Developing the ontological model for research and representation of Commemoration Speeches in Croatia using a graph property database

Benedikt Perak

Faculty of Humanities and Social Sciences, University of Rijeka

Abstract

This paper demonstrates the ontological model of commemoration speeches given in Croatia that was developed as a part of the project "Framing the Nation and Collective Identity in Croatia: Political Rituals and the Cultural Memory of Twentieth Century Traumas". The research gathers linguistic, media and social data about seven commemoration practices connected to the wars of the 20th century using various methodologies involving audio-visual field recordings, transcription of commemoration speeches and the creation of a text corpus. This chapter presents digital humanities methods used to connect various levels of data analysis and digital resources, from natural language processing (NLP) of Croatian to conceptual enrichment through the processing of conceptual metaphors. These resources are integrated using the Neo4j database and the Python Py2Neo library for data manipulation of the ontological model that embeds the linguistic and media research in the context respectively of the social identity of actors, their interaction, institutional affiliations and cultural models they represent and express. The value of this ontological model is in fostering an interdisciplinary approach through the contextualization of data and targeted usage of digital resources.

Key words

Cultural memory, commemoration, textual analysis, monument representation, social ontology, complex system, graph database, Neo4j

Introduction

This paper shows the methods and resources used for the digitization of commemoration speeches given in Croatia and the development of an ontological model for the retrieval and analysis of the complex socio-cultural data within the project "Framing the Nation and Collective Identity in Croatia: Political Rituals and the Cultural Memory of Twentieth Century Traumas "(FRAMNAT)[1]. The FRAMNAT project seeks to develop digital humanities methodologies applicable for research on cultural memory and socio-cognitive linguistics analysis. The research involves gathering linguistic, media and social data about commemoration practices connected to the traumatic events and atrocities of the 20th century wars: Jasenovac, Bleiburg, Brezovica and Jazovka, Srb, Knin and Vukovar. The paper presents the procedures of audio-visual and textual data gathering as well the digital humanities methods used to store and enrich various levels of data analysis using the property graph database model and resources for natural language processing of Croatian, as well as data integration and information enrichment.

The motivation for the development of digital methods and tools to study state remembrance practices is to institute an empirical method of cultural memory research as well as to offer resources and methods to other researchers of cultural memory in the region and beyond. The systematic analysis, from fieldwork at the sites of memory to studying the social networks and the role of media in transmitting narratives, provides an insight into the construction of a society's "collective memory." Although commemorations and political speeches are not the only communication strategies used in constructing a story of the recent past, commemorations provide a highly visible stage for political elites and other memory actors to re-perform and conceptualize the past and define their political agendas within that frame. We identified seven commemorations that were relevant because they either attracted the country's political leadership and were of national significance, or were particularly controversial and therefore provoked debates that would reveal how various actors framed the nation through rival "truths" about the past. Five commemorations are related to the Second World War: Bleiburg, Brezovica, Jazovka, Jasenovac and Srb. The two other commemorations represent both the victim and victor narratives of the Croatian War of Independence (referred to as the Homeland War, or *Domovinski rat*, in Croatia): Vukovar and Knin. During the course of 3 years (2014-2017) the members of the FramNat research team gathered audio-visual evidence of the public speeches delivered at the 7 commemorations. The speeches were transcribed, tokenized, morphosyntactically tagged, lemmatized and syntactically parsed using the Reldi service[2] and then published as a searchable corpus. By studying the data provided by the analyses of the corpus, we offer conclusions about the conceptual representation and cultural distribution of collective (national) identities in discourse and in official narratives of the past. The main questions that are addressed include the following: what salient concepts and mappings are applied by the speakers to frame conceptualizations of the nation and national identity? What are the differences in framing the nation and national identity between different political actors, institutions and options? Who are the individuals, organizations and institutions that produce dominant models of representing cultural memory, how and why? One of the objectives of this chapter is to provide insights into methods and processes that were involved in gathering and analyzing information on the memory agents and their role in commemorative speeches. The second section of the chapter discusses the methods and resources used for the gathering and storing the data, while the third section demonstrates some of the analytic uses of the data for reseach in the conceptual framing methods.

Gathering the data about the commemorations

The project FRAMNAT gathered empirical data about seven commemoration practices connected to the wars of the 20th century: Jasenovac, Bleiburg, Brezovica and Jazovka, Srb, Knin and Vukovar using various methodologies involving audio-visual field recordings, transcription of commemoration speeches and creation of a text corpus.

Socio-cultural data about the speakers in the commemorations

Most of the commemorations involve a public speech delivered by the speaker, who acts as a memory agent. The speaker's role is to conceptualize the historical traumatic event by captivating the attention and raising the motivation of the listeners, while at the same time providing reasoning patterns and establishing the culturally normative values[3]. The conceptualization is performed by the speech delivered by a speaker and addressed primarily to the assembled audience at the commemoration site, and secondarily to the wider national audience through the media coverage. Most of the speakers are connected with some institution or organization, such as the Croatian Parliament, President or Government, as well as other social organizations such as Catholic, Orthodox church, Anti-fascist organizations, veteran organizations etc., that partake in the political agenda of the commemoration. The network of the sixty-four speakers who delivered speeches at seven commemoration sites from 2014-2016. is represented in Illustration 1.

