



Ljudmila PETKOVIĆ

✉ ljudmila.petkovic@sorbonne-universite.fr

🆔 0000-0003-2694-4096

Research Interests

My research focuses on Digital Humanities, Natural Language Processing and Deep Learning. I am particularly interested in the methods of dissemination, sharing and valorisation of heritage collections through the prism of Digital Humanities in order to track the circulation of knowledge.

Digital Humanities: optical character recognition (OCR), handwritten text recognition (HTR), data structuring in XML-TEI format, publication of queryable digital editions

Natural Language Processing: text mining, computational linguistics

Deep Learning: Transformers Models

Education and Qualifications

2021 – **PhD in Digital Humanities**, Sorbonne University, Paris (France), Faculty of Arts, Languages, Literature and Humanities, Doctoral School III “French and Comparative Literature” (ED 019), UMR 8599, CELLF (Center for the Study of French Language and Literature), OB TIC (Observatory of Texts, Ideas and Corpora)

2019 – 2024 **Certificate of specialisation in Linguistics**, University of Geneva (Switzerland), Faculty of Humanities, Department of Linguistics

Dissertation title : Analysis of lexical fields in the heritage collection of Jean-Martin Charcot

2017 – 2019 **Master in Social Sciences and Computing**, University of Belgrade (Serbia), interdisciplinary study program

Master’s Thesis title: CREATION AND ANALYSIS OF THE YUGOSLAV ROCK SONG LYRICS CORPUS FROM 1945 TO 2003 [[Zenodo](#)]

Abstract: The thesis analyzes, from the theoretical and practical perspective, the creation and processing of corpus of rock songs’ lyrics originating from the former Yugoslavia in the period 1945-2003. The lyrics are obtained from the LyricWiki website using the Python library `lyricsmaster` in the *web scraping* process. The collected texts are then merged into a single XML file and automatically annotated with the `yattag` Python tool. Afterwards, the data preprocessing was conducted at the formal and content level. Furthermore, the XML document is transformed into XHTML format applying XSLT processor, in order to generate basic corpus data. The diacritic restoration process with the “Slovo Majstor” application and morphological electronic dictionaries of Serbian language in the LeXimir software package, is also automated. The *text mining* process encompassed retrieving socio-political and patriotic topics using NLTK library in Python, while romantic and other topics were visualized using the TreeCloud and WordItOut software. The similarity between authors represented in the corpus was measured using `stylo` package in the programming language R. Finally, an overview of the today’s most relevant programming libraries in the field of *natural language processing* is provided, which, at the same time, serves as a guideline for the future work.

Keywords: corpus linguistics, Yugoslav rock and roll, web scraping, natural language processing, text mining.

2017 – 2018 **Master in Language, Literature, Culture**, University of Belgrade (Serbia), Faculty of Philology

Master’s Thesis title: USE OF THE MOODLE PLATFORM IN THE REFORMED TEACHING OF NEOHELLENISTIC STUDIES [[Zenodo](#)]

Abstract: This paper deals with the Modern Greek language and literature learning using the Moodle electronic platform within the *hybrid* and *online* learning model. The subject of the discussion are the main features of this teaching method, as well as the application of Moodle in working with the first-year undergraduate students of the Department of Neohellenistic Studies at the Faculty of Philology, University of Belgrade. We examine the effectiveness of the mixed-mode learning by presenting the electronic contents that are available on the online courses *Contemporary Greek Language G-1* and *G-2*, *Greek Literature 1* and *Practicum in Neohellenistics 2*. Special attention is devoted to the analysis of the course interactivity, multimediality, and flexibility. Finally, we also expose the results of the accompanying empirical research, which led to the conclusion that the implementation of Moodle in teaching significantly helps students in learning.

- 2013 – 2017 **Bachelor in Modern Greek language, literature, culture**, *University of Belgrade (Serbia), Faculty of Philology, Department of Neohellenic Studies*
- 2009 – 2013 **Philological High School, Belgrade (Serbia), French class**

Work Experience

The volumes are reported in equivalent TD hours (HETD).

