EMPLOYING DIGITAL PROSOPOGRAPHY IN THE STUDY OF MID- AND UPPER SOCIAL STRATA IN TRANSYLVANIA (MID-EIGHTEENTH TO MID-TWENTIETH CENTURIES): TOOLS AND APPROACHES

VLAD POPOVICI
AND ANGELA LUMEZEANU*

Prosopographical research on the mid- and upper social strata in Transylvania, in a time of massive digitisation

Until not so long ago, researching rural society's upper strata, the petty bourgeoisie, but sometimes even members of the elite in Transylvania, from mid-nineteenth to mid-twentieth century was a difficult task, due to the scarcity of biographical information. Leaving aside the vital data provided by parish registers, one was usually lucky to find a mention in newspaper about a person holding a particular position at a certain time, to find a funeral announcement that helped reconstruct a family, or to find a particular person's name mentioned in the yearbook of an educational institution. Prosopographies covering non-elite social strata are few in number and deal mainly with university students,[1] while the bulk of biographical dictionaries deal mainly with political or cultural figures. Historical writing was, and remains to date, interested mainly in politics and culture, thus the closer the research subjects got to either of these two fields of activity, the higher their chances of stirring historians' interest.[2]

* Vlad Popovici, PhD, Researcher, Masaryk Institute and Archives of the Czech Academy of Sciences, Prague, Czech Republic; Lecturer, Babeş-Bolyai University, Cluj-Napoca, Romania; e-mail: vladutpopovici@yahoo.com. Angela Lumezeanu, PhD, Researcher, Centre for Population Studies, Babeş-Bolyai University, Cluj-Napoca, Romania; e-mail: angela.lumezeanu@ubbcluj.ro.

[1] Selectively: Cornel Sigmirean, *Istoria formării intelectualităţii româneşti din Transilvania şi Banat în epoca modernă* [History of the Formation of the Romanian Intellectuality in Transylvania and Banat in the Modern Era] (Cluj-Napoca: Presa Universitară Clujeană, 2000); László Szögi, József Mihály Kiss, *Magyarországi diákok bécsi egyetemeken és akadémiákon 1849-1867. Ungarländische Studenten an den Universitäten und Hochschulen in Wien 1849-1867* (Budapest: ELTE Levéltára, 2003); Miklós Szabó, Zsolt Simon, László Szögi, *Erdélyiek külföldi egyetemjárása 1849-1919 között*, [Transylvanian Students Abroad between 1849-1919], vol. 1-2 (Marosvásárhely: Mentor Kiadó, 2014). An exception is represented by the prosopographies of the delegates at the Great National Assembly in Alba Iulia (1 December 1918). See the most recent one and the listed literature on the topic: Dragoş Ursu, Tudor Roşu, ed., *Dicţionarul personalităţilor unirii* [The Dictionary of the Personalities of the Union] (Cluj-Napoca: Mega, 2019), 653-663.

[2] Vlad Popovici, "Does History Remember Civil Servants? A Case Study on the Regime Change of 1918-1920 in Transylvania," in Dumitru-Cătălin Rogojanu, Cosmin-Ştefan Dogaru, ed., *Elites, Networks of Power and Citizens (19th-21st Centuries)* (Cluj-Napoca: Presa Universitară Clujeană, 2019), 121-140.

Annales Universitatis Apulensis Series Historica 25, I (2021): 123-143;
https://doi.org/10.29302/auash.2021.25.1.7.

The situation has changed dramatically, and for the better, in the last ten years. Today, with the help of massive digitisation projects (such as Arcanum,[3] the Digital Library of Bucharest,[4] the digital repositories of the University Library in Cluj),[5] genealogical websites[6] or even through a simple search on Google, piecemeal information can be found about almost anybody who was not a mere villager in Transylvania between the mid-1800s to mid-1900s. The digitisation of the archives moves at a slow but still steady pace,[7] bringing most of the nineteenth-century population onto the historian's keyboard and computer screen. However, this inflow of data, waiting to be grasped and used by researchers, comes with a number of caveats.

First and foremost, the development of proper tools for extracting and structuring data from digitised narrative sources is still far from satisfactory. In fact, for the territory of Transylvania it has barely started, and most of the databases currently functioning are built by using manually extracted data, which implies a certain level of subjectivity, since the data is selected according to the current database creator's needs, and not always serving the prospective needs of other potential users. Exceptions exist, the best example being the "Historical Population Database of Transylvania",[8] which accumulates all the information from parish registers. Nevertheless, such an undertaking inevitably takes a heavy toll on the pace of database creation and on databases' chronological and geographical coverage.

In other cases, even if the database has been built to cover the primary sources as thoroughly as possible,[9] the sources' diversity may be overwhelming, requiring, in the medium and long term, constant adaptations which suck up time and resources away from content-building and redirect them towards the reshaping and maintenance of the digital tool. This brings up a different issue, namely the availability of financial resources, the necessity of collaborating with trained computer specialists and, most importantly, the willingness on the part

[3] "Arcanum Digitheca," accessed 10 September 2021, https://www.arcanum.com/hu/adt/.
[4] "Biblioteca Digitală a Bucureștilor" [The Digital Library of Bucharest], accessed 10 September 2021, http://www.digibuc.ro/.
[5] "Biblioteca Digitală BCU Cluj" [The Digital Library of the University Library in Cluj], accessed 10 September 2021, http://dspace.bcucluj.ro.
[6] "Family Search," accessed 10 September 2021, https://www.familysearch.org/en/, "Ancestry", accessed 10 September 2021, https://www.ancestry.com, "MyHeritage", accessed 10 September 2021, https://www.myheritage.ro, "RadixIndex", accessed 10 September 2021, https://www.radixindex.com.
[7] "Arhivele Naționale ale României, eANR" [National Archives of Romania, eANR], accessed 10 September 2021, http://arhivelenationale.ro/site/eanr/.
[8] Ioan Bolovan et al., "Historical Population Database of Transylvania. A Database Manual," *Studia dig.* 64, no. 1 (2020): 10-85.
[9] Sandra Hirsch et al., "Digital Framework for the History of the Austrian Military Border in Transylvania," *Studia dig.* 64, no. 2 (2020): 5-53.

of humanist scholars to step out of their mainly narrative universe into a differently structured, data-focused environment.

Secondly, biographical information comes in bits and pieces, usually linking a person to a context (e.g., individuals holding an office, or running for office, or having a cultural or economic or social initiative registered in the newspapers), or a person to another person, or small groups at best (e.g. baptism, wedding or funerary announcements). The researcher rarely finds a string of biographical data, let alone cross-generational data over more than two generations, and the biographical narrative needs to be rebuilt from this jigsaw puzzle of information. To further complicate things, due to the various onomastic approaches of the period under study and the practice of registering names according to the language of the source (which in Transylvania might be Latin, German, Hungarian or Romanian), different individuals bearing the same name, especially very common names, can be easily mistaken for one another. Conversely, the same person registered under different names can be mistaken for multiple persons, or his/her ethnic or confessional identity wrongly interpreted.

Thirdly, the process of biographical reconstruction is not only time-consuming, but quite often involves a mixture of primary and secondary sources, that is, original source-based information and other people's interpretations (inferred or deducted information). Conflicting information about the same person can be found in different sources, and the historian is not always capable of distinguishing which is correct. This raises methodological and hermeneutical issues, as information of varying reliability is pieced together in order to build the profile of a single historical character. In traditional historical research, the accuracy of one source or another is inferred based on the critique of the sources, which in turn relies on the historian's knowledge and experience; but as all of us are aware, the process itself is not always thoroughly documented for the reader. More importantly, source/information that is considered to be false is not always accurately referenced and discussed in the published version, sometimes being plainly omitted, thus leaving open the possibility that others might approach it less critically – especially since entering a piece of information into a database can be incorrectly perceived by lay persons as validating the information as true or accurate.