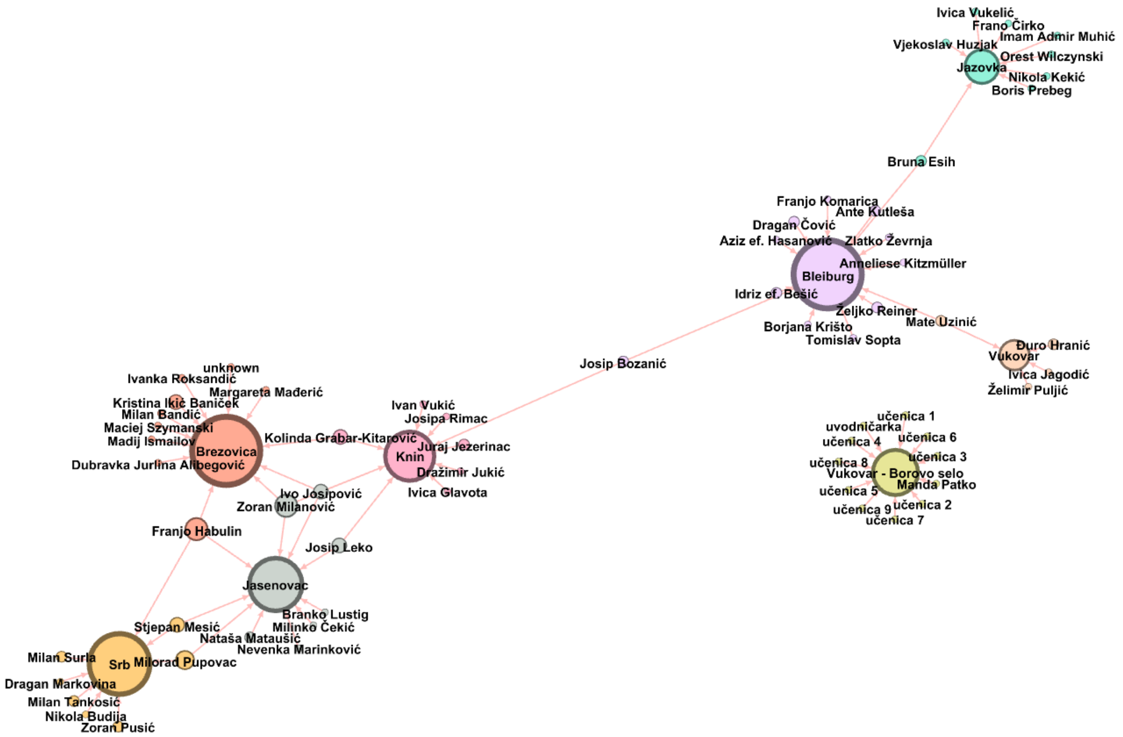


Illustration 1: Network representation of the speakers at each commemoration. The size of the nodes is represented relative to the amount of connections with other nodes (degree).

The layout of the graph is produced by connecting a speaker to the commemoration site where the speech was delivered. The majority of the speakers have delivered speeches at only one commemoration site, but some of them, mostly high ranking political officials, have appeared at several commemorations, such as the former president Ivo Josipović, who delivered speeches in Knin, Brezovica, and Jasenovac, as did former Prime Minister Zoran Milanović. Kolinda Grabar-Kitarović, elected president in January 2015, appeared as a speaker in Knin and Brezovica. Cardinal Josip Bozanić and other members of the Catholic Church also appeared at several commemorations including Knin, Vukovar, and Bleiburg. The following three networks represent the presence of the speakers on a yearly basis from the year 2014 (illustration 2), 2015 (illustration 3), to 2016 (illustration 4).

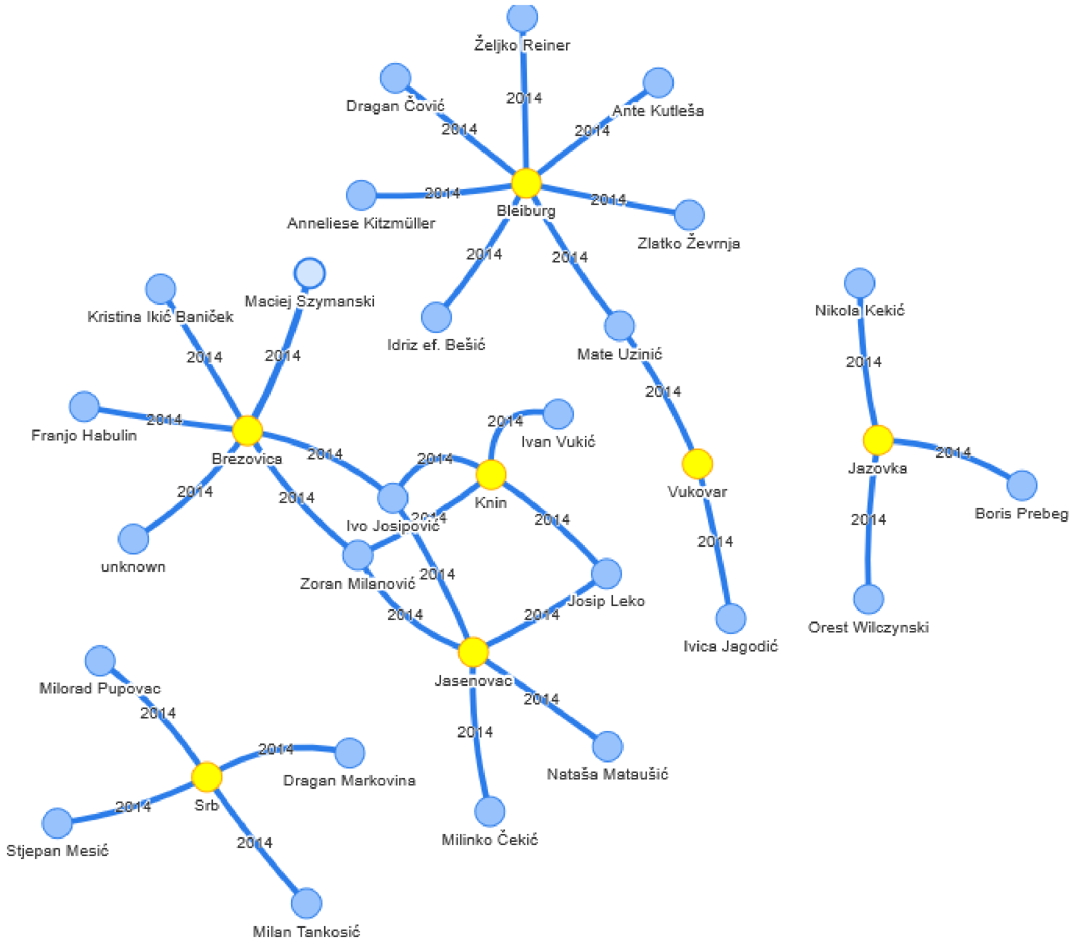


Illustration 2: Network representation of the speaker’s attendance at commemorations during the year 2014.