- 2024-2026 ATER (Temporary Teaching and Research Associate), Sorbonne University, UFR Sociology and Computer Science for the Humanities, Bachelor's in Language Sciences and Master's in Language and Computing.

Label ELP	Code ELP	Audience	Year	Volume
Représentation et traitement des documents électroniques	M1SOL041	M1	2025-2026	CM: 10,5 TD : 10,5
Introduction au traitement automatique des langues	M2SOL035	M1	2025-2026	CM: 16,5 TD : 16,5
Corpus, ressources et linguistique outillée	M2SOL034	M1	2024-2025	CM: 10,5 TD : 10,5
IA hybride et machine learning	L6SOIAAU	L3	2025-2026	CM: 15 TD : 24
Statistiques inférentielles	L6SOSTAT	L3	2024-2025	TD : 10,5
Mathématiques	L5SOMATH	L3	2024-2025	TD : 9
Recherche d'information et site web	L3SOWEB1	L2	2024-2025	TD : 76
PIX – Formation aux compétences numériques	PIX@S2N2	L1 / L2 / L3	2024-2026	CM: 3 TD : 166,5
Analyse de discours avec programmation	L2SOPROG	L1	2024-2025	TD : 2
TOTAL : 381				

- 2023-2024 TEACHING ASSISTANT (TA), course “In-depth Studies – Digital Corpora (H5MUAB7)”, in the Master 2 Archives and the Master 2 Information and Library Sciences, (SIB), University of Angers, Angers (France), Faculty of Languages, Humanities and Social Sciences (LLSH), 3h tutorial classes, [link](#).

- 2023-2024 TEACHING ASSISTANT (TA), Sorbonne Nouvelle, Paris (France), UFR Literature, Linguistics, Didactics (LLD) – Institute of General and Applied Linguistics and Phonetics (ILPGA), Academic Minor *Digital Humanities*

Label ELP	Code ELP	Audience	Year	Volume
Introduction to Digital Humanities	L1HN001	L1	2023-2024	TD : 26
Advanced Digital Cultures	L2HN001	L1	2023-2024	TD : 26
TOTAL : 52				

- 2021 – PhD in Digital Humanities, Sorbonne University, Paris (France), Faculty of Arts & Humanities, Doctoral School III “French and Comparative Literature” (ED 019), UMR 8599, CELLF (Center for the Study of French Language and Literature), OB TIC (Observatory of Texts, Ideas and Corpora).

- 2021 – EXTERNAL CONTRIBUTOR within the [IDEA](#) project for the linguistic annotation task, Faculty of Psychology and Educational Sciences, University of Geneva (Switzerland), under the supervision of D^r Joanna Blochowiak.

- 2021 – Intern at the Observatory of Literary Life (laboratory of excellence OBVIL), as part of the *Very Large Library* project (Très Grande Bibliothèque – TGB) to improve the quality of OCR in the TGB for the task of Named Entity Recognition (NER), Sorbonne University, Paris (France), under the supervision of Prof. D^r Glenn Roe and D^r Motasem Alrahabi.
- 2021 Source entry operator for the project FNS *Visual Contagions* (exhibition catalogues, OCR-XML-TEI-CSV) [[GitLab](#)], Faculty of Humanities, Chair of Digital Humanities, University of Geneva (Switzerland), under the supervision of Prof. D^r Béatrice Joyeux-Prunel.
- 2020 Scientific collaborator and research engineer within the project FNS *Manuscripts for Sale: Automatic Structuration and Annotation of Catalogues with Complex Layouts (Katabase)*, Faculty of Arts and Humanities, Institute of French Literature, University of Neuchâtel : OCRisation (in Transkribus and Kraken)/OCR correction of the XIXth c. manuscript sales catalogues ; semi-automatic semantic encoding and data restructuring in TEI-XML generated by *GROBID-dictionaries* for filling in the dedicated database, under the supervision of D^r Simon Gabay.
- 2021 Peer reviewer of the article Rozovskaya, A. (2021). “Spelling Correction for Russian: A Comparative Study of Datasets and Methods”. *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2021)* [[link](#)] and [[link](#)].
- 2019 Peer reviewer of the article Zosa, E., & Granroth-Wilding, M. (2019). “Multilingual dynamic topic model”. *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2019)* [[link](#)] and [[link](#)].
- 2016 – English/Serbian/Modern Greek translations of financial contracts, Faculty of Philology, Department of Neohellenic Studies, University of Belgrade (Serbia), coordinator: D^r Anka Radenović.
- 2017 – 2018 Participation in the Portuguese-Serbian translation project of Portuguese-speaking literary works, Faculty of Philology, Instituto Camões, University of Belgrade (Serbia), coordinator: D^r Anamarija Marinović.
- 2015 – 2016 Private lessons in Modern Greek language for high school students.
- 2015 – 2016 Student project in lexicography – creation of the Greek-Serbian dictionary for the students of the Department of Neohellenic Studies, Faculty of Philology, University of Belgrade (Serbia), coordinator: D^r Vojkan Stojičić.

