Assistant Professor | Maître de conférences HDR
Centre de Recherches Interdisciplinaires et Transculturelles (CRIT)
Université de Bourgogne Franche-Comté, Besançon, France
✉ *iana.atanassova@univ-fcomte.fr*
🖅 *http://tesniere.univ-fcomte.fr/iana*

Iana Atanassova

Current position

Since 09/2014   **Assistant Professor**, *Université de Bourgogne Franche-Comté*, Besançon, France.
Centre de Recherches Interdisciplinaires et Transculturelles (CRIT)

Research topics

My research is in the field of Natural Language Processing (NLP) and more specifically full-text scientific paper processing. It is part of the Research topic "Science and textuality" in CRIT, which develops and applies linguistic approaches in NLP in order to analyse scientific discourse. I explore the problems of semantic annotation, ontology population, linguistic modeling and information extraction from scientific texts. I have used NLP methods to process scientific corpora with applications in bibliometrics and recommender systems. I participate in projects on text mining for competitive intelligence and personal data identification and representations. My recent research projects consist in studying the rhetorical structure of articles, and the expression of uncertainty as integral part of the process of the contruction of scientific knowledge.

Projects

2019–2021   Principal Investigator (PI) of the project DecRIPT - *"Detection of the Representations of Personal Data in Texts"*, funded by Interreg France-Switzerland 2014-2020 (307 313 euros) and Communauté du Savoir (9 409 euros).
  ◦ Academic partners: HEG (Switzerland), HEG Arc (Switzerland)
  ◦ Private partners: ERDIL (France), Global Data Excellence (Switzerland)
  ◦ Objectives: Building tools for the robust processing of personal data in texts to facilitate compliance with GDPR (http://tesniere.univ-fcomte.fr/projet-decript/).

2020–2024   Supervisor of project EMONTAL - *"Extraction and Ontology Modeling of Subjects and Places for the Exploitation of the Documentary Funds of Bourgogne Franche-Comté"*: supervision of a PhD thesis, funded by Région Bourgogne Franche-Comté (http://tesniere.univ-fcomte.fr/projet-emontal/).

2016–2021   Coordinator of project *"Researcher-entrepreneur"*: supervision of a PhD thesis and a post-doctoral fellow on "Linguistic analysis and automatic extraction of semantic relations in Arabic", funded by Université de Bourgogne Franche-Comté, with the objective of technology transfer to a start-up.

2020–2023   Participant in the project *TheoScit*, funded by French ANR Agency.

2016–2018   Participant in the project *WebSO+: Competitive Intelligence Platform*, funded by Interreg France-Switzerland 2014-2020. Development of a platform for strategic technological and competitive intelligence and e-reputation integrating semantic processing, text classification and sentiment analysis.

2015–2017   Participant in the projet SARS - "System for assisted scientific writing in the biomedical domain", funded by Région Franche-Comté. Tools for scientific writing using lexical databases and Information Retrieval.

Research supervision

Post-doctoral supervision

Youcef Morsi (09/2020-08/2021): PostDoc resercher-entrepreneur, with the objective to create of a start-up that exploits research results obtained during the PhD thesis.

François-C. Rey (01/2020-12/2020): as part of DecRIPT project.

PhD thesis supervision

Youcef Morsi (2016-2020): Thesis "Information extraction in Arabic". We design a linguistically-motivated approach for the morphosyntactic and semantic processing using the specific structure of Arabic. The results are compared with those of state of the art parsers such as Universal Dependency parcer. Funded by project researcher-entrepreneur.

François-C. Rey (2016-): Thesis "Categorisation of speculative sentences in scientific texts". We study the expression of speculations and uncertainty in scientific papers and design methods for the automatic extraction and classification of text segments in a corpus of scientific papers on climate change.

Séda Ozturk (2017-): Thesis "Analysing Scientific Papers for the Extraction and Characterization of Datasets". In the context of Open Science, we design methods to extract and classify information on datasets and results obtained by using them in scientific papers.

Salah Yahiaoui (2019-): Thesis "Information extraction of spatio-temporal data from scientific texts". The objective is to exploit the full text of papers in order to produce spatial and temporal visual representation of research results. Funded by French Ministry for Education and Research.

Nicolas Guterhlé (2020-): Thesis "Extraction and ontological modelling of actors and places for the exploitation of the documentary funds of the Bourgogne Franche-Comté Region". Project EMONTAL, funded by Bourgogne Franche-Comté Region.

