

INFORMAȚII
PERSONALE

Mititelu Verginica

 0740.079.117 vergi@racai.ro <https://www.racai.ro/en/about-us/racai-staff/verginica-barbu-mititelu/>LOCUL DE MUNCA
PENTRU CARE SE
CANDIDEAZĂ
POZIȚIALOCUL DE MUNCĂ
DORITSTUDIILE PENTRU CARE
SE CANDIDEAZĂ
PROFILUL PERSONAL

Membru în comisia de susținere publică a tezei dnei Alexandra Lazăr (Ilie)

EXPERIENȚA
PROFESIONALĂ

Febr 2024 – prezent

Director adjunct

Institutul de Cercetări pentru Inteligență Artificială „Mihai Drăgănescu”, București
(Calea 13 Septembrie nr 13, sector 5, www.racai.ro)

Coordonarea activității de cercetare

Tipul sau sectorul de activitate cercetare

Febr 2003 – ian 2024

Cercetător științific II (1 nov 2017-prezent), cercetător științific III (ian 2014-oct 2017), cercetător științific (martie 2004-ianuarie 2014), asistent de cercetare științifică (feb 2003-feb 2004)

Institutul de Cercetări pentru Inteligență Artificială „Mihai Drăgănescu”, București
(Calea 13 Septembrie nr 13, sector 5, www.racai.ro)

- Cercetare fundamentală
- coordonarea temei de plan „Resurse lingvistice relevante pentru caracteristicile specifice și cele universale ale limbii române în contextul multilingv actual” (2024-2026)
- coordonarea temei de plan „Resurse lingvistice pentru prelucrarea limbii române în contextul global actual” (2020-2023)

Tipul sau sectorul de activitate cercetare

Febr 2003 – martie 2007

cercetător științific

Institutul de Lingvistică „Iorgu Iordan – Al. Rosetti”, București (Calea 13 Septembrie, nr 13, sector 5, www.lingv.ro)

- Cercetare fundamentală
- participare, în calitate de co-autor, la redactarea lucrării „Atlasul lingvistic român pe regiuni. Sinteză” – volumele I și II

Tipul sau sectorul de activitate cercetare

Sept 2002 – ian 2003

Profesor în învățământul preuniversitar

Școala Generală nr. 202, București (Calea Crângași, nr. 140, sector 6)

- Predarea limbii engleze

Sept 2000 – august 2002

Tipul sau sectorul de activitate educație/învățământ preuniversitar

Profesor în învățământul preuniversitar

Grup Școlar „Iuliu Maniu”, București (Bd Iuliu Maniui, nr. 381-391, sector 6)

- Predarea limbii engleze

Tipul sau sectorul de activitate educație/învățământ preuniversitar

EDUCAȚIE ȘI FORMARE

2011-2013 Studii postdoctorale

i

Școala postdoctorală „Valorificarea identităților culturale în procesele globale” (POSDRU/89/1.5/S/59758) din cadrul Academiei Române

2002-2009

- tema de cercetare: *Rețea semantico-derivațională pentru limba română*

Studii doctorale: diplomă de doctor în filologie (Summa cum laude)

Facultatea de Litere din cadrul Universității din București

- titlul tezei: *Relații semantice (sinonimie, antonimie, hiponimie). Abordare din perspectiva lingvisticii teoretice și a lingvisticii computaționale*

2001-2002

Studii aprofundate: programul „Lingvistică teoretică”

Facultatea de Litere și Facultatea de Limbi și Literaturi Străine din cadrul Universității din București

Titlul disertației: *Negația verbală în limba română*

- În perioada studiilor aprofundate, am beneficiat de două burse NATO (15 oct – 15 dec 2001 și 15 apr – 28 iunie 2002) la Universitatea Tuebingen, Germania.

1996-2000

Facultatea de Litere, specializările: limba și literatura română și limba și literatura engleză

Universitatea din București

- titlul lucrării de licență: *Pasivul în limba română*

- În perioada 1 martie – 30 iunie 2000, am fost student Erasmus la Universitatea din Viena, Institutul de Anglistică și Americanistică

28 iunie 2023

COST Academy training

Science Diplomacy workshop, Bruxelles, Belgia

- Workshop dedicat acțiunilor COST pentru a înțelege interacțiunea dintre diplomație și știință

27 iunie 2023

COST Academy training

Science for Policy workshop, Bruxelles, Belgia

- Workshop dedicat acțiunilor COST care doresc să împărtășească rezultatele obținute pentru un impact mai mare asupra factorilor de decizie politică

17-24 aprilie 2023

Fundația Centrul de Formare APSAP

Curs Manager Proiect, online

- Cod COR 242101, curs autorizat de Ministerul Educației și Cercetării și Ministerul Muncii și Protecției

