

INFORMAȚII PERSONALE



Tufis Dan Ioan

 Heleşteului nr. 9A, Com. Mogoşoaia, Jud. Ilfov, Romania

 +40 740165400

 tufis@racai.ro

 <http://www.racai.ro/about-us/racai-staff/tufis/>

Sexul masculin | Data naşterii 05 februarie 1954 | Naţionalitatea română

EXPERIENȚA PROFESIONALĂ

august 2002 – prezent

Director

ICIA – Research Institute for Artificial Intelligence „Mihai Drăgănescu” of the Romanian Academy

[http:// www.racai.ro](http://www.racai.ro)

decembrie 1998- iulie 2002

Director

CCAIAPLPMC – Romanian Academy

septembrie 1989 - decembrie 1998

Sef laborator

ICI – National Institute for Research and Development in Informatics

septembrie 1980 – septembrie 1989

Research Group Lead

ICI – Institutul Național de Cercetare Dezvoltare în Informatică

EDUCAȚIE ȘI FORMARE

Ianuarie 1990- martie 1992

Phd in Computer Science

Universitatea Politehnică in Bucureşti – Faculty of Automation and Computers, Bucureşti, România

Iunie-august 1991

Summer Linguistic Institute Santa Cruz, California

Septembrie 1974- Iulie 1979

Universitatea Politehnică in Bucureşti – Faculty of Automation and Computers, Bucureşti, România , diploma engineer

PERSONAL COMPETENCES

Maternal Language)

Romanian

Other known languages

UNDERSTANDING		CONVERSATION		WRITING
Listening	Reading	Participating in conversation	Listening	Reading
C2	C2	C2	C2	C2



English

French	A2	B1	A1	A1	A1
--------	----	----	----	----	----

Other mentions

Full Member of the Romanian Academy (Academician)
 Order "Steaua României" (Knight degree) awarded by the President of Romania
 Doctor Honoris Causa of the University "Agora" of Oradea
 Honorary Professor of the University "A.I. Cuza" din Iași
 The Romanian Academy prize „Traian Vuia” for research in the area of computational morphology (paradigmatic morphology and the PARADIGM system)
 The Romanian Academy prize „Tudor Tănăsescu” for research in the area of linguistic modelling based on unification (MAC-ELU system)
 Merit Diplomas for essential contributions to the research programs at the National Institute for Research and Development in Informatics and Romanian Academy.

More than 3150 de citations, h-index 26.

PhD Director (din 2003)

Member in professional organisations:

- Professional Member of ACM
- European Language Grid
- European Language Resource Coordination (National Technological Contact Point, member of the board LRB)
- European Association for Computational Linguistics **EACL**,
- European Association for Machine Translation, **EAMT**
- Expert Advisory Group on Language Engineering Standards **EAGLES**,
- Trans European Language Resources Infrastructure **TELRI Gmb**,
- Association for Computational Linguistics Special Interes Group on Lexicons **ACL-SIGLEX**,
- Global WordNet Association **GWA**,
- Forumul Român pentru Societatea Cunoașterii (founding member)
- Asociația Română de inteligență Artificială -**ARIA** (founding member)
- Asociația Română de Lingvistică Computațională – **ARLC** (founding member)
- Comisia Academiei Române de Informatizare pentru Limba Română (President)

Member of the Editorial Boards of several international journals:

- International Journal on Language Resources and Evaluation, Springer Verlag, USA (from 2002)
- The Computer Science Journal of Moldova, Republica Moldova (from 2005)
- Proceedings of the Romanian Academy, Series A (din 1999)- topical editor
- Romanian Journal on Science and Technology of Information, Academia Română (from 1997)
- International Journal on Information and Control, ICI, România (from 1992)
- International Journal of Computers, Communication & Control, Universitatea „Agora” din Oradea (from 2006)
- Romanian Journal of Human – Computer Interaction, ACM SIGCHI Romania (from 2008)
- NOESIS – Travaux du comite roumain d’histoire et de philosophie de sciences, Academia Româna (from 2007) –director (from 2015)