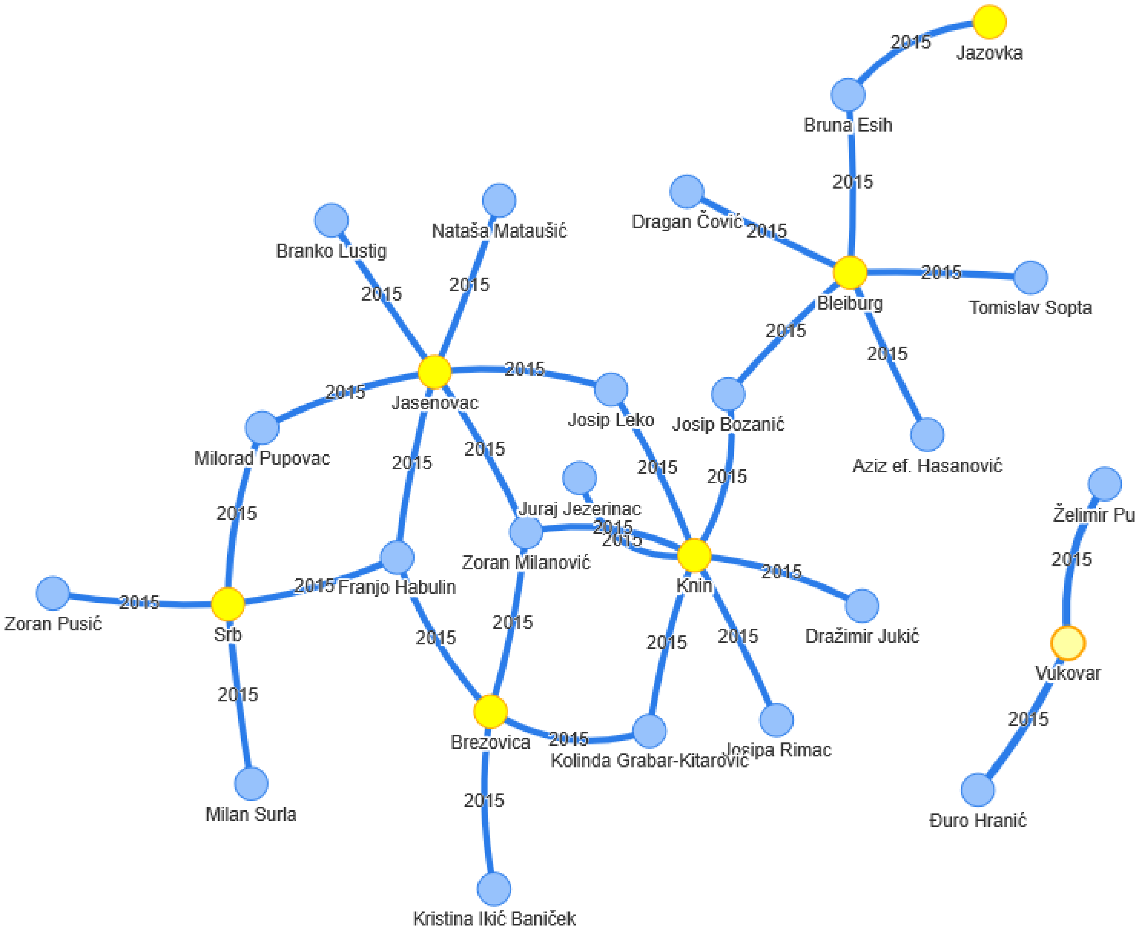


Illustration 3: Network representation of the speaker’s attendance at commemorations during the year 2015.

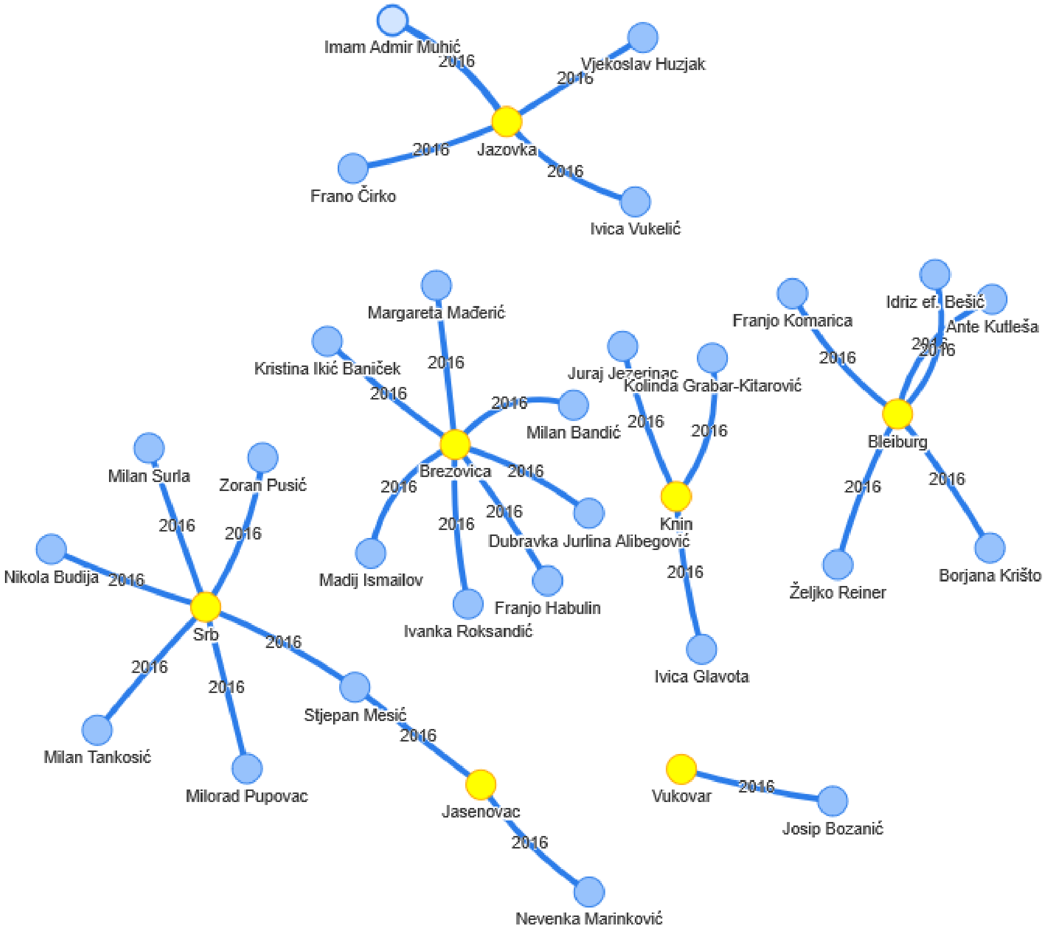
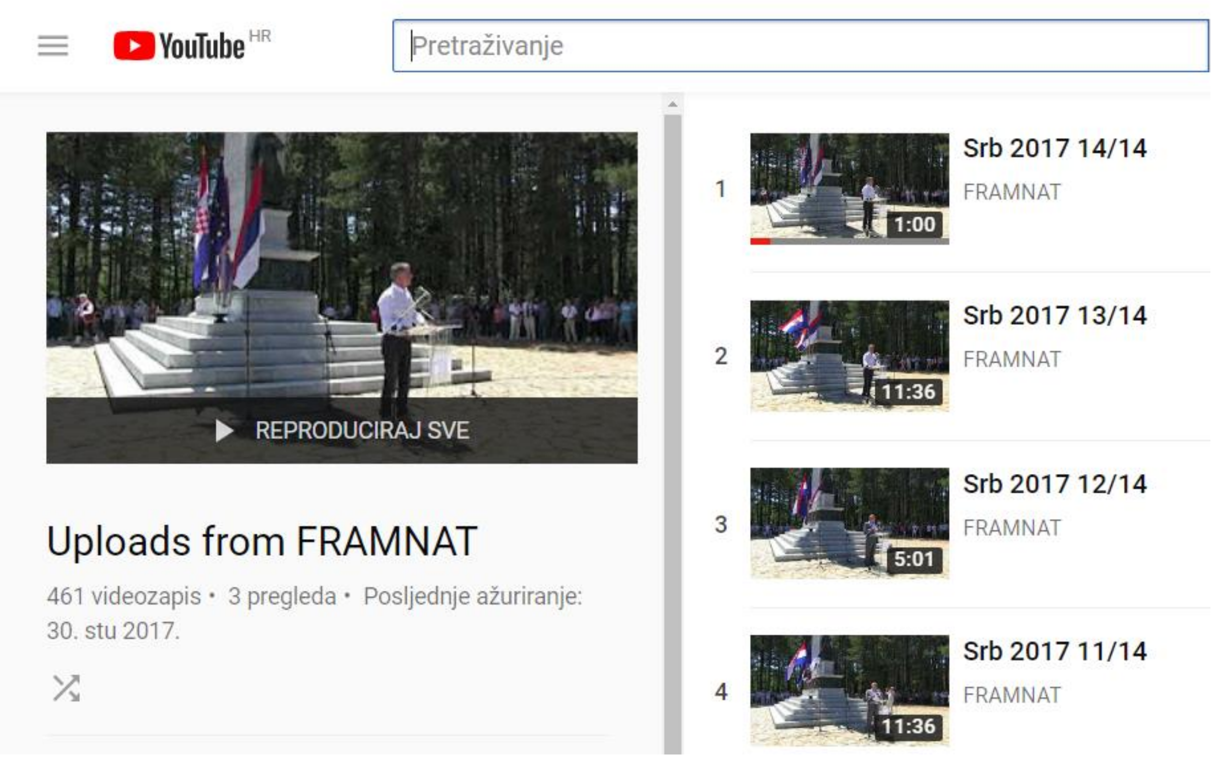