8-12 februarie 2021

EuroLan 2021

Training school on Linked Data în Linguistics, online

15-19 ianuarie 2020

- Cursuri de introducere în metodologiile de reprezentare a resurselor lingvistice folosind tehnologiile Webului Semantic

SC Soft Skills Training SRL

Curs de management de inovare, București, România

27 iunie – 1 iulie 2016

- Cod COR 242106, curs autorizat de Ministerul Muncii și Justiției Sociale și Ministerul Educației Naționale

PARSEME COST Action

Second training school, La Rochelle, Franța

19-23 ianuarie 2015

- Cursuri de frazeologie, adnotarea sintactică a expresiilor, traducere automată și problemele ridicate de frazeologie

PARSEME COST Action

First training school, Praga, Cehia

9-20 septembrie 2002

- Cursuri de frazeologie, gramatici de dependențe și adnotarea sintactică a expresiilor, Head-driven Phrase Structure Grammar și reprezentarea expresiilor, reprezentarea expresiilor în băncile de arbori sintactici

International Doctoral Program „Empirical Linguistics and Natural Language Processing”

Fall school, Sozopol, Bulgaria

- Cursuri de la interfața dintre lingvistica computațională, lingvistica corpusului și formalisme gramaticale

COMPETENTE PERSONALE

Limba maternă

Limba română

Alte limbi străine cunoscute

	INTELEGERE		VORBIRE		SCRIERE
	Ascultare	Citire	Participare la conversație	Discurs oral	
engleză	C1	C1	C1	C1	C1
germană	A2	B1	A2	A2	A2
	Zertifikat Deutsch				
franceză	A2	C1	A1	A1	B1

Niveluri: A1/2: Utilizator elementar - B1/2: Utilizator independent - C1/2: Utilizator experimentat

Cadrul european comun de referință pentru limbi străine

Competențe de
comunicare

bune competențe de comunicare dobândite prin experiența proprie de lucru în diverse colective de cercetători

Competențe organizaționale/manageriale

Competențe organizaționale dobândite și îmbunătățite prin implicarea în organizarea următoarelor evenimente științifice:

- ELRC 2021 <https://lr-coordination.eu/romania3rd> online, 29 sept 2021
- ConsILR 2020, online, 14-15 dec 2020 <https://profs.info.uaic.ro/~consilr/>
- 4th ReTeRom workshop, online, 16 dec 2020, <https://profs.info.uaic.ro/~consilr/>
- Approaches to Discourse-Relational Devices (DRDs): Textual Connectors, Discourse Markers, Modal Particles, București, 15-16 Oct 2019.
- Joint Workshop on Multiword Expressions and WordNet (MWE-WN 2019), 2 august 2019, satelit al ACL 2019
- ELRC workshop în București, 23 martie 2016
- The 8th Global WordNet Conference, 27-30 ianuarie 2016, București
- EuroLAN 2015
- Errors by Humans and Machines in multimedia, multimodal and multilingual data processing – ERRARE 2015 workshop, 11-13 sept 2015, Sinaia, Romania
- Semantic Relations-II. Enhancing Resources and Applications, 22 May 2012, workshop satelit al LREC 2012, Istanbul, Turkey
- Semantic Relations. Theory and Applications, 18 mai 2010, workshop satelit al LREC 2010, Valletta, Malta

Am dobândit **competențe manageriale** prin:

- urmarea celor două cursuri menționate la secțiunea Educație și formare, precum și
- prin implicarea în următoarele activități:
 - Working Group 2 leader în cadrul Acțiunii COST UniDive (<https://unidive.lisn.upsaclay.fr/doku.php?id=wg2:wg2>) (septembrie 2022-prezent)
 - WG1 Task 1.2. task co-leader în cadrul Acțiunii COST European network for Web-centred linguistic data science (CA 18209) - Nexus Linguarum (<https://nexuslinguarum.eu/>)
 - responsabil pentru partea românească în cadrul proiectului „Enhancing Language Resources with Derivationally Linked Multiword Expressions” – proiect bilateral cu Institutul de Lingvistică al Academiei Bulgare de Științe (2018-2021)
 - director al proiectului științific „Analizor sintactic (pentru limba română) bazat pe cunoștințe semantice” (PN-II-RU-TE-2014) (2015-2017);
 - responsabil pentru partea românească în cadrul proiectului „Enhanced Knowledge Bases for Bulgarian and Romanian” – proiect bilateral cu Institutul de Lingvistică al Academiei Bulgare de Științe (2015-2017)

