CURRICULUM VITAE

FLORENTINA HRISTEA

Contact: University of Bucharest, Faculty of Mathematics and Computer Science, Department of Computer Science, 14 Academiei Str., Bucharest, sector 1, cod 010014, Romania
e-mail: **fhristea@mailbox.ro**; **fhristea@fmi.unibuc.ro**
cell : 40722458143

Date of birth : April 7, 1960

Education:
- Habilitation in computer science, University of Bucharest, June 21, 2017;
- Ph.D. in mathematics, University of Bucharest, September 13, 1996;
- B.S. in mathematics and computer science, University of Bucharest, Faculty of Mathematics and Computer Science, June 1984.

Dissertations:
- Habilitation thesis: "Word Sense Disambiguation with Application in Information Retrieval" (2017)
-Ph.D. thesis: "Analysis of the Sensitivity of Some Statistical Models to the Presence of Outlying Observations (1996)

Foreign languages:
- English – spoken (very good); written (very good)
- French – spoken (good); written (good)

Positions held: Programmer (1984-1993);
Assistant Professor (1993-1997);
Lecturer (1997-2003);
Associate Professor (2003 – 2017);
Full Professor (2018 – present).

Current position: Full Professor at the University of Bucharest, Faculty of Mathematics and Computer Science, Department of Computer Science

TEACHING: gives the following (undergraduate and graduate) courses at the Faculty of Mathematics and Computer Science of the University of Bucharest, Dept. of Computer Science:

- Artificial Intelligence - for **undergraduate** students in computer science;
- Natural Language Processing - for **Master of Science** students in artificial intelligence;
- Special Topics in NLP and HLT – for **Master of Science** students in NLP.

RESEARCH FIELDS: Artificial Intelligence; **Specializations:** Knowledge Representation; Expert Systems; Computational Linguistics; Natural Language Processing and Human Language Technologies; Computational Statistics and Data Analysis applied in natural language processing and human language technologies

RESEARCH GRANTS: FULBRIGHT Research Fellow at **PRINCETON University, Cognitive Science Lab** (February 1 – July 31, 2004). Study of the semantic network of the English WordNet; co-author, together with George Miller and his group, of WordNet 2.1.

MEMBERSHIP IN PROFESSIONAL ASSOCIATIONS:

- **Elected member of ISI** (International Statistical Institute)
- Member of **GWA** (Global WordNet Association)
- Member of **ARIA** – the Romanian Association for Artificial Intelligence
- **Expert evaluator of the European Commission** in the field of Computational Linguistics, Natural language processing and human language technologies
- **Coeditor** of "Open Computer Science", published by De Gruyter Open, Germany (formerly "Central European Journal of Computer Science", published by Versita and Springer Verlag)
- **Topic editor** for "Mathematics" (MDPI)
- **Coeditor** of "International Journal on Information Technology", published by Praise Worthy Prize, Italy
- **Coeditor** of "International Journal on Computer and Communication Networks, Computational Intelligence and Data Analytics", published by Praise Worthy Prize, Italy
- **Editorial Review Board** of "Artificial Intelligence Research", published by Sciedu Press, Canada

INVITED PROFESSOR: University of Toulouse Paul Sabatier III, France, 2011

VISITING SCIENTIST at: University of Toulouse Paul Sabatier III, France; Institut de Recherche en Informatique de Toulouse, France; L'ecole Polytechnique "Polytech Montpellier", France; Heidelberg Institute for Theoretical Studies, Germany; Hamburg University, Germany; Princeton University, U.S.A.

INTERNATIONAL RESEARCH-DEVELOPMENT PROJECTS; PRINCIPAL INVESTIGATOR:

1. Participation in the international research - development project **DBR-MAT** ***(German – Bulgarian – Romanian Machine Aided Translation)***, funded by the Volkswagen Foundation, Germany. Research program including the University of Hamburg, Germany, the Bulgarian Academy of Science and the University of Bucharest, Faculty of Mathematics and Computer Science, for a period of two years (1996 – 1998). The general aim of DBR-MAT was the investigation and pilot implementation of a machine aided translation (MAT) system combining a knowledge-based approach with statistical methods for Natural Language Processing. **Principal Investigator of the Romanian team.**

2. Participation in the international research - development project **BALRIC-LING** ***(BALkan Regional Information Centers for awareness and standardization of LINGuistic resources and tools for advanced HLT applications)***, funded by the European Commission (IST – 2000- 26454; 2001-2003). The project involved the participation of The University of Bucharest, the Bulgarian Academy of Science, Sofia University, The University of Sheffield, U.K. and the Institute for Language and Speech Processing from Athens, Greece. Within the framework of this project the European Commission has organized two **Regional Information Centers** in the field of Natural Language Processing (NLP) and Human Language

Technologies (HLT) for the entire Balkan area. One of these centers was located at the University of Bucharest. **Principal Investigator of the Romanian project team and Coordinator of the Regional Information Center located in Romania**.