Illustration 4: Network representation of the speaker’s attendance at commemorations during the year 2016.

Audio-visual material

The primary method of the data gathering was related to the fieldwork observation and audio-visual recordings of the commemorations. Every speech and related commemoration ritual have been audio-visually recorded by researchers of the FRAMNAT project. The recordings were edited and stored on a local hard drive as well as published on the FRAMNAT Youtube channel[4].



Uploads from FRAMNAT

461 videozapis • 3 pregleda • Posljednje ažuriranje: 30. stu 2017.

1 Srb 2017 14/14 FRAMNAT 1:00

2 Srb 2017 13/14 FRAMNAT 11:36

3 Srb 2017 12/14 FRAMNAT 5:01

4 Srb 2017 11/14 FRAMNAT 11:36

Illustration 5. Screenshot of the FRAMNAT YouTube channel.

The audio-video material allows for the multimodal research of commemoration speeches and related rituals such as wreath laying ceremonies, clerical rituals and processions. The published material is organized and tagged according to the year and the commemoration.

Transcriptions of the speeches and corpus creation

The audio-visual data was the basis for the transcription of the texts. The transcription has been performed by researchers and collaborators on the FRAMNAT project: Iva Davorija, Renato Stanković and Mirna Gurdon. In order to allow for consistent text analysis, the texts have been stored as .txt files along with their unique metadata filename: the name of the speaker, date of the speech and the name of the commemoration. For instance, the speech delivered by Bozanić in Knin in the year 2014 is stored as “bozanic_kn_2014.txt” file. The texts of this FRAMNAT corpus are stored locally in the Neo4j property database and published on the SketchEngine cloud service[5].

FramNat corpus on SketchEngine

The SketchEngine service is used for the text storage, tokenization of text, morphosyntactic tagging, lemmatization and parsing of the texts using the IHJJ sketch grammar for Croatian[6]. The FRAMNAT corpus is published on the SketchEngine service at the following address: https://the.sketchengine.co.uk/auth/corpus/140810/search.

Tagged FramNat corpus stored in a Neo4j graph database

A locally stored FramNat Corpus for additional ontological analysis of the texts is created using a custom-developed software application for storing linguistically tokenized, lemmatized and syntactically parsed digital texts of the Croatian language using the Reldi application service[7], py2neo Python library[8] and graph property database Neo4j[9]. The application uses a pipeline of several automated processes that comprise: 1) ingesting the texts as data, 2) tokenizing, lemmatizing and parsing the texts, and 3) storing multiple texts and tokenized, lemmatized and parsed morphosyntactic information in the graph database. The application pipeline is presented in Illustration 6.

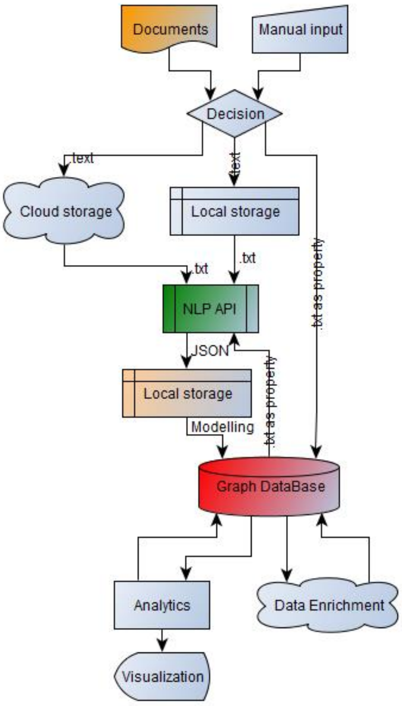


Illustration 6. The pipeline for creating tokenized, lemmatized and syntactically parsed corpus using the Reldi Api, Neo4j graph database and Py2Neo application.

In the first phase, texts are processed in the form of .txt raw data and sent to the Reldi application programming interface (https://github.com/clarinsi/reldi-lib-doc). In the second step, the Reldi service parses the text data and sends back the lexical, morphosyntactic and syntactic values about tokens, lemmas and dependency structures in the form of a JSON data file. Next, the Python py2neo application converts the JSON data values stored on a local drive to a Neo4j database using a custom-made linguistic schema model. The model extracts the JSON key: value structure to represent each text as an entity with the properties, such as the name of the text, creation date, link to the resource, etc. The schema of the model stores the linguistic structures of tokens, words, lemmas as entities while the grammatical and syntactical relations are stored as connections between those entities. Each text is therefore decomposed into corresponding tokens, words and lemmas according to the described linguistic schema (Illustration 7).

Data integration

The data about the audio-visual resources, texts and metadata are published on the Google Sheet: https://docs.google.com/spreadsheets/d/1rXV9x9-Jdpw84nmcOTEJBHnd-S5nu7-YDYk8zj06sN8/edit?usp=sharing.

This datasheet represents the structure of the information for each commemoration speech, including the name of the commemoration, description, available information about the commemoration on Wikipedia, available information about the historical reference on the commemoration topic, place of the commemoration, GPS coordinates of the commemoration site, year of the commemoration, speaker name, speaker party affiliation, Wikipedia resources on the speaker, additional resources on the speaker from media, picture, gender, function, affiliated institution, institution information, text name, text, estimated number of attendees, video link, and order of speech in the commemoration event. The various data points are integrated using an ontological model that embeds the commemoration metadata, and textual and morphosyntactic data with the resources for media research on the social context. The enriched information is stored in the graph property database Neo4j and connected via the ontological model shown in Illustration 7. The graph type of database is used to represent connected data. The graph representation of the commemorations communicates the structure of connectedness between data points at different phenomenological levels and enables the quantitative and qualitative exploration of the correlational structure and their construals.