- Member of more than 120 de Program Committees for international conferences
- Member of more than 150 de Program Committees to international conferences.
- Member in more than 60 international and national de PhD awarding committees at:
 - „A. I. Cuza” University in Iaşi
 - “Politehnica” University in Bucureşti,
 - University Bucureşti,
 - University of Craiova,
 - Romanian Academy,
 - University “Luis Pasteur” of Strasbourg, Franţa
 - Science and Technology University de of Lille, Franţa
 - University of Pisa, Italia
 - Mumbai Institute of Technology,
 - Institute for Mathematics and Informatics of the Academy of Republic of Moldova in Chişinău
 - Politechnic Institute in Grenoble, Franţa
 - Technical University in Strasbourg, Franţa
 - University of Alicante, Spania

Driving License

B

ADDITIONAL INFORMATION

Publications of the last 5 years (2019-2014):

1. Andrei Coman, Maria Mitrofan, Dan Tufiş. (2019). Automatic Identification and Classification of Legal Terms in Romanian Law Texts. In Proceedings of the 14th Edition of the **International Conference on Linguistic Resources and Tools for Natural Language Processing**, Iaşi, Noiembrie 2019
2. Vasile Păiş, Dan Tufiş, Radu Ion. (2019). Integration of Romanian NLP tools into the RELATE platform. In Proceedings of the 14th Edition of the **International Conference on Linguistic Resources and Tools for Natural Language Processing**, Iaşi, Noiembrie 2019
3. Verginica Barbu Mititelu, Elena Irimia, Dan Tufiş. (2019). Linking the Romanian Wordnet to the Open Multilingual Wordnet. Special Issue on Linking, Integrating and Extending Wordnets. CSLI Publications, Stanford, 2019 (in print)
4. Dan Tufiş, Verginica Barbu Mititelu, Elena Irimia, Vasile Păiş, Radu Ion, Nils Diewald, Maria Mitrofan, Mihaela Onofrei. (2019). Little Strokes Fell Great Oaks. Creating Corola, The Reference Corpus of Contemporary Romanian. *Revue roumaine de linguistique*, No./Issue 3, 2019
5. Verginica Barbu Mititelu, Dan Tufiş. 2019 Dificultăţi în prelucrarea morfologică automată a limbii române. În Gh. Chivu, Cătălina Vătăşescu (eds) *Omagiu Profesorului Grigore Brâncuş*. Editura Universităţii Bucureşti, ISBN 978-606-16-1068-6
6. Dan Tufiş, Verginica Barbu Mititelu, Elena Irimia, Maria Mitrofan, Radu Ion, George Cioroiu. (2019). Making Pepper Understand and Respond in Romanian. Proceedings of the 22nd International Conference on Control Systems and Computer Science, Bucureşti, 2019
7. Inguna Skandina, Robert Gaizauskas, Bogdan Babych, Nicola Ljubesic, Dan Tufiş, Andrejs Vasiljevs (2019) (eds). Using Comparable Corpora for Under-Resourced Areas of Machine Translation. In *series Theory and Applications of Natural Language Processing*, Springer, 2019, 978-3-319-99003-3