List of Publications

Peer-reviewed Journals

- 2025 Petkovic, Lj., Koudoro-Parfait, C., Desmarest, M.-S. & Lejeune, G. (2025). Bruit de fond ou valeur ajoutée ? Gérer le bruit lors des traitements informatiques des corpus linguistiques. *Revue Corpus 26 – 2025*. DOI : <https://doi.org/10.4000/1364s>
- 2024 Koudoro-Parfait, C., Petkovic, Lj., & Roe, G. (2024). Analyse multilingue de l'impact de la correction automatique de la ROC sur la reconnaissance d'entités nommées spatiales dans des corpus littéraires. *Revue TAL : Robustesse et limites des modèles de traitement automatique des langues*, 64(2). https://www.atala.org/sites/default/files/TAL_64_2_2.pdf
- 2022 Petkovic, Lj., Alrahabi, M. & Roe, G. (2022). Impact de la correction automatique de l'OCR/HTR sur la tâche de reconnaissance d'entités nommées dans un corpus bruité. *Journal of Information Sciences* 21(2), pp. 1-18. <https://doi.org/10.34874/IMIST.PRSM/jis-v21i2.36599>.
- 2019 Petković, Lj. (2019). Creation and Analysis of the Yugoslav Rock Song Lyrics Corpus from 1967 to 2003. *Infotheca – Journal for Digital Humanities*, 19(1), pp. 5-29. <hal-03091121>
- 2018 Petković, Lj. (2018). Upotreba onlajn platforme Mudl (Moodle) u reformisanoj nastavi Katedre za neohelenske studije Filološkog fakulteta Univerziteta u Beogradu (Use of the Online Platform Moodle in the Reformed Teaching within the Department of Neohellenic Studies at the Faculty of Philology, University of Belgrade). *Metodički vidici*, 9(9), pp. 179-200. <hal-03091176>

Peer-reviewed Colloquia

- 2023 Petkovic, Lj., Alrahabi, M., & Roe, G. (2023). Circulation du discours médical de Jean-Martin Charcot : premières observations. *Humanistica 2023*, Humanistica, June 2023, Geneva, Switzerland <hal-04107099>.
- 2020 Gabay, S., Petkovic, Lj., Bartz, A., Levenson, M. G., & du Noyer, L. R. (2020). Katabase: À la recherche des manuscrits vendus. *Humanistica 2021*, Humanistica, May 2021, Rennes, France. <hal-03066108>

Peer-reviewed Study Days

- 2024 Petkovic, Lj., Koudoro-Parfait, C., & Lejeune, G. (2024). [Mes données ou tes méthodes ? Comment trancher la responsabilité côté TAL ou côté HN quand les résultats ne sont pas au rendez-vous ?](#). Poster presented at the CNRS GdR TAL study day “Natural Language Processing and Digital Humanities”, 7th November 2024, University of La Rochelle.
- 2022 Petkovic, Lj., Alrahabi, M., & Roe, G. (2022). Impact de la correction automatique de l'OCR/HTR sur la tâche de reconnaissance d'entités nommées dans un corpus bruité. *Actes de la journée d'étude sur la robustesse des systèmes de TAL*, p. 22. HAL : <hal-03853541v1>

Peer-reviewed Congresses

- 2023 Alrahabi, M., Cordova, J., Fedchenko, V., Petkovic, Lj., & Roe, G. (2023). Pandore: a toolbox for digital humanities text-based workflows [résumé]. *Digital Humanities 2023: Book of Abstracts*, Digital Humanities 2023. Collaboration as Opportunity. (DH2023), Graz, Austria [doi].