3. Participation in the **WordNet** project of the **Cognitive Science Laboratory of Princeton University**, U.S.A. (2004). Took part in the design of version 2.1 of WordNet (**WN 2.1**). Research project funded by "Advanced Research Development Activity", U.S.A. (Program AQUAINT, Phase 2, Contract No. NBCHC40012).

4. Participation in the European project **MULTI3GENERATION: MULTI-TASK, MULTILINGUAL, MULTI-MODAL LANGUAGE GENERATION (Multi3Generation)** – COST Action CA18231 (starting **2019**; lasting 4 years; 7 countries)

NATIONAL RESEARCH-DEVELOPMENT PROJECTS; PRINCIPAL INVESTIGATOR:

5. AUTOMATIC WORD SENSE DISAMBIGUATION - funded by the Ministry of Education, Research and Youth of Romania, The National Authority for Scientific Research (2009–2011; **Program "IDEAS"**); **Principal Investigator**

6. INFLUENTIAL OBSERVATIONS AND OUTLIERS IN LINEAR REGRESSION. DIAGNOSTIC CRITERIA - funded by the Ministry of Research and Technology, 2001; **Principal Investigator**

- MEMBER OF NUMEROUS OTHER PROJECT TEAMS FOR NATIONAL RESEARCH-DEVELOPMENT PROJECTS

SERVICES TO PROFESSIONAL ORGANIZATIONS:

- Served as a **reviewer** for international journal papers (Springer, De Gruyter) and as **evaluator** of European research-development projects in the field of Natural Language Processing (last evaluation: 23 Latvian research-development projects; September 2019);

- **Program Committee** and/or **reviewer** of international conferences and related events, among which:

  - The Summer 4-th International Conference on Knowledge Generation, Communication and Management: KGCM, Orlando, Florida, U.S.A., 2010;
  - The ACL-IJCNLP Conference (Association for Computational Linguistics), Singapore, 6-7 august 2009;
  - The 3rd International Conference on Knowledge Generation, Communication and Management: KGCM, Orlando, Florida, U.S.A., 2009;
  - The 4-th Balkan Conference in Informatics, Thessaloniki, Greece, 17-19 sept. 2009;
  - The 2nd International Conference on Knowledge Generation, Communication and Management: KGCM, Orlando, Florida, U.S.A., 2008;
  - The 2007 World Congress in Computer Science, Computer Engineering, and Applied Computing - WORLDCOMP '07, 25-28 iulie, 2007, Las Vegas, S.U.A.;

- the ROMAND workshop on Robust Methods in Analysis of Natural Language Data, in conjunction with the EACL 2006 conference (11-th Conference of the European Chapter of the Association for Computational Linguistics), Trento, Italy, April 3-7, 2006;
- the Student Research Workshop at the EACL 2006 conference (11-th Conference of the European Chapter of the Association for Computational Linguistics), Trento, Italy, April 3-7, 2006;
- the Language and Speech Infrastructure for Information Access in the Balkan Countries workshop, in conjunction with the international conference RANLP-2005 ("Recent Advances in Natural Language Processing"), Borovets, Bulgaria, September 21-23, 2005;
- ACL/SIGLEX SENSEVAL-3 ("Evaluating Word Sense Disambiguation Systems"), Barcelona, Spain, July 2004;
- Workshop on Balkan Language Resources and Tools, in conjunction with the Balkan Conference on Informatics, November 21, 2003, Thessaloniki, Greece.
- The 7-th Balkan Conference on Informatics, BCI 2015, Craiova, Romania, September 2-4, 2015.
- The 8-th International Global WordNet Conference (GWC 2016), Bucharest, Romania, January 27-30, 2016.
- Chair - The 2-nd World Congress on Electrical Engineering and Computer Systems and Science (EECSS '16), Budapest, Hungary, August 16-17, 2016.
- Chair - The 1-st Conference on Recent Advances in Artificial Intelligence (RAAI 2017), Bucharest, Romania, June 19-20, 2017.
- The 2-nd Conference on Recent Advances in Artificial Intelligence (RAAI 2018), Bucharest, Romania, June 25-27, 2018.
- The 3-rd Conference on Recent Advances in Artificial Intelligence (RAAI 2019), Bucharest, Romania, June 25-27, 2018.