8. Monica Lestari Paramita, Ahmet Aker, Paul Clough, Robert Gaizauskas, Nikos Glaros, Nikos Mastropavlos, Olga Yannoutsou, Radu Ion, Dan Ştefănescu, Alexandru Ceauşu, Dan Tufiş (2019). Collecting Comparable Corpora. În *Inguna Skandina*, Robert Gaizauskas, Bogdan Babych, Nicola Ljubesic, Dan Tufiş, Andrejs Vasiljevs (eds). Using Comparable Corpora for Under-Resourced Areas of Machine Translation. In *series Theory and Applications of Natural Language Processing*, Springer, 2019, 978-3-319-99003-3, pp. 55-88.
9. Ahmet Aker, Alexandru Ceauşu, Yang Feng, Robert Gaizauskas, Sabine Hunsicker, Radu Ion, Elena Irimia, Dan Ştefănescu and Dan Tufiş (2019). Mapping and aligning units from comparable corpora. În *Inguna Skandina*, Robert Gaizauskas, Bogdan Babych, Nicola Ljubesic, Dan Tufiş, Andrejs Vasiljevs (eds). Using Comparable Corpora for Under-Resourced Areas of Machine Translation. In *series Theory and Applications of Natural Language Processing*, Springer, 2019, 978-3-319-99003-3
10. Daniela Gîfu, Diana Trandabăţ, Vasile Păiş, Dan Cristea, Dan Tufiş (eds) (2018). The Proceedings of the 13th Edition of the **International Conference on Linguistic Resources and Tools for Processing Romanian Language.**, Iaşi, Noiembrie 2018
11. Dan Tufiş, Dan Cristea (2018). A Bird's-eye View of Language Processing Projects at Romanian Academy, Miyazaki, Japan, LREC 2018, pp. 2445-2451, ISBN 979-10-95546-00-9
12. Maria Mitrofan, Dan Tufiş (2018). BioRo: The Biomedical Corpus for the Romanian Language, Miyazaki, Japan, LREC 2018, pp. 1192-1196, ISBN 979-10-95546-00-9
13. Verginica Barbu Mititelu, Dan Tufiş, Elena Irimia (2018). The Reference Corpus of the Contemporary Romanian Language (CoRoLa), Miyazaki, Japan, LREC 2018, pp. 1178-1185, ISBN 979-10-95546-00-9
14. Vasile Păiş, Dan Tufiş (2018). Computing distributed representations of words using the CoRoLa corpus. In *Proceedings of the Romanian Academy, series A*, pp. 403-410, 2018
15. Marc Kupietz, Andreas Witt, Piotr Bański, Dan Tufiş, Dan Cristea, Tamas Varadi. (2017). EuReCo - Joining Forces for a European Reference Corpus as a sustainable base for cross-linguistic research in Europe. In *Proceedings of the 5th Workshop on Challenges in the Management of Large Corpora + Big Data and Natural Language Processing, Corpus Linguistics Conference*, Birmingham, 24-28 July, 2017.
16. Verginica Barbu Mititelu, Svetlozara Leseva, Dan Tufiş (2017). The Bilateral Collaboration for the Post-BalkaNet Extension of the Bulgarian and the Romanian Wordnets, ILB, 15-16 May, Sofia, 2017
17. Tiberiu Boros, Sonia Pipa, Verginica Barbu Mititelu, Dan Tufiş. (2017). A data-driven approach to verbal multiword expression detection. PARSEME Shared Task system description paper. ACL, Valencia, 2017.
18. Dan Tufiş, Dan Cristea (2017). An Outlook over CoRoLa: The Reference Corpus of Contemporary Written and Spoken Corpus, *Proceedings of SpeD conference*, Bucharest, 6-9 July, 2017
19. Dan Tufiş, Radu Ion. Part of Speech Tagging (2017). In Ruslan Mitkov (ed) *The Oxford Handbook of Computational Linguistics*, second edition, Oct 2017, DOI: 10.1093/oxfordhb/9780199573691.013.51, 25 pages
20. Markus Dickinson, Dan Tufiş (2017). Iterative Enhancement. In Nancy Ide, James Pustejovsky (eds): *Handbook of Linguistic Annotation*. Springer, pp. 257-276, 2017, ISBN: 978-94-024-0879-9, DOI: 10.1007/978-024-0881-2
21. Dan Tufiş (2016). An Overview of Data-Driven Part-of-Speech Tagging. *ROMANIAN JOURNAL OF INFORMATION SCIENCE AND TECHNOLOGY*, Volume 19, Numbers 1-2, 2016, 78-97
22. Maria Mitrofan, Daniela Gifu, Dan Tufiş, Dan Cristea (eds) (2016). The Proceedings of the 11th Edition of the **International Conference on Linguistic Resources and Tools for Processing Romanian Language.**, Mălini, Noiembrie 2016
23. Maria Mitrofan and Dan Tufiş (2016). Building and Evaluating the Romanian Medical Corpus. In *Proceedings of the 11th Edition of the International Conference on Linguistic Resources and Tools for Processing Romanian Language.*, Mălini, Noiembrie 2016