- 2018 Stojičić, V., & Petković, Lj. (2018). Od oglasa do platforme Moodle – primer dobre prakse u nastavi modernog grčkog kao stranog (From an Advertisement to the Moodle Platform – an Example of Good Practice in Teaching Modern Greek as a Foreign Language): Paper presented at the 6th International Congress “Primenjena lingvistika danas – jezik, književnost i interdisciplinarnost” (“Applied Linguistics Today 6 – Language, Literature and Interdisciplinarity”) Faculty of Philology, University of Belgrade, Serbia, 12–13.10.2018. [[abstract](#)]

Peer-reviewed Conferences

- 2022 Cordova, J. M., Dupont, Y., Petkovic, Lj., Gawley, J., Alrahabi, M. & Roe, G. (2022). Toolbox : une chaîne de traitement de corpus pour les humanités numériques. *Traitement Automatique des Langues Naturelles*, July 2022, Avignon, France. pp. 11-13. [<hal-03701464>](#)
- 2020 Gabay, S., du Noyer, L. R., Levenson, M. G., Petkovic, Lj. & Bartz, A. (2020), Quantifying the Unknown: How many manuscripts of the marquise de Sévigné still exist?. In *Digital Humanities DH2020*, ADHO, July 2020, Ottawa, Canada. [<hal-02898929>](#)
- 2018 Petković, Lj. (2018). Implementacija tehnika analize teksta u programskom jeziku Python pomoću NLTK-a (Implementation of text mining techniques using NLTK in Python programming language). In *2018 26th Telecommunications Forum (TELFOR 2018)*, pp. 975-979. IEEE. [<hal-03091167>](#)

Peer-reviewed Workshops

- 2021 Scius Bertrand, A., Gabay, S., Petkovic, Lj., Janes, J., Corbières, C., & Clérice, T. (2021). The BIR database – Identifying typographic emphasis in list-like historical documents. *HIP@ICDAR21 — The 6th International Workshop on Historical Document Imaging and Processing*, September 2021, Lausanne, Switzerland. [<hal-03355683>](#)

Talks

Conferences

- 2025 **THE TEXT OF THE OTHER. INTERDISCIPLINARY DIALOGUE ON INTERTEXTUALITY AND REPORTED SPEECH** – “Analysis of the Scientific Impact of Jean-Martin Charcot on French Medical Literature from the 19th to the 20th century”, organised by the Center for Interlingual Research on Meaning in Context (CRISCO). Sorbonne Nouvelle, Maison de la Recherche | École des Chartes, 1-3 July 2025, Paris (France).
- 2023 **Humanistica 2023** – 4th edition of the Humanistica conference. 23th-28th June 2023, Geneva (Switzerland) [[link](#)].
- 2022 **CIHN’22** – International Conference CIHN’22 “Digital humanities at the centre of the management of cultural heritage, data and information. Methods, devices and perspectives”. 22-23 November 2022, Rabat (Morocco) [[link](#)].
- 2022 **UNIGE Data Science Day** – “Promises of Artificial Intelligence”, 15th September 2022, Geneva (Switzerland) [[link](#)].
- 2022 **TALN-RÉCITAL** – “Towards an inclusive NLP”, 27th June – 1st July 2022, Avignon (France) [[link](#)].
- 2021 **Humanistica 2021**, Francophone Association of Digital Humanities, 10th-12th May 2021, Rennes (France) [[link](#)].
- 2020 **DH2020** – Digital Humanities DH2020, “carrefours/intersections”, ADHO, 22th-24th July 2020, Ottawa (Canada) [[link](#)].
- 2018 **TELFOR** – 26th Telecommunications Forum, Centre Sava, Belgrade (Serbia), 26th-27th November 2018 [[link](#)].