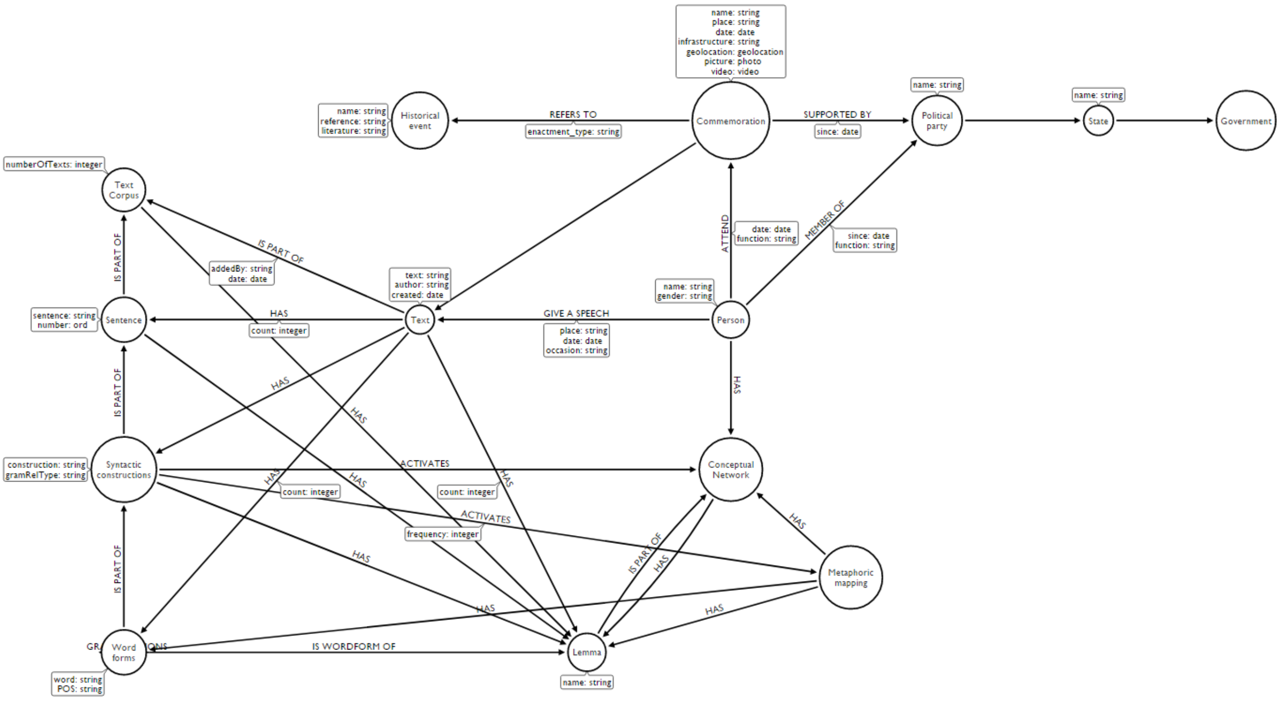


Illustration 7. Ontological model and Database Schema of Commemorative Speech Analysis

The nodes represent the structurally different categories of entities and their directional relations to component categories. For instance, a text has sentences, and syntactic constructions are part of sentences. Word forms are part of lemmas, while lemmas are part of texts. Furthermore, each node and relations have some properties that are expressed in the key: value format. For instance, the Person category has properties name and gender. Person is connected with the category Commemoration with a relation ATTEND that has properties: name, place, date, geolocation, picture and video. The ontologically different levels of the digital phenomena of commemorations are stored in a single database that enables complex queries within a certain category or between categories. The directed graph property structure supports the complex analysis of the social identity of actors, their interactions and institutional affiliations, and the cultural models they express in their speeches.

Analysis of the data

The analytic system comprises of queries that are formulated using the Cypher native programming language for Neo4j database[10] and Py2Neo Python library. The results can be represented as tabular sets of data or in a network. The basic use of the analytic tools is related to various types of summarization and statistics. For instance, the thirty most frequent lemmas (basic word forms) or entity concepts in the FramNat corpora 2014-2016 can be identified with the following query:

```
MATCH p=(t:Texts)-[r:HAS_lemma]->(l:Lemmas)
WHERE l.lempos ENDS WITH “-n”
WITH l.lempos as Lemma,r.lemmaCountInFile as count
RETURN lema, sum(count) as Sum
ORDER BY Sum DESC
LIMIT 30
```

The output of this query is represented in a Table 1.

Table 1. Thirty most frequent noun lexical concepts in the FRAMNAT 2014-2016 corpus.

Lemma (lexical concept)	Translation	Sum
Hrvatska	Croatia	486
narod	People	323
godina	Year	322
čovjek	Man	308

žrtva	Victim	269
dan	Day	226
rat	War	219
život	Life	195
država	State	188
istina	Truth	182
mjesto	Place	174
zločin	Crime	157
sloboda	Freedom	154
borba	Struggle	150
domovina	Homeland	144
branitelj	Defender	137
put	Path	135
zlo	Evil	122
grad	City	121
povijest	History	121
mir	Peace	102
zemlja	Land	100
republika	Republic	99
ime	Name	97
Hrvat	Croat	97
riječ	Word	95
kraj	End	93
Vukovar	Vukovar	90
vrijeme	Time	89
fašizam	Fascism	87

Furthermore, if we want to extract similarities between two speakers on the level of the conceptual use we can formulate a query that looks for the overlapping lexical concepts (lemmas). The following Cypher query displays these differences between texts produced by Ivica Glavota and Dražimir Jukić

```
MATCH p=(s1:Speakers{name:"Zoran Milanović"})-[d1:DELIVERED_SPEECH]->(t1:Texts)-
[l1:HAS_lemma]->(lema:Lemmas)<-[l2:HAS_lemma]-(t2:Texts)<-[d2:DELIVERED_SPEECH]-
(s2:Speakers{name:"Želimir Puljić"})
WHERE lema.lempos ENDS WITH "-n"
WITH distinct(lema.lempos) AS lemm
RETURN lemm
```

The query returns following nouns: *mjesto* “place”; *naš* “our”; *budućnost* “future“; *branitelj* “defender“; *izraz* “expression“; *vjera* “faith“; *zajedništvo* “community“; *dijete* “child“; *dobro* “good“; *osjećaj* “feeling“; *vlast* “government“; *oko* “eye“; *srce* “heart“; *republika* “republic“; *Hrvatska* “Croatia“; *predsjednik* “president“; *rat* “war“; *gospodin* “mr.“; *dan* “day“; *sloboda* “freedom“; *poštovanje* “respect“; *čovjek* “man“; *godina* “year“; *vojska* “army“; *Europa* “Europe“; *svijet* “world“; *građanin* “citizen“; *zemlja* “land“; *dio* “part“; *obitelj* “family“; *mir* “peace“; *život* “life“; *država* “state“; *ime* “name“; *ponos* “pride“; *put* “path“; *prostor* “space“; *otpor* “resistance“; *strana* “side“; *kraj* “end“; *prijatelj* “friend“; *događaj* “event“; *Hrvat* “Croat“; *vrijeme* “time“; *tisuća* “thousand“; *želja* “wish“; *narod* “people“; *mjesec* “moon“; *čelo* “forehead“; *trenutak* “moment“; *sila* “force“; *kultura* “culture“; *početak* “start“; *ruka* “hand“; *riječ* “word“; *sestra* “sister“; *zajednica* “community“; *poziv* “call“; *pad* “fall“; *način* “means“; *svećenik* “priest“; *dom* “home“; *temelj* “ground“; *čast* “honour“; *sredstvo* “means“; *Radić* “Radić“; *pozdrav* “salute“; *inozemstvo* “foreign country“; *tijelo* “body“; *rodbina* “family“.