INFORMATII SUPLIMENTARE

Prezentări invitate

- Verginica Barbu Mititelu, *On the Reference Corpus of the Contemporary Romanian Language (CoRoLa)*, Digital Humanities la ICUB, 3 decembrie 2018.
- Openly Released Corpora for Romanian, invited talk at the 5th Conference of the Mathematical Society of the Republic of Moldova, dedicated to the 55th anniversary of the foundation of Vladimir Andrunachievici Institute of Mathematics and Computer Science, Chișinău, 28 Sept-2 Oct 2019.
- CoRoLa. Accesul la textele scrise și la cele orale, la Conferințele Centrului de Lingvistică Teoretică și Aplicată (CLTA), Universitatea din București, 21 ian 2021.
- *Resurse lingvistice dezvoltate de ICIA pentru îmbogățirea e-Translation* în cadrul “Annual Conference of the Romanian Language Department: AI, big data and political correctness – arranged marriage or a match made in heaven?”, 25 nov

2022.

Proiecte

Pentru proiectele coordonate, a se vedea secțiunea „Competențe organizaționale/manageriale”.

Proiecte în care am fost executant:

A. proiecte internaționale:

- System for the anonymisation of Romanian Jurisprudence (SARoj), finanțat în proiectul TJENI de EEA and Norway Grants
- Underrepresented speech dataset from open data: case study on the Romanian language (USPDATRO), finanțat în ELE 2
- Universality, diversity and idiosyncrasy in language technology (UniDive) COST Action
- Multilingual Semantic Textual Similarity (<http://tharindu.co.uk/Multilingual-Semantic-Textual-Similarity/>)
- ENRICH4ALL (E-goverNment [RI] CHatbot for ALL) (<https://www.enrich4all.eu/>)
- LITHME (Language in the Human-Machine Era) COST Action
- CURLICAT (Curated Multilingual Language Resources for CEF AT) - CEF AT, <https://curlicat-project.eu/>
- ELG (European Language Grid)
- European network for Web-centred linguistic data science (CA 18209) - Nexus Linguarum (<https://nexuslinguarum.eu/>)
- Translation Automation Services for EU Council Presidency (2018-EU-IA-0079), <https://ec.europa.eu/inea/en/connecting-europe-facility/cef-telecom/2018-eu-ia-0079>
- Multilingual Resources for CEF.AT in the legal domain (2017-EU-IA-0136) (MARCELL), <https://marcell-project.eu/>
- DRuKoLA - Sprachvergleich korpus-technologisch. Deutsch – Rumänisch, Research Group Linkage Programme, funded by the Alexander von Humboldt-Foundation.
- Adding Morpho-Semantic Information to the Bulgarian and Romanian Wordnets – proiect bilateral cu Institutul de Lingvistică al Academiei Bulgare de Științe (2012-2014)
- COST Action PARSEing and Multi-word Expressions (PARSEME)
- METANET4U – Enhancing the European Linguistic Infrastructure (ICT PSP Objective identifier: 6.1 Open linguistic infrastructure) (Grant Agreement No 270893)
- ACCURAT - Analysis and evaluation of Comparable Corpora for Under Resourced Areas of machine Translation (FP7/2007-2013, Grant Agreement no. 248347)
- BALKANET - IST European project, <http://www.dblab.upatras.gr/balkanet/>
- REGEN – proiect la Research Group in Computational Linguistics, University of Wolverhampton

B. proiecte naționale:

- Roboții și societatea. Sisteme cognitive pentru Roboți Personali și Vehicule Autonome (ROBIN) (72PCCDI/2018)
- Resurse și tehnologii pentru dezvoltarea interfețelor om-mașină în limba română (ReTeRom) (73PCCDI/2018)
- ACCURAT-Ro - Analiza și Evaluarea Rolului în Traducerea Automată a Corpusurilor Comparabile pentru Limbi cu Puține Resurse Lingvistice (PN2, Capacitati)

- eDTLR, Dictionarul Tezaur al Limbii Române în format electronic (PN2, D91013/18.09.2007)
- SIR-RESDEC - Sistem de Întrebare-Răspuns în limbile Română și Engleză cu Spații Deschise de Căutare (PN2, D11007/18.09.2007)
- STAR - Sistem de Traducere Automat pentru limba Română (PN2, Idei)
- CLARIN-RO - Infrastructură pentru resurse lingvistice interoperabil pentru limba română (PN2, Capacitati)
- ROTEL - Sisteme inteligente pentru Webul semantic, bazate pe logica ontologiilor și tehnologiile limbajului. Aplicații pentru limba română – finanțat de MEC (CEEX 29)
- Bază de date sintactico-semantică în format XML: valențele combinatorii ale verbelor românești în reprezentare HPSG – finanțat de CNCSIS, 2005-2007
- Dicționar morfologic în format electronic al limbii române contemporane – finanțat de CNCSIS, 2002-2004