Study Days

- 2025 **STUDY DAY: DIGITAL HUMANITIES AND ARTIFICIAL INTELLIGENCE: INTERACTIONS / ALTERCATIONS** – poster session “Analysis of Jean-Martin Charcot’s Influence through Medical Terminology Extraction: a *PatternRank* Approach”. Study day launching the Digital Humanities Initiative of Sorbonne University. Maison de la recherche, Paris (France), 7th November 2025.
- 2025 **STUDY DAY OF THE ARIANE CONSORTIUM (WG4)** “AI-Assisted Literary Analysis” – “...”, organised by the **ARIANE Consortium**. SCAI, Paris (France), 17 June 2025.
- 2024 **ARTIFICIAL INTELLIGENCE AND DIGITAL HUMANITIES** – “Measuring Charcot’s Influence on his Contemporaries Using Keyphrase Extraction”. Study day organized under the aegis of the French Association for Artificial Intelligence (AFIA), and the RG CNRS MADICS and MAGIS, National Library of France, Paris (France), 3rd May 2024.
- 2023 **Background noise or added value? Managing noise during computer processing of linguistic corpora** – co-organised by the University of Grenoble Alpes and Roma La Sapienza, House of Languages and Cultures, University Grenoble Alpes, Grenoble (France), 28th April 2023 [[link](#)].
- 2023 **Digital Humanities Study Day at Sorbonne University** – 1st study day on Digital Humanities, Maison de la recherche, Paris (France), 18th January 2023 [[link](#)].
- 2022 **RobusTAL22** – 1st study day on the robustness of the NLP systems, Maison de la recherche, Paris (France), 25th November 2022 [[link](#)].

Congresses

- 2023 **DH2023** : International Conference “DH2023 – Collaboration as Opportunity”, University of Graz | Messe Congress Graz, Graz (Austria), 10th–14th July 2023 [[link](#)].
- 2018 **ALT 6**: 6th International Congress “Primenjena lingvistika danas – jezik, književnost i interdisciplinarnost” (“Applied Linguistics Today 6 – Language, Literature and Interdisciplinarity”), Faculty of Philology, Belgrade (Serbia), 12th–13th October 2018 [[link](#)].

Seminars

- 2026 **Seminar of the “Linguistic Variation and Computational Linguistics” team**, organised by Dr Gaël Lejeune, “Tracing the circulation of concepts in noisy heritage corpora: contributions and limits of automatic tools using the Jean-Martin Charcot collection”. 11th June 2026, Maison de la Recherche, Paris (France).
- 2025 **ObTIC Team Doctoral Seminar**, SCAI, Paris (France). Seminar organised by Prof. Dr Glenn Roe and Dr Motasem Alrahabi.
- “Measuring the Impact of Jean-Martin Charcot *via* the Extraction of Medical Terms: Which Approach to Adopt?”, 6th February 2025.
 - “Comparison of Keyphrase Extraction Approaches”, 13th March 2025.
 - “Keyphrase Extraction from the Charcot Corpus: Quantitative and Qualitative Evaluation”, 15th April 2025.
- 2024 **Seminar of the “Computational Linguistics – Speed Dating” team**, organised by Dr Gaël Lejeune, “Origins and Progress of the Project Charcot”. 12th September 2024, Maison de la Recherche, Paris (France).
- 2024 **PhD Seminar/Workshop on HTR and eScriptorium**, organised by Centre Roland Mousnier of the Faculty of Arts, Languages, Literature and Humanities, Sorbonne University, Feedback Session, 26th April 2024, Centre Roland Mousnier (France).