The above-mentioned aspects were reflected in our own research, initially dedicated to the Romanian political elite of Transylvania, and then to the multipositionality of the members of the region's mid- and upper social strata (what Romanian historiography usually calls the "national elite"). This highlighted the need to build a digital tool capable of storing the wealth and growing variety of information that has surfaced since 2010 as a result of the aforementioned mass digitization process. The result was "Historical Data

Grinder" (HDG),[10] initially developed as a concept by the authors of this paper, and later implemented, tested and perfected by Angela Lumezeanu within her doctoral research project.[11]

In what follows, we shall clarify the circumstances that led us to use this digital tool for another research project, the scientific premises on which we built our research, the impact of the new study's requirements and expectations on HDG, and, in this context, the relationship between HDG and similar models of recording and storing prosopographical historical data.

January 2020 marked the beginning of a research project hosted by the Charles University in Prague in partnership with the Masaryk Institute and Archive of the Czech Academy of Sciences, the aim of which is to study comparatively the political and administrative elites from Bohemia and Transylvania (i.e., from two Habsburg provinces[12] with different historical traditions which nevertheless display a number of similarities), between 1860 and 1920, in terms of demographic behaviour, social mobility, and strategies for integration and advancement in the professional and public space. To this end, the datasets to be compared include vital statistics (e.g., date of birth, marriage and death; children's dates of birth); information on kinship relations, education, and professional and political career (positions and ranks); and the individual's involvement in civil society and economic enterprises (for Members of Parliament elected in Transylvania and Bohemia, as well as for senior civil servants of the period).

In the case of both provinces, the starting point was a series of datasets collected during previous research projects, mainly covering the political and, partially, the professional careers of Transylvanian and Bohemian elites, as well as vital statistics found in the literature (usually the date and place of birth and death). The publication of prosopographical data for political and administrative elites in Bohemia is much more advanced than for Transylvania.[13] However, data

[10] "Historical Data Grinder," accessed 9 September 2021, http://hdgrinder.ro/.

[11] Angela Cristina Lumezeanu, "A Database Model for Social History: Historical Data Grinder and the Transylvanian Society of the 19th and 20th Centuries," *Trans R* 28, no. 2 (Summer 2019): 100-111.

[12] Although Transylvania officially lost its status as a province in 1867, after the union with Hungary, regional particularities and even legal frameworks differentiating it from Hungary proper (e.g. the provisions of the electoral law) continued up to the First World War and beyond.

[13] Selectively, for Bohemia: Josef Tomeš, *Slovník k politickým dějinám Československa 1918-1992* [Dictionary of Political History of Czechoslovakia 1918-1992] (Praha: Pražská edice, 1998); Josef Tomeš et al., ed., *Tváře našich parlamentů 1861-2011. 150 let parlamentarismu v českých zemích* [Faces of our Parliaments from 1861 to 2011. 150 Years of Parliamentarism in the Czech Lands] (Praha: Nakladatelství Lidových novin, 2012); Franz Adlgasser, *Die Mitglieder der österreichischen Zentralparlamente 1848-1918. Konstituierender Reichstag 1848-1849. Reichsrat 1861-1918. Ein biographisches Lexikon*, vol. 1–2 (Wien: Österreichische Akademie der Wissenschaften, 2014); Martin Klečacký, *Slovník představitelů politické správy v Čechách v letech 1849-1918. The Biographical Dictionary of Political Administration Officials in Bohemia in 1849-1918* (Praha:

concerning the social origin of members of the elite and their wives (profession and status of the father, grandfather and father-in-law) was either missing or only partially collected. Moreover, data on other family members, especially the siblings of the research subjects – who were to form the control sample according to the research methodology – were completely missing. The intention was to extract this information from parish registers. In the case of Bohemia, all the extant data is collected in an existing database which was adapted to the needs of the project by making it possible to view data from parish registers alongside information on careers and social involvement. For Transylvania, there is an existing database that could have accommodated the information from the parish registers (HPDT),[14] but its complexity goes beyond the needs of the project and, moreover, its design did not allow for the accommodation of socio-professional information. The solution of using two databases (one for demographic data and one for socio-occupational data) and subsequently consolidating the information had several obvious drawbacks: besides increasing the working time, it would have involved operating with two different tools built on different principles. Moreover, while for Bohemia the collections of parish registers are kept in a relatively complete series and are fully accessible online, in the case of Transylvania, the series of parish registers is often incomplete, the registers are only partially accessible online, and those that are online are solely from one county.[15] Last but not least, due to the unique situation created by the pandemic, physical access to the county archives was impossible for an extended period during 2020, with a short break during the summer.

Due to these circumstances, until the reopening of the archives to the public in the summer of 2021, the extraction of vital statistics on family relationships primarily focused on two categories of source, both incompatible with HPDT: obituaries (funeral announcements)[16] and the major genealogical collections concerning the noble and aristocratic families of Hungary.[17] The latter

Masarykův ústav a Archiv AV ČR, v.v.i. – Národní archiv, 2020). For Transylvania, the main prosopographic work remains Adalbert Toth, *Parteien und Reichstagswahlen in Ungarn 1848-1892* (München: R. Oldenbourg, 1973), but the first complete set of election results was published only recently: Judit Pál et al., ed., *Parliamentary Elections in Eastern Hungary and Transylvania (1865-1918)* (Berlin: Peter Lang, 2018), and the prosopographic data is still being processed.

[14] "Historical Population Database of Transylvania," accessed 7 September 2021, http://hpdt.ro.

[15] Brașov County, but only partially, through the "Portal of the Romanian National Archives," accessed 7 September 2021, http://cautare-bv.arhivelenationale.ro/cautare-bv/detail.aspx?id=32759.

[16] "Hungary Funeral Notices, 1840-1990." Images. FamilySearch, accessed 7 September 2021, http://FamilySearch.org. Original collection: National Széchényi Library, Budapest.

[17] Selectively: Iván Nagy, *Magyarország családai. Czímerekkel és nemzedékrendi táblákkal* [Hungarian Families. With Coats of Arms and Generational Signs], vol. I–XIII (Budapest, 1857-1865); Béla Kempelen, *Magyar nemes családok* [Hungarian Noble Families], vol. I-X (Budapest, 1911-1931); János József Gudenus, *A magyarországi főnemesség XX. századi genealógiája* [The Nobility of Hungary in the 20th Century Genealogy], vol. I-V (Budapest: Heraldika, 1990-1999).

proved particularly useful, as the share of the nobility among Hungarian members of parliament before 1919 was very high.[18]

These premises forced us to opt for storing the information intended for the project – namely, the data on the political and administrative elite of Transylvania – in HDG, from which the datasets necessary for comparative analyses could be extracted. The strongest argument for this was precisely the mosaic-like nature of both the information and the sources. For instance, four main types of sources are used for vital and kinship information alone: parish and civil registers (which in themselves feature a wide structural variety),[19] funeral announcements, genealogical collections, and marriage and death notices published in the press of the time. Parliamentary careers up to 1918 are recorded in a single secondary source which was, however, built upon dozens of primary sources.[20] On the other hand, professional careers, including political positions held outside Parliament, were reconstructed using a multitude of yearbooks of the period, or, depending on the circumstances, biographical or prosopographical secondary literature.[21] Last but not least, the involvement in civil society and the economic initiatives are captured both in a series of publications of the period and by numerous mentions in the press.[22] Therefore, on each member of the elite, dozens (sometimes even hundreds) of pieces of concise information are recorded. These pieces of information can, when concatenated, be used to reconstruct the life and career of the individual in question and, at least in part, those of his or her family members. This type of information – diverse and with fluctuating levels of accuracy – could be described in terms of factoids, a term launched in digital prosopography by John Bradley (see below), but also relevant to current tools in transnational historical demography research, on which we will dwell briefly below.