Due to the similarity between lexical concepts expressed by all speakers we can calculate the proximity of different speakers and represent it by means of the speaker community graph. This community identification method can be used for discerning the Cultural Model of conceptualization related to a particular Speaker (Illustration 8).

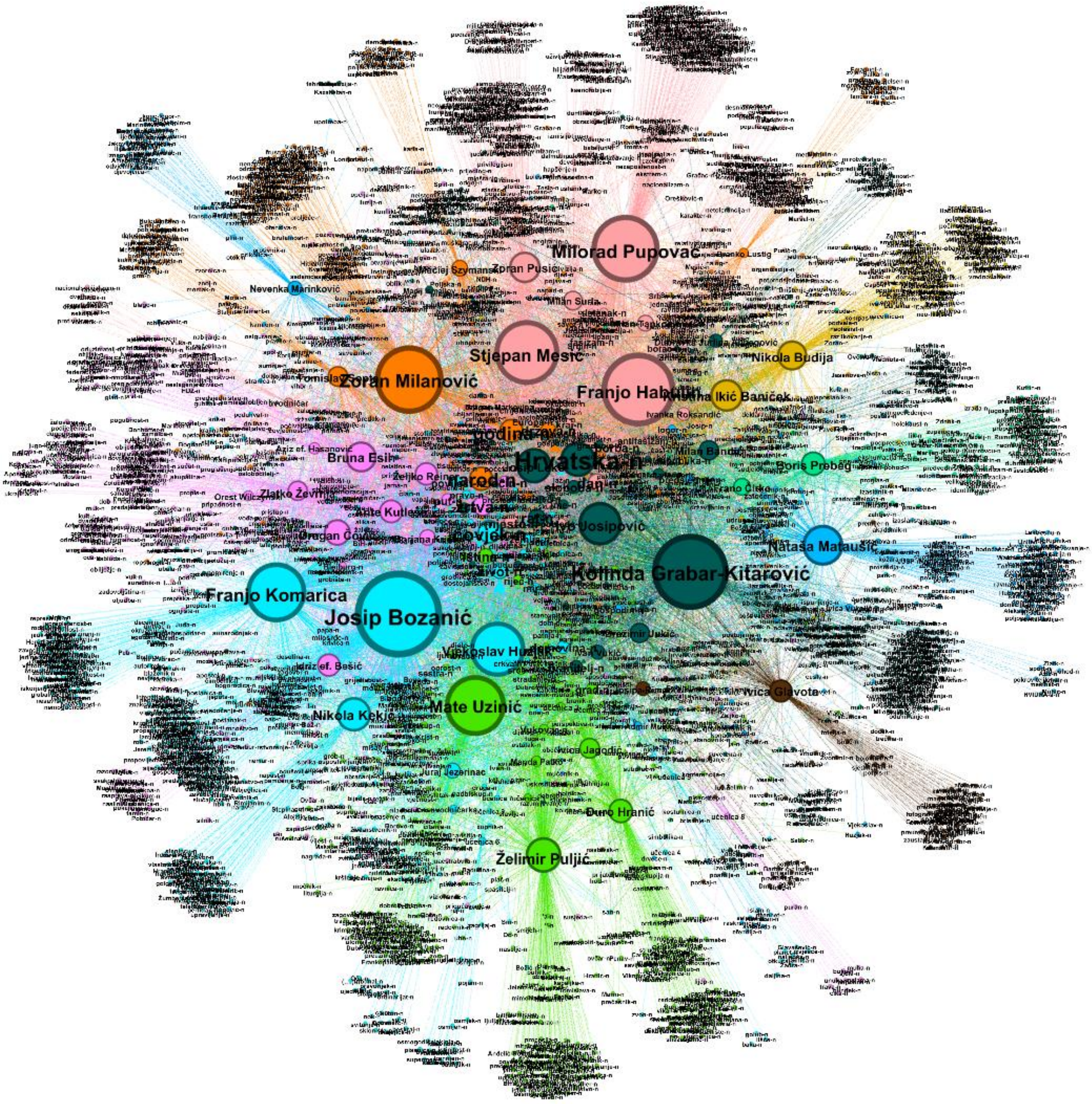


Illustration 8: The graph of relationships between the 3370 noun lemmas expressed by the 64 speakers. The size of the labels corresponds to the overall frequency of the nouns connected with the speaker.

The Louvain algorithm for detecting communities[11] was applied to the network represented in Illustration 8 and distinguishes ten communities of speakers. The communities, clustered according to the similarity of the nouns they used in their speeches, are shown in Table 2.

Table 2. Ten communities of the speakers clustered according to the similarity of the nouns used in their speeches.

Community	Speakers	% of the network activation
1	Josip Bozanić, Franjo Komarica, Vjekoslav Huzjak, Nikola Kekić, Juraj Jezerinac, student 6, student 5, student 7, student 9, student 4	17,6 %
2	Franjo Habulin, Milorad Pupovac, Stjepan Mesić, Zoran Pusić, Milan Surla, Milan Tankosić, Ivanka Roksandić, Dragan Markovina	18,9 %
3	Mate Uzinić, Želimir Puljić, Đuro Hranić, Ivica Jagodić, Manda Patko, student 3, student 1, student 2	12,6 %
4	Bruna Esih, Dragan Čović, Ante Kutleša, Željko Reiner, Idriz ef. Bešić, Zlatko Ževrnja, Borjana Krišto, Aziz ef. Hasanović, Orest Wilczynski, student 8	11,7 %
5	Kolinda Grabar-Kitarović, Ivo Josipović, Josip Leko, Milan Bandić, Dražimir Jukić, Ivan Vukić, Dubravka Jurlina Alibegović, Margareta Mađerić, Madij Ismailov	11,3 %
6	Kristina Ikić Baniček, Nikola Budija	5, 1 %
7	Zoran Milanović, Tomislav Sopta, Maciej Szymanski, Branko Lustig	8,2 %
8	Boris Prebeg, Frano Čirko	3,7 %
9	Ivica Glavota, Josipa Rimac	3,4 %
10	Nataša Mataušić, Nevenka Marinković, Imam Admir Muhić, Ivica Vukelić	7,5 %

Lastly, we can perform similar types of queries by connecting the Speaker entities with the affiliated Institutions. In this way we get the representation of the influence of the institutions in the overall corpus (Illustration 9)