Distincții și premii

- Premiul Bogdan Petriceicu Hasdeu al Academiei Române
- Best paper award la International Conference on Intelligent Information Systems, 20-23 august 2013, Chișinău, Republica Moldova pentru lucrarea *Increasing the Effectiveness of the Romanian Wordnet in NLP applications*.
- Best poster award la LDK 2023 – 4th Conference on Language, Data and Knowledge

Publicații (ultimele zece)

1. Verginica Mititelu, Voula Giouli, Gražina Korvel, Chaya Liebeskind, Irina Lobzhanidze, Rusudan Makhachashvili, Stella Markantonatou, Aleksandra Markovic, and Ivelina Stoyanova. 2025. [Survey on Lexical Resources Focused on Multiword Expressions for the Purposes of NLP](#). In *Proceedings of the 21st Workshop on Multiword Expressions (MWE 2025)*, pages 41–57, Albuquerque, New Mexico, U.S.A.. Association for Computational Linguistics.
2. Ion, Radu, Vasile Păiș, Verginica Barbu Mititelu, Elena Irimia, Maria Mitrofan, Valentin Badea, and Dan Tufiș. 2025. "Unsupervised Word Sense Disambiguation Using Transformer's Attention Mechanism" *Machine Learning and Knowledge Extraction* 7, no. 1: 10. <https://doi.org/10.3390/make7010010>
3. Păiș, Vasile, Verginica Barbu Mititelu, Elena Irimia, Radu Ion, and Dan Tufiș. 2024. "Under-Represented Speech Dataset from Open Data: Case Study on the Romanian Language" *Applied Sciences* 14, no. 19: 9043. <https://doi.org/10.3390/app14199043>
4. Verginica Barbu Mititelu, Tudor Voicu, Function Multiword Expressions Annotated with Discourse Relations in the Romanian Reference Treebank. In *Proceedings of CLiB 2024*, Sofia, Bulgaria, p. 90-97.
5. Radu Ion, Verginica Barbu Mititelu, Vasile Păiș, Elena Irimia, Valentin Badea, A cross-model study on learning Romanian parts of speech with Transformer models. In *Proceedings of CLiB 2024*, Sofia, Bulgaria, p. 6-13.
6. Păiș, V., Ion, R., Irimia, E., Barbu Mititelu, V., Badea, V., Tufiș, D. System for the anonymization of Romanian jurisprudence. *Artificial Intelligence and Law* (2024). <https://doi.org/10.1007/s10506-024-09420-y>
7. Maria Pia di Buono, Blerina Spahiu, and Verginica Barbu Mititelu. 2024. [Evaluating Large Language Models for Linguistic Linked Data Generation](#). In *Proceedings of the Workshop on Deep Learning and Linked Data (DLnLD) @ LREC-COLING 2024*, pages 66–75, Torino, Italia. ELRA and ICCL.
8. Verginica Barbu Mititelu, Voula Giouli, Kilian Evang, Daniel Zeman, Petya Osenova, Carole Tiberius, Simon Krek, Stella Markantonatou, Ivelina Stoyanova,

Ranka Stanković, and Christian Chiarcos. 2024. [Multiword Expressions between the Corpus and the Lexicon: Universality, Idiosyncrasy, and the Lexicon-Corpus Interface](#). In *Proceedings of the Joint Workshop on Multiword Expressions and Universal Dependencies (MWE-UD) @ LREC-COLING 2024*, pages 147–153, Torino, Italia. ELRA and ICCL.

9. Agata Savary, Daniel Zeman, Verginica Barbu Mititelu, Anabela Barreiro, Olesea Caftanator, Marie-Catherine de Marneffe, Kaja Dobrovoljc, Gülşen Eryiğit, Voula Giouli, Bruno Guillaume, Stella Markantonatou, Nurit Melnik, Joakim Nivre, Atul Kr. Ojha, Carlos Ramisch, Abigail Walsh, Beata Wójtowicz, and Alina Wróblewska. 2024. [UniDive: A COST Action on Universality, Diversity and Idiosyncrasy in Language Technology](#). In *Proceedings of the 3rd Annual Meeting of the Special Interest Group on Under-resourced Languages @ LREC-COLING 2024*, pages 372–382, Torino, Italia. ELRA and ICCL.

10. Andrei-Marius Avram, Verginica Barbu Mititelu, Vasile Păiș, Dumitru-Clementin Cercel and Ștefan Trăușan-Matu, Multilingual Multiword Expression Identification Using Lateral Inhibition and Domain Adaptation, in Florentina Hristea (ed.), *Current Trends in Natural Language Processing (NLP) and Human Language Technology (HLT)*, MDPI, 2024, ISBN 978-3-7258-0085-8 (Hbk), ISBN 978-3-7258-0086-5 (PDF), doi.org/10.3390/books978-3-7258-0086-5.