Digital tools and approaches well suited to research on mid- and upper-level social strata

In 2005, John Bradley and Harold Short drew attention to the benefits of rethinking the way in which information from historical sources is registered into databases for prosopographical purposes. They highlighted, based on three

[18] József Pap, *Parliamentary Representatives and Parliamentary Elections in Hungary (1848-1918)* (Frankfurt am Main: Peter Lang, 2017), 36-53.

[19] Bolovan et al., "Historical Population Database of Transylvania," 11.

[20] Pál et al., ed., *Parliamentary Elections*.

[21] Mainly detailed in ibid., 356-366. The most useful are *Magyarország tiszti cím- és névtára 1884-1918* [List of Civil Servants in Hungary] (Budapest: Arcanum Adatbázis, 2011, Blu-ray).

[22] Most of the press, including yearbooks of various cultural and civil associations, is digitised on "Arcanum Digitheca" and "Biblioteca Digitală BCU Cluj." Of special importance are the financial and banking yearbooks: *Magyar Compass* [Hungarian Compass] (1874-1941, accessed 7 September 2021, https://adt.arcanum.com/ro/collection/MagyarCompass/), respectively *Anuarul Băncilor Române* [Yearbook of the Romanian Banks] (1900-1918, accessed 7 September 2021, http://dspace.bcucluj.ro/handle/123456789/82910).

major prosopographical projects focusing on late antiquity and the medieval period, that the variety, inconsistency and sometimes conflictual character of information from different sources demands the use of a new approach, based on what they called the 'factoid' model (i.e., something that might be so, but there is no absolute certainty about it). Basically, the factoid model implies that "a source 'S', at location 'L' states something ('F') about person 'P'".[23] This enables the breakdown of historical information, usually found in narrative form, into small pieces of data and the relationships between them. Furthermore, if two historical sources provide conflicting information about the same entity/person, both are to be registered into the database, leaving the historian to compare and choose which information is the most reliable. The employment of this model transforms the historical narrative into a structured graph (a network),[24] which can be further transformed into a "proto-narrative", that is, a list of chronologically, or just logically ordered events which form the backbone of the future historical narrative.[25]

Another paper, by Michele Pasin and J. Bradley, building on the experience of over a decade of work with the factoid model, explored the possibility of integrating it with computer ontologies (i.e., RDF and CIDOC-CRM), thus making it available on the Semantic Web. The authors also addressed the issue of interpretation, or rather interpretations, of the primary sources, concluding that a factoid is not only the mere recording of a historical situation following a given ontology, but that it represents in itself an act of interpretation (factoid: Document-Interpretation-Act) and that the factoid approach can and should accommodate distinct perspectives on the same historical event.[26]

As such, the factoid approach can prove most useful when it comes to dealing with prosopographical information (i.e., a large number of persons featured in an even larger number of sources), as it enables consolidation of the original source information within a single tool, following the same logics of structuring, and allowing the user to focus and give thought to the way in which she/he, but also previous historians, relate to the presented information and sources. However, as stated above, one of the main research aims of our current project relates to the demographic patterns of the elite strata under investigation, an aspect to which the factoid model could certainly be adapted (a 'fact' can

[23] John Bradley and Harold Short, "Texts into Databases: The Evolving Field of New-style Prosopography," *Literary and Linguistic Computing* 20, Suppl. Issue (2005): 8, doi:10.1093/llc/fqi022.
[24] Ibid., 10.
[25] Ibid., 21.
[26] Michele Pasin and John Bradley, "Factoid-Based Prosopography and Computer Ontologies: Towards an Integrated Approach," *Digital Scholarship in the Humanities* 30, no. 1 (2015): 86-97. See also "Factoid Prosopography," accessed 25 August 2021, https://www.kcl.ac.uk/factoid-prosopography.

always be a vital event), but for which other, more specific tools have already been built.

Since the 1970s, multiple historical population databases covering different periods stretching from the eighteenth century (sometimes even earlier) to the mid-twentieth century, have been built in Europe and North America.[27] These are source-oriented databases (e.g., censuses, civil registers, parish registers, tax registers) and, therefore, each of them has a different architecture and sometimes even different data-structuring principles. In order to be able to carry out cross-national demographic analyses, an aggregation system called the Intermediate Data Structure (IDS) was developed (and is currently being improved) the aim of which is to allow for information from different databases to be integrated into a common reference and analysis framework. The IDS is based on the structuring principle of entity-attribute-value (EAV) and identifies multiple records on the same person in different spatial, social and temporal contexts according to a predefined set of attributes. Each IDS record is associated with a single person and attribute, and the information from the primary source is then recreated through the relationships between individuals and contexts (INDIV_INDIV, INDIV_CONTEXT, CONTEXT_CONTEXT), to which the chronological marker (TIME STAMP) and the source of information are added. The attributes the individuals are associated with could be various types of vital events (e.g., birth, marriage, death), social ones (e.g. change of residence), or even characteristics (e.g. sex), to which values (e.g. male, female) are assigned, thus enabling the reproduction, in a standardized manner and within a pre-existing classifications, of the information from the original source.[28]

Both the factoid prosopography model and the IDS were created in response to the challenges posed by aggregating, standardizing and analyzing information from multiple sources which, methodologically speaking, is also one of our project's major challenges. From a basic technical perspective, the factoid approach definitely comes closer to the "classical" EAV database structure than most historical databases do, while the creators of IDS explicitly assert the EAV model as a foundation of this instrument.[29] In general, the EAV model is considered the best solution when dealing with inconsistent data and a continuous change of attributes. At the base of the system stands what is known as "row modelling" – a standard component of a database design set. While the row modelling is homogenous with the facts that it describes, the EAV system means that all the information related to an entity is recorded in a single table

[27] A large part of these databases, in their current form, can be accessed via "European Historical Populations Sample Network," accessed 27 August 2021, https://ehps-net.eu/databases.

[28] George Alter and Kees Mandemakers, "The Intermediate Data Structure (IDS) for Longitudinal Historical Microdata, version 4," *Historical Life Course Studies* 1, no. 1 (2014): 1-26, accessed 27 September 2021, https://ehps-net.eu/sites/default/files/hislives_vol_1._23-05-14.pdf.

[29] Ibid., 4 referring the work of W.W. Stead, W.E. Hammond and M.J. Straube in the early 1980s.

across the database containing one column for the entity, one for the attribute and another one for the value. Each attribute of the same entity is recorded on a separate row and is assigned a single value. Thus, a visual projection of an entity's full range of attributes, and their respective values, resembles very much the "proto-narrative" projected by Bradley's factoid approach and is largely similar to the basic IDS structure.