MSc thesis supervision

L. Lograda (2016), F.-C. Rey (2016), M. Delhotal (2016), Y. Morsi (2016) N. Guterhlé (2018), C. El Cadi (2018), I. Hatira (2018), L. Annebi (2018) S. Yahiaoui (2019), A. Calatayud (2019), R. Hasnaoui (2020), J. Bergoend (2020)

Service

Course management

Since 09/2020, manager of the department of Natural Language Processing, Université de Franche-Comté, France.

Since 09/2017, course manager of the specialty *"Natural Langauge Processing (NLP)"* of Master of Science "Languages and Foreign Cultures (LLCER)", Université de Franche-Comté, France.

ISO/AFNOR commission

Since 09/2014, member of the ISO/AFNOR commission, work group ISO/TC 37/SC 4 on terminology and linguistic ressource managment. National referent for ISO 24617 "Semantic Annotation Framework".

Organisation of workshops and scientific events

○ Workshop "Mining Scientific Papers: Computational Linguistics and Bibliometrics (CLBib)", 16th International Conference on Scientometrics and Informetrics (ISSI), Wuhan, China, 2017 (https://easychair.org/cfp/CLBib2017).

○ Workshop "Traitement Automatique des Langues Slaves (TASLA)", conference TALN 2015, France (http://tesniere.univ-fcomte.fr/tasla/).

○ Workshop "Mining Scientific Papers: Computational Linguistics and Bibliometrics(CLBib)", 15th International Society of Scientometrics and Informetrics Conference (ISSI), Istanbul, Trukey, 2015.

○ Workshop "Analyses logométriques des revues québécoises : traitements et visualisations", colloque "Relire les revues québécoises : histoires, formes et pratiques", Montreal, Canada, October 2015.

Advisory boards

○ Since 2016, elected member of the advisory board of the Doctoral School LECLA, Université de Bourgogne Franche-Comté (UBFC), France.

○ Since 2018, member of the advisory board of the *"Natural Language Processing"* series, John Benjamins publishing (https://benjamins.com/catalog/nlp).

○ Since 2019, member of the advisory board of the SIPS platform - *"Information System on Phylosophy of Science"* (https://sips.univ-fcomte.fr/).

Editorial work, reviews, programme committees

○ Editor of the Research Topic *"Mining Scientific Papers: NLP-enhanced Bibliometrics"* in "Frontiers in Research Metrics and Analytics" (https://www.frontiersin.org/research-topics/7043/mining-scientific-papers-nlp-enhanced-bibliometrics).

○ Reviewer for *Journal of Infometrics, Aslib Journal of Information Management, Journal of Natural Language Engineering, Computational Intelligence, Scientometrics*.

○ Member of the programme committees of international workshops and conferences: Bibliometrics and Information Retrieval (BIR, since 2014); Workshop On Mining Scientific Publications (WOSP-2018), LREC 2018; Recent Advances in Natural Language Processing (RANLP-2017); The International Florida Artificial Intelligence Research Society Conference (FLAIRS); The American Association for the Advancement of Science (AAAS) 2015 Annual Meeting; Bibliometric-enhanced Information Retrieval and Natural Language Processing for Digital Libraries (BIRNDL, since 2016); The International ACM Conference on Management of computational and collective intElligence in Digital EcoSystems (MEDES 2014 et 2015); Workshop on Scholarly Web Mining (SWM 2017); Computational Linguistics in Bulgaria (CLIB 2018, 2019).

Grants and prizes

2016: Nomination for the annual International Society for Scientometrics and Infometrics (ISSI) Paper of the year award: Marc Bertin, Iana Atanassova, Vincent Larivière, and Yves Gingras. The Invariant Distribution of References in Scientific Papers. JASIST, 2016

2017: First prize on the InovHackTion hackathon organised by the Air Force France on the Intelligent analysis of documents (http://actu.univ-fcomte.fr/article/le-traitement-automatique-des-langues-seduit-larmee-005704).

2019–2023: Grant for doctoral supervision and research (PEDR), France.

Science Dissemination

TDM Stories "The structure of papers": OpenMinTeD Interview (http://openminted.eu/tdm-stories-structure-papers/)

TEDonnées: Workshop for the general public "Emerging Data Processing Tools (TEDonnées)", in 2017 and 2018, Besançon, France (http://tesniere.univ-fcomte.fr/tedonnees/). These workshops aim to inform the general public and also promote the interactions between students in NLP, researchers and firms that work in this domain. The first two editions were on the topics of data mining, e-reputation and smart cities.

Teaching activity

Non-exhaustive list.

Since 09/2014 **In MSc "Languages and Foreign Cultures" (LLCER), specialty Natural Language Processing**, *Université de Franche-Comté (UFC)*, Beançon, France.
Courses (lectures & tutorial classes): Methodology in NLP, Computer Programming for NLP, Information extraction and IR, Etymology, Cognitive Methods and ML, Software engineering for NLP.