24. Rehm G., Uszkoreit H., Ananiadou S., Bel N., Bielevičienė A., Borin Lars, Branco, A., Budin G., Calzolari N., Daelemans W., Garabík R., Grobelnik M., García-Mateo C., Genabith J. van, Hajič J., Hernáez I., Judge J., Koeva S., Krek S., Krstev C., Lindén K., Magnini B., Mariani J., McNaught J., Melero M., Monachini M., Moreno A., Odijk J., Ogrodniczuk Maciej, Pezik Piotr, Piperidis Stelios, Przepiórkowski Adam, Rögnvaldsson Eiríkur, Rosner Mike, Sandford Pedersen B., Skadiņa I., Smedt K. de, Tadić M., Thompson P., Tufiş D., Váradi T., Vasiljevs A., Vider K., Zabarskaitė J. (2016). *The strategic impact of META-NET on the regional, national and international level*. In Language Resources and Evaluation, 50(2), pp. 351-374, DOI 10.1007/s10579-015-9333-4.
25. Ruxandra Cosma, Dan Cristea, Marc Kupietz, Dan Tufiş, Andreas Witt (2016). *DRuKoLA – Towards Contrastive German-Romanian Research based on Comparable Corpora*. In Proceedings of the 4th Workshop on Challenges in the Management of Large Corpora, pp 28-32, 2016, Portorož, Slovenia
26. Gilles Adda, Verginica Barbu Mititelu, Joseph Mariani, Dan Tufiş, Ioana Vasilescu (eds) (2016). *Errors by Humans and Machines in multimedia, multimodal and multilingual data processing*. Editura Academiei, 2016
27. Dan Tufiş., Verginica Barbu Mititelu, Elena Irimia, Ştefan Daniel Dumitrescu, Tiberiu Boroş (2016). *The IPR-cleared Corpus of Contemporary Written and Spoken Romanian Language*. In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016), European Language Resources Association (ELRA), ISBN: 978-2-9517408-9-1, 2016, pp. 2516-2521, Portorož, Slovenia
28. Dan Tufiş (coord.) (2016). "Siguranța Informatică – Protecția Cibernetică, Protecția Proprietății Intelectuale în Proiecte și în Publicarea Electronică", partea a 3-a, apărută în volumul 3 al STRATEGIEI DE DEZVOLTARE A ROMÂNIEI ÎN URMĂTORII 20 ANI, Ionel-Valentin Vlad (coord), vol. 3, Editura Academiei, ISBN 978-973-27-2706-5, pp. 199-256, 2016
29. Dan Tufiş (coord.) (2016). "Siguranța Informatică – Protecția Cibernetică, Protecția Proprietății Intelectuale în Proiecte și în Publicarea Electronică", partea a 2-a, apărută în volumul 2 al STRATEGIEI DE DEZVOLTARE A ROMÂNIEI ÎN URMĂTORII 20 ANI, Ionel-Valentin Vlad (coord), vol. 2, Editura Academiei, ISBN 978-973-27-2615-0, pp. 121-149, 2016
30. Dan Tufiş (coord.) (2015). "Siguranța Informatică – Protecția Cibernetică, Protecția Proprietății Intelectuale în Proiecte și în Publicarea Electronică", prima parte, apărută în volumul 2 al STRATEGIEI DE DEZVOLTARE A ROMÂNIEI ÎN URMĂTORII 20 ANI, Ionel-Valentin Vlad (coord), vol.1, Editura Academiei, ISBN 978-973-27-2556-6, pp. 129-168, 2015
31. Dan Tufiş et al. (2015). - CoRoLa Starts Blooming – An update on the Reference Corpus of Contemporary Romanian Language, in Proceedings of the Workshop on "Challenges in the management of large corpora", Lancaster, 20th July, 2015, pp. 5-10
32. Tiberiu Boroş, Ştefan Daniel Dumitrescu, Adrian Zafiu, Dan Tufiş, Verginica Barbu Mititelu, Paul Ionuț Văduva (2014). RACAI GEC – A hybrid approach to Grammatical Error Correction. In Eighteenth Conference on Computational Natural Language Learning (CoNLL 2014). Baltimore, USA, June 2014
33. Dan Tufiş, Radu Ion, Ştefan Daniel Dumitrescu, Dan Ştefănescu (2014). Large SMT data-sets extracted from Wikipedia. In Language Resources and Evaluation Conference (LREC 14). Reykjavik, Iceland, May 2014
34. Verginica Barbu Mititelu, Ştefan Daniel Dumitrescu, Dan Tufiş (2014). News about the Romanian Wordnet. In Proceedings of the 7th International Global WordNet Conference. Tartu, Estonia, 2014

National and International projects (coordinated for the Romanian part) for the last 5 years (2014-2019):