Historical Data Grinder

The two previous examples, together with the MOSAIC project,[30] served as sources of inspiration for HDG, and we have dwelled on them precisely because the principles of approaching the historical sources and the implemented technical model are compatible with the historical realities from which the latter emerged. Since a full technical description of HDG 1.0 has already been published,[31] and one for HDG 2.0 is on the way, the technical aspects are here only be addressed insofar as they facilitate the understanding of the principles under which the database operates, enable its comparison with the previously presented tools, and illustrate the way in which the data for the elites' social mobility research project is structured and can be analyzed according to these principles.

The HDG database generalizes the row-modelling standard using different tables for *Entities, Attributes and Values*, to which *Source* and *Relations* are added. For enhanced structuring, entities from HDG are organized into several classes found in the table *EntityType,* which in time can be extended, detailed and scaled, if required, into an ontology. In the present state of development, HDG stores several entity classes, the most frequently used ones being: Locality, Person, Administrative Unit and Juridical Person. Accordingly, *Entities* can be Individuals, Localities, Institutions, Events, etc. This is one of the main differences between HDG and the factoid model: while currently employed mainly as a prosopographical database, HDG does not focus exclusively on people, starting from the assumption that beings (human and non-human alike), juridical entities, informal collective entities and even places can be the object of historical research. Which of these types of entity the researcher should focus on depends entirely on their interests, since all of them can be assigned *Attributes* and *Values* which recreate part of their past.

Attributes are parameters that define the relationship between the *Entities* and *Values* (e.g., a person – *EntityType*: Individual – can be defined by attributes such as name, profession, status or residence, but also height, weight, eye colour, etc.). *Values* are the qualitative or quantitative characteristics of *Attributes*. Each entity will then have a value for each attribute (e.g., a person

[30] Mikołaj Szołtysek and Siegfried Gruber, "Mosaic: Recovering Surviving Census Records and Reconstructing the Familial History of Europe," *The History of the Family* 21, no. 1 (2016): 38-60.
[31] Lumezeanu, "A Database Model," 100-111.

will have a particular name, but also a particular eye colour).[32] Each of these tables is connected into a metadata table (*Relationship*). Each row represents a single factoid-like record, based on a single relation, which can be of two types: *Entity-Attribute-Value-Source* or *Entity-Entity-Attribute-Value-Source*. Each relation is time stamped, either with a single date or with a start and end date of a period of observation (similar to IDS). Similar to the factoid model, such records can host contradictory or even conflicting information on the same entity, depending on the content of the original source. Their aim is to capture and integrate into the database the information from the source, not to analyse it in any way. However, the user can make annotations about the respective information in a dedicated field ("Remarks"), so that the inevitable hermeneutical process which develops alongside gathering information can also be preserved and shared. Linked to the metadata table is the *Source* table, which contains information regarding the latter: author, title, editors, archival call no., etc.

This simple and flexible system allows the recording and consolidation of information from many typologically and structurally different sources, thus making it ideal for prosopographical inquiries into the bourgeoisie and the "elites" of the modern period. It should be underlined, at this point, that although HDG can accommodate both serial data and piecemeal information, our aim is to fill it with datasets which cover the sources exhaustively. There are two major projects underway at the time of writing: one related to the members of parliament elected before and after 1918 in Transylvania, and the other focused on the members of "Astra" (The Association for the Culture and Literature of the Romanians in Transylvania). Some of the data has already been imported into the database, other subsets are still in the process of being added, and at the same time, manually inputted information is also being added. Other datasets in earlier stages of construction include those covering the management boards and clerks of Romanian banks in Transylvania between 1900 and 1918, and county-level civil servants in Transylvania between the late 1860s and 1918.

The first dataset currently amounts to ca. 10,000 kinship relations between the members of different families who had at least one person as member of the parliament, to which other ca. 2,000 relations are added regarding the electoral activity of the MPs. The second dataset currently amounts to ca. 40,000 instances of membership of Astra between 1900 and 1921 (i.e., over 100,000 relations between members, their relation with the Association, their residence, profession and time of death where recorded). The third and fourth datasets are projected to be even larger, with their contents amounting to

[32] Usually, empty values are not stored into the database, meaning that HDG only registers what is known of an entity (i.e., provided by a source), without highlighting the attributes for which information is lacking. As trivial as this may seem, it is a step further from the classical database design, in which rows in a table for which the sources do not provide information, remain blank (i.e., no value), but are still recorded in the database, taking up space and slowing down the queries.

thousands of individuals and well over 100,000 relations each. Such information, when combined with genealogical and vital registration data (which comes in separate datasets, but it is also inputted manually, due to the frequent occurrence of new information), can enable the reconstruction of clusters and networks to an extent never before aimed at in Romanian historical research. This is particularly beneficial to our current project on social mobility, due to the high degree of multipositionality among members of the political and administrative elite. The Romanian MPs and other members of their families can be found among Astra's members, but also among the management boards of the Romanian credit institutions of the time. Quite a few of them served for longer or shorter periods as county officials, and some returned to civil service after their parliamentary mandate.[33]

A recurring challenge one faces when collecting historical data in Central Europe is the diversity of linguistic forms of both toponyms and anthroponyms. The same place or person may be recorded under different names, varying according to the language of the source. For example, one might not find *Târnăveni* (the current name of the town), but *Dicsőszentmárton, Szentmárton, Dicső Sz. Márton, Diciosânmărtin, D.S. Mărtin, Sankt-Martin, S-Martin, Martinskirch,* etc. Well-known members of the Romanian political elite can be found under at least two onomastic forms: the Romanian one (e.g., *Iuliu Maniu*) and the Hungarian one (e.g., *Maniu Gyula*). The situation is similar for persons of Hungarian or Saxon origin registered in datasets from before and after 1918, as during the dualist period they were usually recorded with their Hungarian or German name, while after 1918 the Romanian translation of their name, or some Romanianized form, became more widely used.

In order to overcome this challenge, in the case of localities, HDG uses a standardization system based on a pre-existing list of localities. Thus, the original name from the source is kept only in the original dataset and, when transferred into the database, the entity takes a standardized form, recording the official names in Romanian, Hungarian and German. In this way, HDG's principles of collecting information largely adhere to the best practices proposed by Kees Mandemakers and Lisa Dillon.[34] The "database of the sources" remains, in fact, the original dataset (available on request to any researcher), while the publicly accessible version contains standardized forms which facilitate both searches by external users and the linkage procedures. Concerning persons' names, primary standardization of the dataset is also carried out prior to its ingestion. However, this standardization is performed only within the linguistic framework of the

[33] Cf. Ovidiu Iudean, *The Romanian Governmental Representatives in the Budapest Parliament (1881-1918)* (Cluj-Napoca: Mega, 2016), 87-238 passim.

[34] Kees Mandemakers and Lisa Dillon, "Best Practices with Large Database on Historical Populations," *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 1, 37 (2004): 34-38.

source and does not involve the translation of the name. Therefore, the same person can be input into the database under linguistically different onomastic variants, depending on the source from which the original dataset was compiled. To cross-linguistically identify the person, manual linkage is used. Hence, the database may in fact contain one standard variant for each language, and all these entities are recorded and displayed as the same person (fig. 1).