- 2024 **PHD SEMINAR OF THE SERVICE UNIT CENTER FOR EXPERIMENTATION IN DIGITAL METHODS FOR RESEARCH IN THE HUMANITIES AND SOCIAL SCIENCES (CERES)**, “**APPROACHES TO DIGITAL METHODS**” – “**In Charcot’s Private Papers: from Experimentation to the Premises of Modern Neurology**”. Seminar organised by Léa Andolfi, Julien Bezançon, Clara Bordier, Thibault Grison, Rimane Karam, Adélie Laruncet and Marine Tiger, 27th March 2024, Maison de la Recherche, Paris (France).
- 2023 **Seminar for Master’s students, “Literary Digital Humanities”**, organised by Prof. D^r Glenn Roe, “Circulation of the Jean-Martin Charcot’s medical discourse: first observations”, 6th December 2023, Maison de la Recherche, Paris (France) [[link](#)].
- 2023 **Doctoral seminar/workshop on OCR and HTR**, organised by Centre Roland Mousnier of the Faculty of Arts, Languages, Literature and Humanities, Sorbonne University, “Digitisation of documents transcription, annotation, OCR, HTR, correction...”, with D^r Motasem Alrahabi, 22th November 2023, Maison de la Recherche, Paris (France) [[link](#)].
- 2023 **Seminar of the “Computational Linguistics” team**, organised by D^r Gaël Lejeune, “Automatic correction of OCR interferences in the recognition of spatial named entities: real gain or loss of information?”, 13th April 2023, Maison de la Recherche, Paris (France) [[link](#)].
- 2022 **Seminar “Literary Digital Humanities”**, organised by Prof. D^r Glenn Roe, “Impact of automatic correction of the OCR/HTR on the named entity recognition task in noisy corpora”, 16th November 2022, Maison de la Recherche, Paris (France) [[link](#)].
- 2021 **JeRTeh**, “Katabase : In Search of Lost Manuscripts”, organised by JERTEH, Language Resources and Technologies Society in Belgrade (Društvo za Jezičke Resurse i Tehnologije, Beograd), 10th June 2021, Belgrade (Serbia) [en ligne] [[link](#)].

Workshops

- 2024 **Python Workshop** – “**Keyphrase Extraction from the Charcot collection**”, organised by the team-project OB TIC, 30th April 2024, DataLab, National Library of France, Paris (France).
- 2023 – **ANHIMO Workshop**, “**Use of the Automatic Text Transcription Techniques for the Analysis of Typed Maritime Documents of the XXth c.**”, Summer School of Numerical Analysis of the History of the Sea and Oceans (ANHIMO 2023), dedicated to numerical analysis of the History of the Sea and Oceans and organised by the Ocean Institute of Sorbonne University Alliance and Sorbonne Centre for Artificial Intelligence (SCAI), Paris (France), 26th-30th June 2023.
- 2021 **Python Workshop**, “Handling text files”, organised by the team-project OB TIC, 5th December 2021, Paris (France) [[link](#)].
- 2021 **OCR Workshop**, “Automatic correction of the OCR output”, organised by the team-project OB TIC, 25th November 2021, Paris (France) [[link](#)].
- 2021 **OCR Workshop**, “Transkribus, Kraken, eScriptorium, Tesseract”, organised by the team-project OB TIC, 18th November 2021, Paris (France), with Johanna Cordova [[link](#)].
- 2021 **OCR Workshop**, “Presentation and use of OCR/HTR tools: Abbyy FineReader, Transkribus, Kraken and eScriptorium”, organised by team-project OB TIC, 28th October 2021, Paris (France), with D^r Motasem Alrahabi [[link](#)].

Co-supervision of Master's Thesis

- 2025 Camille Fief & Yeleen Canonge (M1 – Sorbonne University) : “**Reconnaissance d’entités nommées et entraînement de modèle sur données bruitées**” (“Named Entity Recognition and Model Training on Noisy Data”), with Gaël Lejeune and Caroline Koudoro-Parfait.
- 2025 Ruixing Zheng et Marina Carvalho Lopes (M1 – Sorbonne University) : “**Reconnaissance d’Entités Nommées sur un corpus de poèmes du 19^e siècle**” (“Named Entity Recognition on a Corpus of 19th Century Poems”), with Gaël Lejeune and Caroline Koudoro-Parfait.

