Alföldy, G., *Bevölkerung und Gesellschaft der römischen Provinz Dalmatien*, Budapest.
- Broekaert (2013): Broekaert, W., *Navicularii et negotiantes: a prosopographical study of Roman merchants and shippers*, Rahden/Westf.
- Cooley (2012): Cooley, A., *The Cambridge Manual of Latin Epigraphy*, Cambridge.
- Herman (1983): Herman, J., *La langue latine dans la Gaule romaine, Aufstieg und Niedergang der römischen Welt II 29 (2)*, 1045–1060.
- Hin et al. (2006): Hin, S., Conde, D. A. and A. Lenart, *New light on Roman census papyri through semi-automated record linkage*, *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 49 (1), 50–65.
- Mandemakers u Dillon (2004): Mandemakers, K. and L. Dillon, *Best Practices with Large Databases on Historical Populations*, *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 37 (1), 34–38.
- Varga (2013): Varga, R., *Two inscriptions from Sarmizegetusa revisited*, *Studia Antiqua et Archaeologica* 19, 79–86.
- Varga (2016): Varga, R., *Aurelius Aquila, negotiator ex provincia Dacia. A prosopographic reconstruction*, in: R. Ardevan, E. Beu-Dachin (eds.), *Mensa rotunda epigraphica Napocensis*, Cluj-Napoca, 27–34.
- Varga (2017): Varga, R., *Romans1by1 v.1.1. New developments in the study of Roman population*, *Digital Classics Online* 3 (2), 44–50.
- Varga et al. (2018): Varga, R., Pázsint, A., Boda, I., Deac, D., “Romans 1by1. Overview of a research project”, *Digital Classics Online* 4 (2), 37–63.

## Corpora

AE – *Année Épigraphique*, Paris.

AIJ – Hoffiller, V. and B. Saria. 1938. *Antike Inschriften aus Jugoslawien*. Zagreb: Druck der Fondsdrukkerei der ‘Narodne novine’.

CIL – *Corpus Inscriptionum Latinarum*, Berlin.

IDR – *Inscriptiones Daciae Romanae*, București-Paris.

ISM – *Inscriptiones Scythiae Minoris*, București-Paris.

RIB – *Roman Inscriptions of Britain*, Oxford.

## Autorenkontakt<sup>24</sup>

### **Dr. Angela Lumezeanu**

Center for Population Studies,  
Babeş-Bolyai University, Cluj-Napoca  
Avram Iancu St. 68  
400083

Email: [angie.lumezeanu@gmail.com](mailto:angie.lumezeanu@gmail.com)

### **Dr. Rada Varga**

Center for Roman Studies,  
Babeş-Bolyai University, Cluj-Napoca  
Moşilor St. 11  
400001

Email: [radavarga@gmail.com](mailto:radavarga@gmail.com)

---

24 The rights pertaining to content, text, graphics, and images, unless otherwise noted, are reserved by the author. This contribution is licensed under CC BY 4.0.