Thus, by means of standardization, the factoids related to the same individual are brought together into the "proto-narrative" regardless of the onomastic form under which she/he is recorded (i.e., the language of the source). With the help of the latter, even the most apparently bone-dry sources, such as long and boring membership lists, start telling stories and reveal important information about life events and decisions which are hard to understand and explain if regarded as particular occurrences, but whose wider meaning can be grasped, or at least inferred, as a result of comparing and analysing serial recordings. In a previous study, also dedicated to a former member of the Hungarian Parliament, we have illustrated how digital prosopography and, particularly, HDG can help the historian comprehend life decisions and identify decades-long dormant social ties (e.g. friendships and collegiality relationships established during study years) which were to be reactivated at key moments of the subjects' lives.[35] In what follows, other cases recorded in HDG will be explored, with the aim of emphasizing its usefulness in the research project dedicated to the social mobility of the elites, but also in historical research in general.

Case studies: Parliamentary Representatives from Transylvania as seen through HDG

One of the main features of HDG is that it provides the user with data structured similarly, in many ways, to J. Bradley's factoid "proto-narrative" (figs. 2 and 3) allowing a concise view of the main biographical events of a single person registered through various sources. One such example is the family and the political career of Lőrinc Mara (HDG[36] = 4891 Lőrinc Mara [Id]), son of a high civil servant (HDG = 4892 Lőrinc Mara [legid]), a former officer in the Székely border regiments before the Revolution of 1848-1849, then in the Hungarian revolutionary army of 1848, himself later a civil servant and afterwards a member of the Hungarian parliament. He was followed in the same constituency (Haţeg / Hátszeg), several years later, by his son László Mara (HDG = 4903), who also became a high civil servant later in life (Lord Lieutenant / *főispán*). Alongside this three-generation line of high officials and deputies, the database records information about their families, studies and, to a certain extent, about their

[35] Vlad Popovici and Rada Varga, "Building Life Courses and Explaining Life Choices with the Help of Digital Prosopography," *Studia dig.* 63, no. 2 (2018): 59-63.

[36] As biographical information is already stored and referenced in HDG (cf. fig. 3) whenever possible from this point onward, we will indicate only the HDG IDs of the selected case studies.

professional careers before they reached their peak political and administrative positions. Cases of possible erroneous information are recorded according to the source, but underlined as such (e.g., fig. 3, row 9, the office of Lőrinc Mara as county commissioner / *alispán,* recorded by the prosopographical work of Gábor Bona on the captains of the Hungarian Revolutionary Army,[37] has been marked by Judit Pál as possibily erroneous and requiring further verification.[38]

This example represents a "classical" translation of an HDG "proto-narrative" into a narrative biographical medallion. The reconstruction begins with information collected primarily by focusing on individuals: it refers to a family of the political and administrative elite of nineteenth-century Transylvania, concerning whom, although no civil registers were preserved in their native locality, a wealth of information enables easy reconstruction of their genealogical relationships, their biographies and their transgenerational social mobility (or in this case, rather, their social immobility). However, as we have already mentioned, HDG equally allows for research focused on institutions or the relationship between individuals and institutions.

One such institution is Astra, a cultural society joined by the majority of the Romanian political leaders partly out of a genuine attachment to the ideas and values promoted by the society and partly for the sake of public image, networking and prestige among their fellow Romanians. In order to be registered as a member of Astra, it was necessary to provide information concerning one's residence and profession. Thus, the annually published membership lists serve as valuable sources in identifying professional and geographical mobility, as well as members of the family.[39] Obviously, as these documents are digitized and open access, such research could in theory be carried out without the help of a database. However, given the high quantity of information to be processed (between 1,500 and 2,000 members annually and a fluctuation in membership that could reach 20-25% per year in some departments), an attempt to analyse these sources coherently and exhaustively in the absence of a dedicated electronic tool and its associated procedures (standardization and linkage) would be almost impossible.

The majority of the Romanian nationalist MPs that were elected to the Budapest Parliament between 1900 and 1918 were members of Astra throughout this period and continued to remain so into the early 1920s. Among the Romanian moderate ("governmental") MPs, membership levels were lower, even

[37] Gábor Bona, *Századosok az 1848/49. évi szabadságharcban* [Captains of the War of Independence from 1848/49] (Budapest: Heraldika, 2008-2009), digital edition on "Arcanum Digitheca."

[38] The authors would like to thank Professor Judit Pál for verifying the biographical data and annotating the possibly erroneous information.

[39] The lists were published initially on a monthly basis, and after 1900 on a yearly basis in Astra's journal, *Transilvania*, which is digitised at "Biblioteca Digitală BCU Cluj," last accessed 8 September 2021, http://dspace.bcucluj.ro/jspui/handle/123456789/7435.

among those whose profession was closely related to cultural activities. For example, Ioan Ciocan, (HDG = 5448), a governmental deputy and holder of the Chair of Romanian Language at the University of Budapest, became a member of Astra only in 1905, most probably in connection with the dissolution of the governmental (until then) Liberal Party and his rapprochement with the leaders of the Romanian national movement.[40] Since the group of the Romanian MPs in Hungary was quite small (less than 30 people between 1901 and 1918), and most of those surviving the war had well-known political and professional careers in the Kingdom of Romania, Astra's membership lists prove to be less useful in their cases, as they contain information which had already entered the historiographical circuit. Occasionally, these documents can reveal their children's career paths. Such are the cases of the lawyer Caius Brediceanu and the secretary, then bank manager, Tiberius Brediceanu, both the sons of the MP Coriolan Brediceanu (HDG = 5435). Interestingly, Caius Brediceanu's (HDG = 6965) entry into Astra (1910) took place immediately after his father's death (1909), illustrating a common practice, namely the preservation/transmission of membership among family members.[41]

Very few deputies' wives became members of Astra and the memberships of those who did were of short duration, lasting no more than a few years. This situation has to do with the general state of Romanian women's associationism in pre-war Hungary, but also with the orientation of the female public towards the other great cultural society of the time: the Society for the Romanian Theatre Fund. For instance, Nicolae Comșa, a physician and MP from Sălișteа (HDG = 5450), was the only Astra member in his family between 1902 and 1919. In contrast, on the membership lists of the Society for the Romanian Theatre Fund, one encounters no fewer than seven individual members of his family – out of which three were women – and two institutional members closely related to the family: the Romanian Casino in Sălișteа (the masculine pole) and the Reunion for the Embellishment of the Church in Sălișteа (the feminine pole).[42]

Astra's membership lists prove to be more useful for inquiries into the deputies elected to the post-1919 Romanian legislatures, as they manage to capture the professional and social mobility brought about by the regime change. Sometimes, even when an individual did not change their main profession, the very decision to join Astra implied a change in status, income and presumably self-perception – an investment in the regional prestige economy of the time. Aurel Socol (HDG = 6966), a lawyer from Bandu de Câmpie, son of a judge, finally managed to graduate from the Faculty of Law of the University of Cluj in 1908, after several interrupted semesters at the Universities of Budapest (1898/99),

[40] Iudean, *The Romanian Governmental Representatives*, 109-114, esp. 113.
[41] *Transilvania* 40, no. 4, (1909): 223; ibid., 41, no. 4 (1910): 221.
[42] *Anuarul Societății pentru Fond de Teatru Român* 4 (1901): 212-220.