Additional Training

- 2025 [DFC] **ETHICS OF SCIENTIFIC RESEARCH**, organised by Sorbonne University, Campus Cordeliers, Sorbonne University, Paris (France) [participation : in person] [**certificate of attendance**].
- 2023 “Summer School on Digital Humanities”, as part of the strategic alliance between National Autonomous University of Mexico (UNAM) and Sorbonne Center for Artificial Intelligence (SCAI). Applied Mathematics and Systems Research Institute (IIMAS), Mexico City, Mexico [**certificate of attendance**].
- 2022 EnExDi (*Encoder, Exploiter, Diffuser*) – “Digital humanities in research projects”, intensive training school in digital humanities, The House of the Humanities and Social Sciences (MSHS), University of Poitiers, France [**certificate of attendance**].
- 2022 “Introduction to Natural Language Processing”, course offered as part of the Master “Digital Humanities” of the PSL University (Paris Sciences & Letters) | ENC (National School of Charters) [**certificate of attendance**].

Scholarships and Awards

- 2023 – Travel Bursary ADHO of the DH2023 Conference, Graz (Austria) [[link](#)].
- 2021 – Funding of the PhD contract by the French Excellence Initiative (Idex), as part of the Institutes and Initiatives Programme of The Sorbonne University Alliance (ASU) (The Heritage Observatory – OPUS) [[link](#)].
- 2017 – 2018 Dositeja – Serbian Government scholarship (Fund for Young Talents of the Republic of Serbia) for the best Master’s students
- 2016 – 2017 Dositeja – Serbian Government scholarship (Fund for Young Talents of the Republic of Serbia) for the best Bachelor’s students
- 2016 Dean’s Honor. Mention for Outstanding academic achievement, University of Belgrade (Serbia), Faculty of Philology
- 2014 – 2016 Serbian Government scholarship (Ministry of Education, Science and Technological Development) for the Bachelor’s students
- 2014 Training certificate from the Center for Lifelong Learning and Training of the University of Ioannina (Center for Greek Language Education and Culture DI.K.E.P.P.E.E. “Stavros Niarchos”) – IKY scholarship for the Summer School of the Greek language and culture in Ioannina (Greece)

Skills

Informatics

- MS Office Word, Excel, Access, PowerPoint
- OS Windows, macOS, Linux

Languages

Script	Python (os, re, numpy, matplotlib, scipy, pandas, scikit-learn, nltk, tensorflow, pytorch, gensim, requests, spacy)..., R (stylo, ggplot2, knitr), Bash, Perl
Development	HTML/CSS
Databases	SQL
Markup	XML, L ^A T _E X, Markdown
Stylesheets	XSLT

Version Control

Git

Other Softwares

Transkribus, Kraken, eScriptorium, TXM, GROBID-dictionaries, Gephi, Docker, Oxygen XML Editor, Anaconda, Visual Studio Code, RenderX XEP Engine, DARIAH-DE Topics Explorer, SPSS, SQL Server Management Studio, Azure Data Studio, Unitex, PyCharm, TreeCloud, Zotero, PIX, Overleaf, Texmaker, TeXstudio

Familiar with

DTS, IIIF, SPARQL, Flask, Protégé, MATLAB, Isabelle

Audiovisual Arts

Final Cut Pro X, Sony Vegas, FL Studio, MAGIX, Audacity, Virtual DJ

Languages

Serbian	Native
Français	Fluent
Anglais	Fluent
Grec moderne	Fluent
Portugais	Intermediate
Coréen	Beginner
Latin	Reading, writing
Grec ancien	Reading, writing

Sports Experience

- 2024 – taekwondo – violet belt (XX *keup*), club SP Training, Bourg-la-Reine (France).
- 2021 – 2023 taekwondo – blue belt-two red stripes (4th *keup*), club Paris Taekwondo Elite, Paris (France).
- 2019 – 2021 taekwondo – green belt-blue stripe (5th *keup*), club Il Gi Dojang, Geneva (Switzerland).
- 2015 – 2019 taiji quan – club “Peng”, Belgrade (Serbia).
- 2015 – 2019 qigong – certificate of participation in modules I and II for the instructor of qigong, Serbian qigong federation, Belgrade (Serbia).

Other Interests

- 2018 – 2019 Member of the Society for Language Resources and Technologies in Belgrade (Serbia) JERTEH (Društvo za Jezičke Resurse i Tehnologije, Beograd [[link](#)])
 - Audiovisual arts
 - Eastern Asian languages and cultures (Chinese, Japanese, Korean)