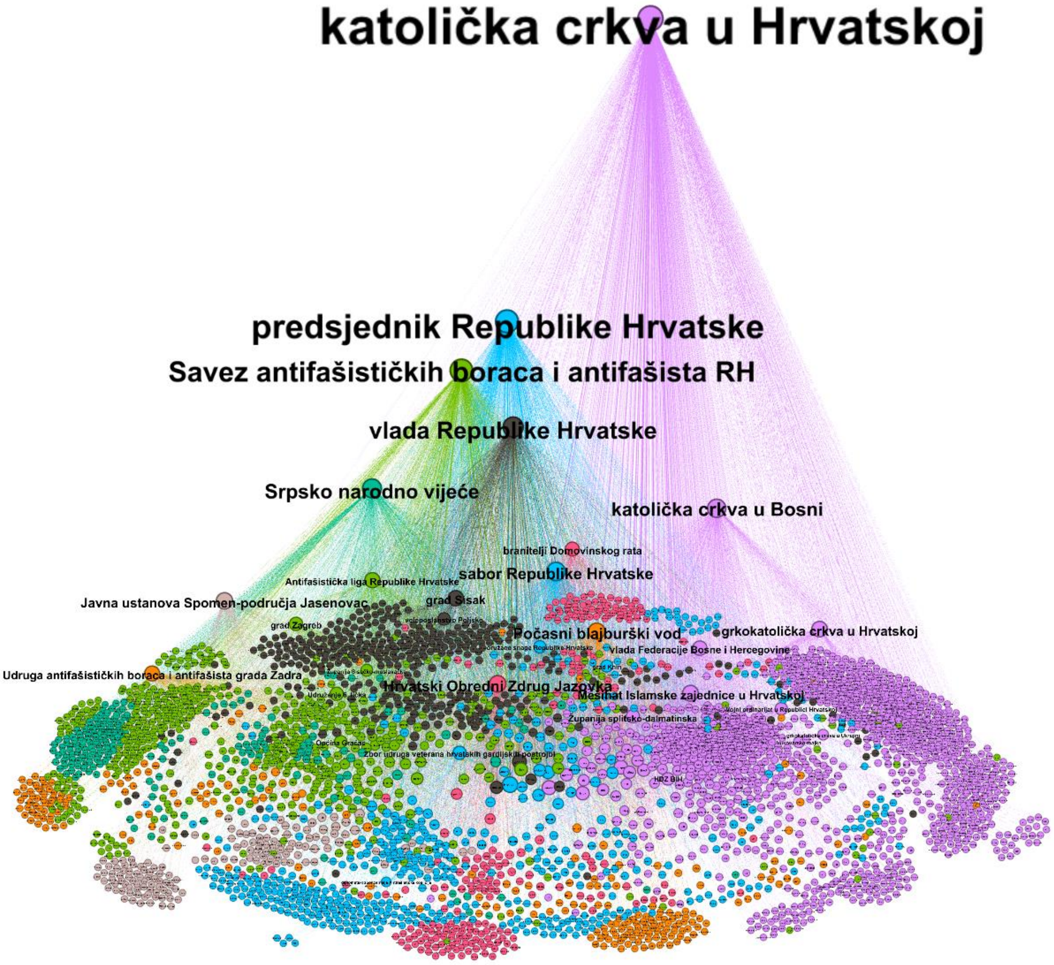

katolička crkva u Hrvatskoj
predsjednik Republike Hrvatske
Savez antifašističkih boraca i antifašista RH
vlada Republike Hrvatske
Srpsko narodno vijeće
katolička crkva u Bosni
branitelji Domovinskog rata
Antifašistička liga Republike Hrvatske
sabor Republike Hrvatske
Javna ustanova Spomen-područja Jasenovac
grad Sisak
grad Zagreb
veleposlanstvo Poljske
Počasni blajburški vod
grkokatolička crkva u Hrvatskoj
oružane snage Republike Hrvatske
vlada Federacije Bosne i Hercegovine
grad Knin
Udruga antifašističkih boraca i antifašista grada Zadra
Županija Sisačko-moslavačka
Hrvatski Obredni Zdrug Jazovka
Mešihat Islamske zajednice u Hrvatskoj
Udruženje 6. lička
Vojni ordinarijat u Republici Hrvatskoj
Županija splitsko-dalmatinska
grkokatolička crkva u Ukrajini
Općina Gračac
Zbor udruga veterana hrvatskih gardijskih postrojbi
HDZ BiH

Illustration 9. The graph of the relationships between the 3370 noun Lemmas expressed by the representatives of 31 Institutions. The size of the labels and nodes corresponds to the overall frequency of the nouns connected with a particular Institution. The graph is projected in 3 ordinates: x, y and z. The projected z ordinate, perceived as the height of the institution nodes, corresponds to the number of different words connected in the graph.

The graph representation can be interactively explored at the FRAMNAT web site. It is important to note the connectedness and the structure between the institutions and concepts, represented by the Force Layout with z ordinate in Illustration 9. The nouns commonly used by many representatives of the institutions are located in the oval centre of the graph due to the many connections with different representatives, while the nouns specific and unique to a certain institution extend to the margins.

The data integration allows for the refinement of the various dimensions of the exploration within the structure of the dataset. For instance, the previous graph (illustration 9) states that the Catholic Church produced most of the concepts in the dataset. By using the graph structure we can query where were these speeches performed?

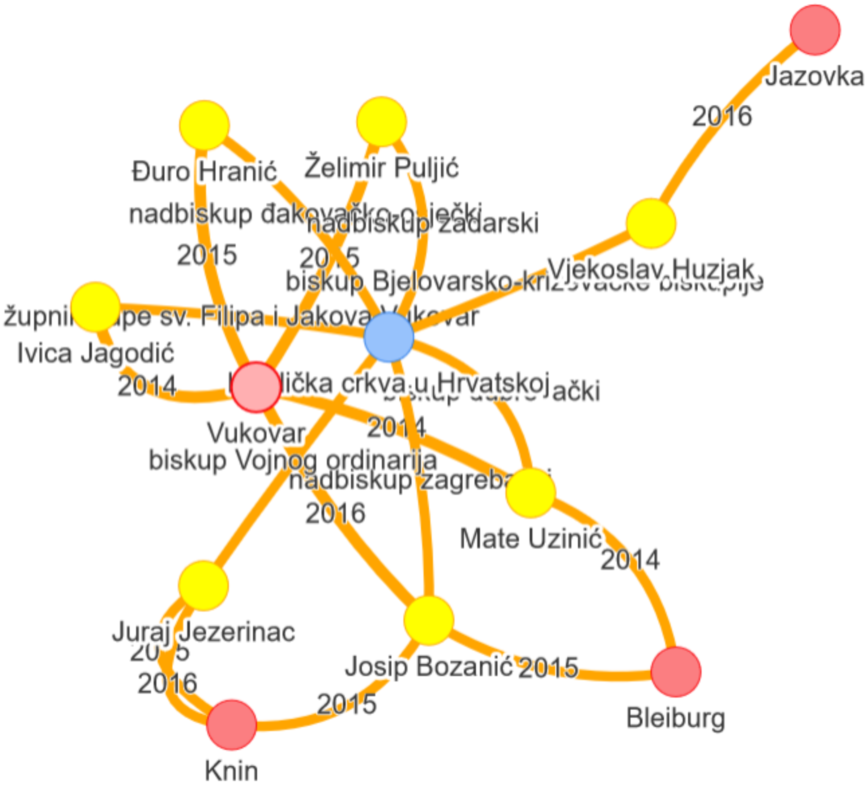


Illustration 10: Network representation of the speaker’s attendance affiliated with the Catholic Church in Croatia.