Prague (1902/03) and possibly Berlin.[43] Although information on his activity prior to 1919 is scarce, we do know that he served as President of the Romanian Casino in Cluj, which means that he enjoyed a high status and prestige within the local community. In spring 1919, he became a member of the new administrative and judicial institutions of the County of Cluj.[44] In the same year, he was elected to the Romanian Parliament as deputy for the urban constituency of Cluj,[45] later pursuing a political career as a Member of Parliament and State Secretary.[46] It is no coincidence, given his newly acquired parliamentary mandate, that he first appears in the membership list of the Cluj department of Astra in the same year. Until his death in 1937, Aurel Socol had a rich parliamentary and political career and was succeeded among the local party leadership by his son, Aurel Socol Jr, who is remembered mainly for his role in helping numerous groups of Jews to cross the frontier from Hungary into Romania in 1944.[47]

Reconstructions similar to the one above (which could have been greatly extended if we were to present the entire political activity of A. Socol) are currently impossible to undertake solely on the basis of the information accommodated by HDG. This is largely due to the fact that most of the data hosted by HDG resulted from the ingestion of large serial sources, which usually capture a very well-defined, even narrow, aspect of a person's biography. Thus, for now, each life course needs to be complemented with manually inserted information from biographical or prosopographical literature, with the aim of enriching the person's biography and generating connections with other biographies already recorded in the database, leading to further extensions. In this way, HDG has developed as a product of the combination of serial and individual sources and sometimes even ego-documents, and the quantitative analyses on extended samples can always be complemented by delving into the biographical narrative, a format much more familiar to historians.

Cross-checking information from multiple sources and comparing serial sources with narrative-biographical ones inevitably leads to the identification of errors and the possibility of correcting them. Astra's membership lists, like any other historical source, may contain questionable information. For example, Emil Andreșan (HDG = 6968), a landowner from Băila, was listed as member of Astra

[43] Sigmirean, *Istoria formării intelectualității*, 375, 480, 696.
[44] Augustin Tataru, "† Dr Aurel Socol," *Gazeta Ilustrată* 6, no. 9-10 (September - October 1937): 121.
[45] Vlad Popovici, "Cojocna," in Bogdan Murgescu and Andrei Florin Sora, ed., *România Mare votează. Alegerile parlamentare din 1919 "la firul ierbii"* [Greater Romania Is Voting. The 1919 Parliamentary Elections "at the Edge of the Grass" (Jassy: Polirom, 2019), 180-181.
[46] Tataru, "† Dr Aurel Socol," 121.
[47] Aurel Socol, *Furtună deasupra Ardealului* [Storm over Transylvania] (Cluj-Napoca: Tribuna, 1991).

from 1911 to 1913, followed by Anastasia Andreșan, wife of a priest (HDG = 6969), between 1913 and 1914. After 1918, neither of them was still a member of the society. Given the lack of correspondence between their professions (landowner vs priest's wife), one may think that the two were not a couple, but rather, at most, members of the same extended family. However, Emil was not only a landlord, but also a priest (an omission in the original source) and held a doctoral degree in law;[48] he was elected Senator to the Parliament of Romania in the constituency of Reghin, Mureș County, in 1920. Anastasia was in fact his wife, with whom he had ten children[49] and their names appearing consecutively in the membership lists indicates again (as in the above-mentioned case of Caius and Coriolan Brediceanu) the transfer of membership within the family, interrupted, in this case, most likely by the outbreak of war. Although the erroneous information provided by the original source has been corrected with the help of biographical medallions, the database still preserves it, albeit with an explicit mention attesting its partial validity (fig. 4).

Conclusions

The process of massive digitization of primary sources paves the way for a more detailed and in-depth understanding of historical events and background relations than would have been possible just a decade ago. At the same time, the information available from digitized narrative sources is continuously growing, and it will certainly surpass the possibility of the human historian to fully harvest it directly. This will, most probably in less than a decade, lead to the use of automated extraction software (i.e., crawlers, ML software)[50] trained to recognize, extract and arrange the data in large datasets. While manual harvesting of data remains the most reliable solution at present, things are already changing and not even large research teams will be able to cope with the research requirements and standards heightened by the entry into use of historical crawlers.

However, this avalanche of information, dissipated among a plethora of typologically different sources, requires adequate (i.e., ordered, critically evaluated, contextualized, linked and analysed) tools to control it. Therefore, in the case of prosopographical research, the use of databases remains for the moment indispensable. Whether source-oriented or event-history-oriented, they involve reorganizing historical information in relation to the structure of the original source. Breaking historical data into small-scale time-stamped relations between entities (humans, non-humans, institutions, etc.), all linked back to the

[48] Sigmirean, *Istoria formării intelectualității*, 380, 460.

[49] Traian Bosoancă, *Mureșenii și Marea Unire* [People from Mureș County and the Great Union] (Târgu Mureș: Ardealul, 2000), 17-20.

[50] E.g., Bruno Rodrigues, "Historical Newspaper Scraping with {Tesseract} and R", accessed 7 September 2021, https://www.brodrigues.co/blog/2019-04-07-historical_newspaper_scraping_tesseract/.

digitized source, represents, in our opinion, the next mandatory step for any research projects aiming at biographical reconstructions, large-scale prosopographies, and the more sociologically oriented research attempting to analyse the nineteenth- to mid-twentieth-century bourgeoisie in its entirety. Due to the simplicity of their structure, reliability and ease of verification, EAV-structured tools such as HDG could play an essential part in this process, both during the testing and training of future crawlers, and as digital repositories of historical data.

By transferring the digitized narratives in the sources into a structured database, which allows automated verification, linkage and comparisons, and enables the researcher to approach the data as "factoids" (whether or not they choose to employ Bradley's original model) rather than given historical facts, historians should be able to improve the selection of the prosopographical samples in view of further analyses, keep track of conflicting information provided by different sources, and revisit any piece of data when required.

Particularly in the case of the project dedicated to the social mobility of the elites, the specificity of the sources, at least for Transylvania (numerous, yet fragmented and uneven in terms of the value and accuracy of the information) lends itself very well to the use of a digital tool such as HDG, which enables the recording of both vital statistics and family and institutional relationships. The data network thus created facilitates biographical reconstruction, but at the same time can be interrogated with respect to demographic aspects (e.g., life-course analysis, marriage strategies in a social context) or socio-political behaviours (e.g. social origin, professional career path, social networks) of larger elite samples.

At this point in the development of HDG, the major challenge remains extracting information from the primary sources and, in particular, corroborating information from serial sources with that from the individual or family biographical sources. However, once a sufficiently large data pool has been created, this will generate, as we have shown in the previous section, sufficient relationships (and the research questions related to them) in order to determine and even force the continuous gathering of new information necessary for obtaining answers. Clearly, HDG cannot function as a perpetuum mobile, yet it is able to generate, thanks to the information it collects, enough scientific energy to justify its long-term operation, collaborations, constant extension, and maintenance at a level of research close to international standards.

Acknowledgement
This study was supported by the project of the Czech Science Foundation, no. 20-19463X, "Social mobility of elites in the Central European regions (1861–1926) and transition of imperial experience and structures in nation-states".

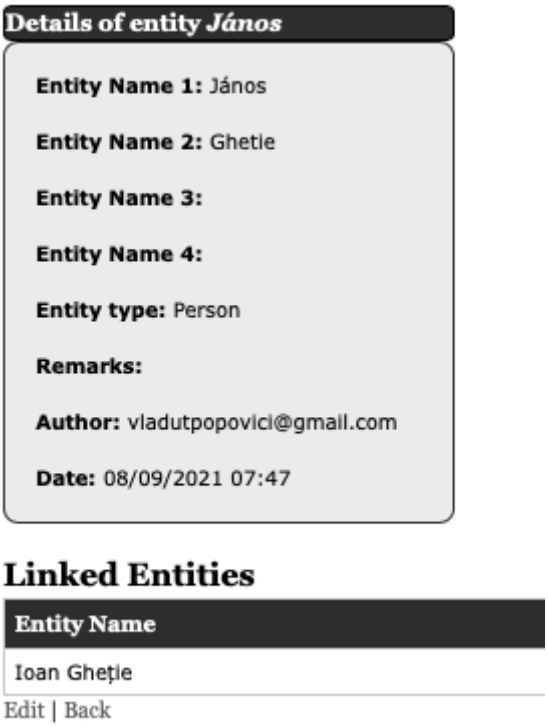

Fig. 1. Record linkage of the same person / entity registered in HDG according to different sources, under different names, in different languages.