1. CURLICAT, European Project (CEF Telecom) Call 2019-1 Automated Translation Action n° 2019-EU-IA-0034
2. "European Language Grid" (2019-2022). Coordonator al Centrului de Competență din România
3. "European Language Resource Coordination - ELRC" (2015-2017), 30 de parteneri. Membru al Language Resource Board (LRB) organismul de conducere al ELRC
4. PARSEME (Parsing Multiword Expressions) –COST Action IC1207 (2013-2017), 30 countries, 29 official languages and 6 dialects
5. Interacademic project (Bulgaria-România) "Enhanced Knowledge Bases for Bulgarian and Romanian" (2015-2017).



6. DRuKoLA – Project funded by Alexander von Humboldt Foundation, in partnership with IDS-Mannheim, University of Bucureşti and ITI-Iaşi (2016-2018)
7. ELRC – European Language Resource Coordination (CEF SMART 2014/1074)– I am National Technological Anchor Point and member of the coordination board (2016-2020)
8. MARCELL - "Multilingual Resources for CEF.AT in the legal domain" (INEA/CEF/ICT/A2017/1565710), 2018-2020
9. ReTeRom - Resurse și tehnologii pentru dezvoltarea interfețelorom-mașina în limba română PN-III-P1-1.2-PCCDI-2018-0818, nr. 73PCCDI/2018, 2018-2020 (for this project I am the coordinator)
10. ROBIN – Roboții și Societatea: Sisteme Cognitive pentru Roboți Personali și Vehicule Autonome PN-III-P1-1.2-PCCDI-2018-0818, nr. 72PCCDI/2018, 2018-2020
11. CoRoLa –Priority Program of the Romanian Academy 2017-2019
12. CEF Automated Translation Toolkit for the EU Council Presidency (2018-2020)
13. NexusLinguarum – European Network for Research and Development in Linguistic Data Science (2019-2022)
14. ELG -European Language Grid (2019-2022)

Invited conferences/tutorials during the last 5 years (2014-2019):

1. Dan Tufiş Bente Maegard. 2014. Language Technology- A key Enabler for Industrial Applications; 2nd June 2014, Copenhagen
2. Dan Tufiş. 2014. Machine Translation – A Look into the Future. Invited talk at Craiova University, September 18th, 2014
3. Dan Tufiş. 2015. Data Hunger: Language Technologies-A Dream Coming True. Invited talk at Bulgarian Academy of Sciences, May 14th, 2015
4. Dan Tufiş. 2015. Challenges in building a publicly available reference corpus. (tutorial) EUROLAN 2015 Summer School, Sibiu, Romania, 13-25 July, 2015
5. Dan Tufiş. 2016. *COROLA – A Large Priority Project of the Romanian Academy on Romanian Language*. Invited talk at 9th International Conference on Interdisciplinary Studies (ICIS 2016), 10-11 iunie 2016, Constanța, Romania
6. Dan Tufiş, Alexandru Ceaușu Platforma eTranslation a MIE @ lucru. 2018. European Commission Workshop ELRC, 1 octombrie 2018
7. Dan Tufiş. 2018. **ELRC în Romania**, European Commission Workshop ELRC,1 octombrie 2018
8. Dan Tufiş. 2019. Language Technology and Digital Culture. European Conference on Exposing Online the European Cultural Heritage: the Impact of Cultural Heritage on the Transformation of the Society, Iași, 17-18 aprilie, 2019
9. Dan Tufiş. 2019. Resurse, tehnologii și rezultate în prelucrarea automată pentru limba română. Conferința Cercetării Științifice din Academia Română (CCSAR-2019), București, 6 mai 2019
10. Dan Tufiş. 2019. Language technologies and the challenges for Digital Single Market. IMPACT 2019, 21-22 May, Kracow, Poland
11. Dan Tufiş. 2019. Resources and Technologies for Language and Speech Processing of Romanian. Plenary talk to SPED 2019 International Conference, Timișoara 10-12 octombrie, 2019
12. Dan Tufiş. 2019. "Inteligența artificială și provocările ingineriei lingvistice". Plenary talk to the international workshop BACStud 2019 (5th edition), 17-19 October, 2019

Acad. TUFIȘ Dan Ioan



Data 27.10.2019