AWARDS: Prize for outstanding results in scientific research, from the University of Bucharest, in 2007.

SELECTED PUBLICATIONS:

I. BOOKS:

1. The Naïve Bayes Model for Unsupervised Word Sense Disambiguation. Aspects Concerning Feature Selection, SpringerBriefs in Statistics Series, Springer, 2012 (in English).
2. Issues of Knowledge Representation with Applications in Natural Language Processing, Bucharest, University of Bucharest Publishing House, 2011 (in Romanian).

3. Introduction to Natural Language Processing, Bucharest, University of Bucharest Publishing House, 2010 (in Romanian).

4. Search and Knowledge Representation in Artificial Intelligence. Theory and Applications (with M.F. Balcan). Bucharest, University of Bucharest Publishing House, 2005 (in Romanian).

5. An Introduction to Natural Language Processing with Applications in Prolog. Bucharest, University of Bucharest Publishing House, 2000 (in Romanian).

II. TEXTBOOKS AND MANUALS:

1. Issues of Search and Knowledge Representation in Artificial Intelligence (with M.F. Balcan). Bucharest, University of Bucharest Publishing House, 2004 (in Romanian).

2. Norton Commander 4.0. Norton Utilities 7.0 and 8.0 (with D. Enachescu and R. Balint). Bucharest, Technical Publishing House, 1994 (in Romanian).

3. The Norton Commander System of Programs. Bucharest, Technical Publishing House, 1992 (in Romanian).

III. BOOK CHAPTERS; EDITOR:

1. "On the Semiautomatic Generation of WordNet Type Synsets and Clusters with Special Reference to Romanian", in: Building Awareness in Language Technology (editors Florentina Hristea and Marius Popescu). Bucharest, University of Bucharest Publishing House, pp. 113-140, 2003 (in English).

2. "A Dependency Grammar Approach to Syntactic Analysis with Special Reference to Romanian" (with M. Popescu), in: Building Awareness in Language Technology (editors Florentina Hristea and Marius Popescu). Bucharest, University of Bucharest Publishing House, pp. 9-34, 2003 (in English).

IV. ENCYCLOPEDIA ENTRIES:

1. **Outlier Detection, Hristea Algorithm**. Encyclopedia of Statistical Sciences, 2nd ed., Vol. 9; N. Balakrishnan, Campbell B. Read, and Brani Vidakovic, Editors-in-Chief. John Wiley & Sons, Inc., New York, p. 5885-5886, 2005.

2. **Statistical Natural Language Processing**. International Encyclopedia of Statistical Sciences. Miodrag Lovric, Editor-in-chief. Springer, U.S., Part 19, p.1452-1453, DOI: 10.1007/978-3-642-04898-2_82, 2011.

V. PAPERS IN REFEREED (PEER REVIEW) JOURNALS:

1. Dascalu, S., Hristea, F., Towards a Benchmarking System for Comparing Automatic Hate Speech Detection with an Intelligent Baseline Proposal. Mathematics, 2022, 10(6), 945; https://doi.org/10.3390/math10060945 In: Natural Language Processing (NLP) and Machine Learning (ML) - Theory and Applications, Special Issue.

2. Feature selection for spectral clustering: to help or not to help spectral clustering when performing sense discrimination for IR? (with A. Chifu). Open Computer Science, Vol. 8, No. 1, 2018, pp. 218-227.

3. The long road from performing word sense disambiguation to successfully using it in information retrieval: An overview of the unsupervised approach (with M. Colhon). Computational Intelligence, Vol. 36, No. 3, 2020, pp. 1026-1062.

4. Word sense discrimination in information retrieval: A spectral clustering-based approach (with A. Chifu, J. Mothe, M. Popescu). Information Processing and Management, Vol. 51, No. 2, 2015, pp. 16-31.
5. Unsupervised Word Sense Disambiguation with N-Gram Features (with D. Preotiuc). Artificial Intelligence Review, Vol. 41, Nr. 2, 2014, pp. 241-260.
6. Feeding Syntactic Versus Semantic Knowledge to a Knowledge-lean Unsupervised Word Sense Disambiguation Algorithm with an Underlying Naïve Bayes Model (with M. Colhon). Fundamenta Informaticae, Vol. 119, Nr. 1, 2012, pp. 61-86.
7. State of the Art Versus Classical Clustering for Unsupervised Word Sense Disambiguation (with M. Popescu). Artificial Intelligence Review, Vol.35, No.3, DOI 10.1007/s10462-010-9193-7, 2011, pp. 241-264.