Entity 1	Entity 2	Attribute	Value (Standard)	Value (English)	Start Observation period	End Observation period	Julian Calendar	Source	Source page	Remarks
4891 Lőrinc Mara [id]	2384 Székelyföldvár Războieni-Cetate	inregistrare vitală	naștere	birth	1823//	//	false	N7 1860	287-288	
4891 Lőrinc Mara [id]	2384 Székelyföldvár Războieni-Cetate	inregistrare vitală	înmormântare	burial	1893//	//	false	FN Mara Lőrinc		
4891 Lőrinc Mara [id]	1190 Felsőszállápatak Sălașu de Sus	inregistrare vitală	deces	death	1893//	//	false	FN Mara Lőrinc		
4891 Lőrinc Mara [id]		statut nobiliar	nobil	nobleman	//	//	false	N7 1860	287-288	felsőszállápataki
4892 Lőrinc Mara [legid]	4891 Lőrinc Mara [id]	rudenie	tată	father	1823/10/21	//	false	N7 1860	287-288	Birthdate from: https://www.arcanum.hu/hu/online-kiadvanyok/Bona-bona-tabornokok-torzstisztek-1/szazadosok-az-184849-evi-szabadsagharcban-96F2/eletrajzi-adatok-989D/m-9EE9/felszoszallaspataki-mara-lorinc-9F0C/
4893 Kata Dániel Mara Lőrincné	4891 Lőrinc Mara [id]	rudenie	mamă	mother	1823/10/21	//	false	N7 1860	287-288	Birthdate from https://www.arcanum.hu/hu/online-kiadvanyok/Bona-bona-tabornokok-torzstisztek-1/szazadosok-az-184849-evi-szabadsagharcban-96F2/eletrajzi-adatok-989D/m-9EE9/felszoszallaspataki-mara-lorinc-9F0C/
4891 Lőrinc Mara [id]	4892 Lőrinc Mara [legid]	rudenie	fiu	son	1823/10/21	//	false	N7 1860	287-288	Birthdate from https://www.arcanum.hu/hu/online-kiadvanyok/Bona-bona-tabornokok-torzstisztek-1/szazadosok-az-184849-evi-szabadsagharcban-96F2/eletrajzi-adatok-989D/m-9EE9/felszoszallaspataki-mara-lorinc-9F0C/
4891 Lőrinc Mara [id]	4893 Kata Dániel Mara Lőrincné	rudenie	fiu	son	1823/10/21	//	false	N7 1860	287-288	Birthdate from https://www.arcanum.hu/hu/online-kiadvanyok/Bona-bona-tabornokok-torzstisztek-1/szazadosok-az-184849-evi-szabadsagharcban-96F2/eletrajzi-adatok-989D/m-9EE9/felszoszallaspataki-mara-lorinc-9F0C/
4892 Lőrinc Mara [legid]	4893 Kata Dániel Mara Lőrincné	rudenie	soț	husband	//	//	false	N7 1860	287-288	
4893 Kata Dániel Mara Lőrincné	4892 Lőrinc Mara [legid]	rudenie	soție	wife	//	//	false	N7 1860	287-288	
4891 Lőrinc Mara [id]	4894 László Mara	rudenie	nepoată (de fiu sau fică)	granddaughter	//	//	false	N7 1860	287-288	
4891 Lőrinc Mara [id]	4895 Kata Dániel [id] Mara Lászlóné	rudenie	nepot (de fiu sau fică)	grandson	//	//	false	N7 1860	287-288	
4894 László Mara	4891 Lőrinc Mara [id]	rudenie	bunic	grandfather	//	//	false	N7 1860	287-288	
4895 Kata Dániel [id] Mara Lászlóné	4891 Lőrinc Mara [id]	rudenie	bunică	grandmother	//	//	false	N7 1860	287-288	
4894 László Mara	4895 Kata Dániel [id] Mara Lászlóné	rudenie	soț	husband	//	//	false	N7 1860	287-288	
4895 Kata Dániel [id] Mara Lászlóné	4894 László Mara	rudenie	soție	wife	//	//	false	N7 1860	287-288	
4891 Lőrinc Mara [id]	4902 Lőrinc Mara [ifj]	rudenie	tată	father	//	1893//	false	FN Mara Lőrinc		
4902 Lőrinc Mara [ifj]	4891 Lőrinc Mara [id]	rudenie	fiu	son	//	1893//	false	FN Mara Lőrinc		
4891 Lőrinc Mara [id]	4904 Berta Zalandak Mara Lőrincné [ifj]	rudenie	socru	father in law	//	//	false	FN Mara Lőrinc		
4904 Berta Zalandak Mara Lőrincné [ifj]	4891 Lőrinc Mara [id]	rudenie	soră	daughter in law	//	//	false	FN Mara Lőrinc		
4902 Lőrinc Mara [ifj]	4904 Berta Zalandak Mara Lőrincné [ifj]	rudenie	soț	husband	//	//	false	FN Mara Lőrinc		
4904 Berta Zalandak Mara Lőrincné [ifj]	4902 Lőrinc Mara [ifj]	rudenie	soție	wife	//	//	false	FN Mara Lőrinc		
4891 Lőrinc Mara [id]	4903 László Mara	rudenie	tată	father	//	//	false	FN Mara Lőrinc		
4903 László Mara	4891 Lőrinc Mara [id]	rudenie	fiu	son	//	//	false	FN Mara Lőrinc		
4891 Lőrinc Mara [id]	4905 Lujza Barcsay Mara Lászlóné	rudenie	socru	father in law	//	//	false	FN Mara Lőrinc		

Fig. 2. Proto-narrative of Lőrinc Mara sr. in HDG (a).