The insight of these two graphs is that the institution Catholic Church was active in the Jazovka and Bleiburg commemorations of the Second World War atrocities, as well as in the Vukovar and Knin commemorations of the Croatian War of Independence.

Another type of the conceptual analysis is related to the construal of the relevant entities. For instance, the “dobj” type of syntactic-semantic construction reveals what the object is of some process. So, we can formulate a query that reveals what is the conceptualization of some entities when construed as a direct object. In other words, we ask, what you can do with some entity? Here is the query for the noun *domovina* or “homeland” construed as a direct object:

```
MATCH (a:Lemmas{lempos:"domovina-n"})-[r:HAS_dependency{function:"dobj"}]-(b:Lemmas)
WHERE b.lempos ENDS WITH "-v"
WITH max(r.countDep) AS freq, a.lempos AS directObject, b.lempos AS process, r
RETURN  process, directObject, r.function AS dependencyType, freq ORDER BY freq DESC
```

The result of the query is represented in Table 3.

Table 3. The processes that conceptualize the noun *domovina* or “homeland” as a direct object:

Process	directObject	dependencyType	Freq
imati “have”	domovina “homeland”	dobj	4
voljeti “love”	domovina “homeland”	dobj	3
obilježiti “mark”	domovina “homeland”	dobj	1
štititi “protect”	domovina “homeland”	dobj	1
uzeti “take”	domovina “homeland”	dobj	1
voditi “lead”	domovina “homeland”	dobj	1
graditi “build”	domovina “homeland”	dobj	1
biti “be”	domovina “homeland”	dobj	1
poštovati “respect”	domovina “homeland”	dobj	1

The problem of construal of the abstract social concept HOMELAND is related to the metonymic and metaphoric conceptual analysis of the cognitive processes[12]. In this case, the underlying metaphorical mappings are produced with the construal of HOMELAND as AN OBJECT THAT SOCIAL AGENT POSSESSES; A THING THAT PSYCHOLOGICAL AGENT LOVES; A VALUABLE GOOD TO BE PROTECTED|RESPECTED, AN OBJECT THAT CAN BE TAKEN, AN OBJECT THAT CAN BE LED, A THING THAT CAN BE BUILT.

Conclusion

This chapter deals with the analysis of commemoration practices from the perspective of the digitization and data integration of various phenomenological levels of analysis. The main goal of this interdisciplinary research is to understand the process of construing the culturally distributed cognition and conceptualizations by means of public communication acts and speeches. We have identified the speakers as the social agents in promoting immediate conceptual and gradual cultural dissemination. The content of their message is framed by the salient concepts from a cultural model, or worldview, that speakers share by institutional affiliation. The message is analyzed as text consisting of tokens, words and lemmas, applying NLP methods to the Croatian language. The value of this multilevel ontological model is in fostering an interdisciplinary approach through contextualization of data and targeted usage of digital resources. The contextualization of the data enables holistic insight into the dynamics of cultural memory practices and their political implications for contemporary culture. The usage of the digital methods allows for a fine grained quantitative analysis that, due to the graph property organization of the model, continues to be qualitatively expressive, flexible and non-reductive.

Bibliography

Ljubešić, Nikola; Klubička, Filip; Agić, Željko; Jazbec, Ivo-Pavao. „New Inflectional Lexicons and Training Corpora for Improved Morphosyntactic Annotation of Croatian and Serbian" In: *Proceedings* *of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*. Portorož, Slovenija: European Language Resources Association (ELRA), 4264-4270, 2016.

Pavlaković, Vjeran, and Perak Benedikt. "How Does This Monument Make You Feel? Measuring Emotional Responses to War Memorials in Croatia.", in: *The Twentieth Century in European Memory: Transcultural Mediation and Reception*, eds. Plewa, Barbara Törnquist, and Tea Sindbæk Andersen, Brill, http://www.brill.com/products/book/twentieth-century-european-memory-2017. www.cost.eu/module/download/62629, 2017.

Štrkalj Despot, K., Brdar, M., Essert, M., Tonković, M., Perak, B., Ostroški Anić, A., Nahod, B., Pandžić, I. *MetaNet.HR – Croatian Metaphor Repository*. http://ihjj.hr/metafore/, 2015.

Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, Etienne Lefebvre „Fast unfolding of communities in large networks“, *Journal of Statistical Mechanics: Theory and Experiment***,** 10, 2008.

[1] The FRAMNAT project is a four-year project, financed by the Croatian Science Foundation under the number HRZZ-3782. This paper was supported by this project.

[2] Ljubešić, Nikola; Klubička, Filip; Agić, Željko; Jazbec, Ivo-Pavao. „New Inflectional Lexicons and Training Corpora for Improved Morphosyntactic Annotation of Croatian and Serbian // Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)". Portorož, Slovenija: European Language Resources Association (ELRA), 4264-4270, 2016.

[3] Pavlaković, Vjeran, and Perak Benedikt. "How Does This Monument Make You Feel? Measuring Emotional Responses to War Memorials in Croatia.", in: *The Twentieth Century in European Memory: Transcultural Mediation and Reception*, eds. Plewa, Barbara Törnquist, and Tea Sindbæk Andersen, Brill, http://www.brill.com/products/book/twentieth-century-european-memory-2017.

[4] FRAMNAT Youtube channel: https://www.youtube.com/channel/UCjarsad6jWsKPi4Z7CdK5Wg.

[5] The SketchEngine cloud service: https://the.sketchengine.co.uk.

[6] The IHJJ sketch grammar for Croatian: https://the.sketchengine.co.uk/auth/sketch_grammar/356/view/.

[7] The Croatian language using Reldi application service: https://github.com/clarinsi/reldi-lib-doc. Ljubešić et al. 2016.

[8] Py2neo Python library: http://py2neo.org/v3/.

[9] Graph property database Neo4j: https://neo4j.com/.

[10] Cypher programming language: https://www.opencypher.org/.

[11] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, Etienne Lefebvre „Fast unfolding of communities in large networks“, *Journal of Statistical Mechanics: Theory and Experiment* 2008 (10), 1000.

[12] Štrkalj Despot, K., Brdar, M., Essert, M., Tonković, M., Perak, B., Ostroški Anić, A., Nahod, B., Pandžić, I. (2015). *MetaNet.HR – Croatian Metaphor Repository.* [Database]. http://ihjj.hr/metafore/.