8. Syntactic Relations in the Context of Performing Word Sense Disambiguation (with M.Colhon). Journal of Knowledge, Communications and Computing Technologies, Vol. II, No. 2, 2010, pp. 1-9.

9. Performing Word Sense Disambiguation at the Border Between Unsupervised and Knowledge-Based Techniques (with M. Popescu and M. Dumitrescu). Artificial Intelligence Review, Vol. 30, No. 1-4, DOI 10.1007/s10462-009-9117-6, 2009, pp. 67-86.

10. Recent Advances Concerning the Usage of the Naïve Bayes Model in Unsupervised Word Sense Disambiguation. International Review on Computers and Software, Vol. 4, No. 1, 2009, pp. 58-67.

11. Adjective Sense Disambiguation at the Border Between Unsupervised and Knowledge-based Techniques (with M. Popescu). Fundamenta Informaticae, Vol. 91, No. 3-4, 2009, pp. 547-562.

12. On the Usage of the Naïve Bayes Model in Unsupervised Word Sense Disambiguation. International Review on Computers and Software, Vol.3, no. 6, 2008, pp. 572-578.

13. Semiautomatic Generation of WordNet Type Synsets and Clusters Using Class Methods. An Overview. Revue Roumaine de Linguistique, VOL. LII, Issues 1-2, 2007, pp. 97-133.

14. Towards Building a WordNet Noun Ontology (with G. Miller). Revue Roumaine de Linguistique, VOL. LI, Issues 3-4, 2006, pp. 405-413.

15. On Some Romanian Text Summarization Contributions (with M. Popescu). Revue Roumaine de Linguistique, VOL. LI, Issues 3-4, 2006, pp. 507-509.

16. WordNet Nouns: Classes and Instances (with G. Miller). Computational Linguistics, Volume 32, Number 1, 2006, pp. 1-3.

17. On the semiautomatic generation of verb synsets in languages other than English. Annals of the University of Bucharest, ANO LII, 2003, pp. 75-86.

18. On the semiautomatic generation of WordNet type synsets and clusters. Journal of Universal Computer Science (J.UCS), Vol.8, no.12, 2002, pp .1047-1064.

19. On the use of MAICE for the detection of outliers in a linear regression model. Mathematical Reports Vol.4 (54), No.2, 2002, pp. 177-190.

20. On the diagnosis of influential observations in multidimensional linear regression. Mathematical Reports, No.4, 1999, pp. 545-550.

21. On WG syntactic analysis with special reference to Romanian. Annals of the University of Bucharest, No.1, 1998, pp. 59-69.

22. A Word Grammar approach to syntactic analysis with special reference to Romanian (with M. Popescu). Annals of the University of Bucharest , Special Issue, 1998, pp. 101-113 .

23. On multiple response outliers of additive nature in a linear regression model. Revue Roumaine de Mathematiques Pures et Appliquees, No.3-4, 1998, pp. 393-405.

24. An algorithm for the detection of outliers in the case of the normal distribution. Mathematical Reports, No.5-6, 1997, pp. 339-342.

25. On single distribution outlying pairs in a linear regression model. Annals of the University of Bucharest, XLVI, 1997, pp. 27-34.

26. On single carrier outliers in a linear regression model. Foundations of Decision and Computing Sciences, Vol.21, No.3, 1996, pp. 125-136.

27. On single additive outliers in the linear regression model. Annals of the University of Bucharest, No.2, 1996, pp. 53-64.

28. On the masking effect concerning some tests for single upper outliers in the case of an exponential sample. Foundations of Decision and Computing Sciences, Vol. 20, No.1, 1995, pp. 37-48.

29. On the performances of an outlier test in the case of the exponential distribution (with D. Enachescu & M. Dumitrescu). Computational Statistics and Data Analysis, No.17, 1994, pp. 119-127.

30. A codification of column-convex polyominoes which generates a regular language. Bulletin of the European Association for Theoretical Computer Science, No.50, 1993, pp.197-208.

31. On the performances of a test for a single upper outlier in an exponential sample (with D. Enachescu). Annals of the University of Bucharest, No.2, 1992, pp. 61-69.

32. On stepwise procedures for the rejection of outliers in a normally distributed population (with R. Niculescu & D. Enachescu). Annals of the University of Bucharest, No.1-2, 1990, pp. 14-26.

May 15, 2022

Florentina Hristea