4906 Leontin Lukács Mara Lőrincné [id]	4891 Lőrinc Mara [id]	rudenie	soție	wife	//	//	false	Bona-1848Századosok		https://www.arcanum.hu/hu/online-kiadvanyok/Bona-bona-tabornokok-torzstisztek-1/szazadosok-az-184849-evi-szabadsagharcban-96F2/eletrajzi-adatok-989D/m-9EE9/felszoszallaspataki-mara-lorinc-9F0C/
4891 Lőrinc Mara [id]	4907 k.u.k. Technische Militärakademie Imperial and Royal Technical Military Academy bécsi hadmérnöki akadémiá	student	generic	generic	1839//	1844//	false	Bona-1848Századosok		https://www.arcanum.hu/hu/online-kiadvanyok/Bona-bona-tabornokok-torzstisztek-1/szazadosok-az-184849-evi-szabadsagharcban-96F2/eletrajzi-adatok-989D/m-9EE9/felszoszallaspataki-mara-lorinc-9F0C/
4891 Lőrinc Mara [id]		religie sau confesiune	reformat	Calvinist	//	//	false	FN Mara Lőrinc		
4891 Lőrinc Mara [id]	4911 A Hunyad megyei Honvédegylet	membru	generic	generic	//	//	false	Bona-1848Századosok		https://www.arcanum.hu/hu/online-kiadvanyok/Bona-bona-tabornokok-torzstisztek-1/szazadosok-az-184849-evi-szabadsagharcban-96F2/eletrajzi-adatok-989D/m-9EE9/felszoszallaspataki-mara-lorinc-9F0C/
4891 Lőrinc Mara [id]	4910 Prison (generic)	profesie	prizonier	prisoner	1849//	1853//	false	Bona-1848Századosok		https://www.arcanum.hu/hu/online-kiadvanyok/Bona-bona-tabornokok-torzstisztek-1/szazadosok-az-184849-evi-szabadsagharcban-96F2/eletrajzi-adatok-989D/m-9EE9/felszoszallaspataki-mara-lorinc-9F0C/
4891 Lőrinc Mara [id]	4908 K.u.k. Dragonerregiment Nr. 5 5th Dragoon Regiment	profesie	cadet	cadet	1844//	1846//	false	Bona-1848Századosok		https://www.arcanum.hu/hu/online-kiadvanyok/Bona-bona-tabornokok-torzstisztek-1/szazadosok-az-184849-evi-szabadsagharcban-96F2/eletrajzi-adatok-989D/m-9EE9/felszoszallaspataki-mara-lorinc-9F0C/
4891 Lőrinc Mara [id]	4909 1. Szekler Grenz-Infanterieregiment Nr. 14 14th (1st Székely) Border Guards Regiment	profesie	sublocotenent	Junior Lieutenant	1846//	1848//	false	Bona-1848Századosok		https://www.arcanum.hu/hu/online-kiadvanyok/Bona-bona-tabornokok-torzstisztek-1/szazadosok-az-184849-evi-szabadsagharcban-96F2/eletrajzi-adatok-989D/m-9EE9/felszoszallaspataki-mara-lorinc-9F0C/
4891 Lőrinc Mara [id]		profesie	pretor	sheriff	1861//	1868//	false	Bona-1848Századosok		https://www.arcanum.hu/hu/online-kiadvanyok/Bona-bona-tabornokok-torzstisztek-1/szazadosok-az-184849-evi-szabadsagharcban-96F2/eletrajzi-adatok-989D/m-9EE9/felszoszallaspataki-mara-lorinc-9F0C/
4891 Lőrinc Mara [id]	1006 Hátszeg Haţeg	profesie	vicecomite	county commissioner	1867//	//	false	Bona-1848Századosok		https://www.arcanum.hu/hu/online-kiadvanyok/Bona-bona-tabornokok-torzstisztek-1/szazadosok-az-184849-evi-szabadsagharcban-96F2/eletrajzi-adatok-989D/m-9EE9/felszoszallaspataki-mara-lorinc-9F0C/ [probabil eronat - JP]
4891 Lőrinc Mara [id]	1006 Hátszeg Haţeg	profesie	deputat	representative (MP)	1875//	1878//	false	Pál et alii 2018	80	
4891 Lőrinc Mara [id]	4912 Hátszeg Haţeg	profesie	deputat	representative (MP)	1878//	1881//	false	Pál et alii 2018	124	
4891 Lőrinc Mara [id]	4912 Hátszeg Haţeg	profesie	deputat	representative (MP)	1881//	1884//	false	Pál et alii 2018	124	
4891 Lőrinc Mara [id]	4912 Hátszeg Haţeg	profesie	deputat	representative (MP)	1884//	1887//	false	Pál et alii 2018	124	
4903 László Mara	4912 Hátszeg Haţeg	profesie	deputat	representative (MP)	1896//	1899/7/12	false	Pál et alii 2018	124	
4891 Lőrinc Mara [id]	4897 Károly Mara	rudenie	frate	brother	//	//	false	FN Mara Lőrinc		
4891 Lőrinc Mara [id]	4896 Miklós Mara	rudenie	frate	brother	//	//	false	FN Mara Lőrinc		
4891 Lőrinc Mara [id]	4899 György Mara	rudenie	frate	brother	//	//	false	N7 1860	287-288	
4891 Lőrinc Mara [id]	4900 Ágnes Mara Kemény Istvánné	rudenie	frate	brother	//	//	false	FN Mara Lőrinc		
4900 Ágnes Mara Kemény Istvánné	4891 Lőrinc Mara [id]	rudenie	soră	sister	//	//	false	FN Mara Lőrinc		
4891 Lőrinc Mara [id]		literatură	personaj	character in a story	//	//	false	MK 68		https://www.arcanum.hu/hu/online-kiadvanyok/Mikszath-mikszath-osszes-muve-2A85B/cikkek-es-karcolatok-5186-kotet-40CD5/1883-parlamenti-karcolatok-68-kotet-4B833/a-t-hazbol-marc-9-4BBA9/iv-a-mara-lorinc-emberei-4BBD8/
4903 László Mara	4914 Hunyad megye Hunedoara (comitat)	profesie	comite suprem	Lord Lieutenant	1904//	1906//	false	N7 1860		
4891 Lőrinc Mara [id]	4913 1848 honvédség armata revoluționară maghiară de la 1848	profesie	căpitan	captain	1849//	1849//	false	Bona-1848Századosok		https://www.arcanum.hu/hu/online-kiadvanyok/Bona-bona-tabornokok-torzstisztek-1/szazadosok-az-184849-evi-szabadsagharcban-96F2/eletrajzi-adatok-989D/m-9EE9/felszoszallaspataki-mara-lorinc-9F0C/
4903 László Mara	4914 Hunyad megye Hunedoara (comitat)	profesie	comite suprem	Lord Lieutenant	1910//	1917//	false	N7 1860		

Fig. 3. Proto-narrative of Lőrinc Mara sr. in HDG (b).

			ID	Entity 1	Entity 2	Attribute	Value (Standard)	Value (English)	Start Observation period	End Observation period	Julian Calendar	Source	Source page	Remarks
Show	Edit	Destroy	1199	6968 Emil Andreșan	1 Asociațiunea pentru Literatura și Cultura Poporului Român Astra	membru	membru ordinar	ordinary member	1911//	1913//	false	Transilvania 1911-4	518	See also Transilvania 43, 1912, 5, p. 337; Transilvania, 44, 1913, 4-5, p. 233.
Show	Edit	Destroy	1200	6969 Anastasia Andreșan	1 Asociațiunea pentru Literatura și Cultura Poporului Român Astra	membru	membru ordinar	ordinary member	1913//	1914//	false	Transilvania 1913-4-5	233	See also Transilvania 45, 1914, 7-9, p. 364.
Show	Edit	Destroy	9927	6968 Emil Andreșan	6969 Anastasia Andreșan	rudenie	soț	husband	1900//	//	false	Bosoancă 2000	17-20	Observation start period: before 1900. Being ordained as priest in 1900, he should have been married at the time
Show	Edit	Destroy	9928	6968 Emil Andreșan		profesie	Proprietar	Owner	1911//	1913//	false	Transilvania 1911-4	518	See also Transilvania 43, 1912, 5, p. 337; Transilvania, 44, 1913, 4-5, p. 233. Main profession: priest, see Bosoancă 2000, 17-20.
Show	Edit	Destroy	9929	6968 Emil Andreșan		profesie	Preot	Priest	1900//	1945//	false	Bosoancă 2000	17-20	

Fig. 4. Biographical data about Emil Andreșan in HDG.