Since 09/2014 **In MSc "Languages and Foreign Cultures" (LLCER), all specialties**, *Université de Franche-Comté (UFC)*, Beançon, France.
Courses (lectures & tutorial classes): Collaborative work and project management, Information Retrieval

Since 09/2014 **In BSc "Languages and Foreign Cultures", specialty Natural Language Processing**, *Université de Franche-Comté (UFC)*, Beançon, France.
Courses (lectures & tutorial classes): Computational Linguistics, Formal linguistics and informatics, Introduction to computer programming for NLP.

2019 & 2020 **InURFIST de Lyon**, France.
Courses: "Introduction to scientific writing with LaTeX"

2011–2012	**In MSc *"French Language and Applications"***, *Paris-Sorbonne University*, Paris, France. Courses: "Introduction to Informatics and Natural Language Processing" (lectures & tutorial classes).
2009–2011	**In BSc *"French Language and Informatics"***, *Paris-Sorbonne University*, Paris, France. Courses: "Logics and Introduction to Informatics" (tutorial classes); "General Mathematics and Analysis - 1" (lectures & tutorial classes); "Preparation for the Certificate of Informatics and Internet (C2i) - level 1" (lectures & tutorial classes).
2009–2011	**In MSc *"Informatics and Language Engineering for Information Management"***, *Paris-Sorbonne University*, Paris, France. Courses: "Methodology and scientific reasoning" (tutorial classes).
2006–2009	**In MSc *"Philosophy and Sociology"***, *Paris-Sorbonne University*, Paris, France. Courses: "Informatic tools" (tutorial classes).

Work experience

02/2017	**International mobility Erasmus+ STA**, *University of Wolverhampton*, Wolverhampton, UK, Research Institute in Information and Language Processing (RIILP), Research Group in Computational Linguistics.
01/2014–08/2014	**Post-doctoral fellow: Semantic processing of corpora in humanities and social sciences**, *Concordia University*, Montreal, Canada. Research director: Pr. Jean-Philippe Warren Objectives: Linguistic analyses for the exploitation of the documentary collections of the Bibliothèque et Archives nationales du Québec (BAnQ); Identification of new concepts, semantic relations and definitions in journal articles; Diachronic analysis of corpora to produce data for the sociological study of ideas and the evolution of concepts.
03/2012–06/2013	**Research and Development manager**, *MyScienceWork*, France/Luxembourg. Implementation of a search engine for scientific papers for the MyScienceWork social network.
09/2009–08/2011	**Teaching and Research Assistant**, *Paris-Sorbonne University, Institute of Applied Human Sciences*, Paris, France.

Education

December 2015	**Habilitation to Direct Research (HDR), specialty "Natural Language Processing"**, *University of Franche-Comté, Centre Tesnière*, Besançon, France. Title: "Qualitative and quantitative analysis of scientific discourse. Applications to Information Extraction, Information Retrieval and the Semantic Web" Supervisor: Pr. Sylviane Cardey (University of Franche-Comté, France) Jury: Pr. R. Mitkov (University of Wolverhampton, UK); Pr. Ch. Roche (Université de Savoie, France); Pr. B. K. Bogacki (Université de Varsovie, Poland); Pr. L. Da Sylva (Université de Montréal, Canada); Pr. J.-P. Desclés (Paris-Sorbonne University, Paris); Pr. P. Pognan (INALCO, Paris).
January 2012	**PhD in "Mathemacs, Informatics and Applications to Human Sciences"**, *Paris-Sorbonne University, Facutly of Letters, LaLIC-STIH laboratory*, Paris, France. Title : "Exploiting semantic annotations for information retrieval and text navigation" Supervisor: Pr. J.-P. Desclés (Paris-Sorbonne University, Paris) Jury: Pr. Th. Poibeau (CNRS, LaTTiCe); Pr. M. Hassoun (ENSSIB); Dr. Ch. Harbulot (EGE); Dr. B. Djioua (STIH, Paris-Sorbonne).
2005–2006	**MSc "Information and Communication"**, *Paris-Sorbonne University, Facutly of Letters*, Paris, France. MSc thesis title: "Semantic annotations of texts in French and in Bulgarian with the Excom engine. Automatic syntheses."

2004–2005   **MSc "Computational Linguistics"**, *Sofia University, Faculty of Classical and Modern Philology*, Sofia, Bulgaria.

2002–2005   **BSc in English Studies**, *Sofia University, Faculty of Classical and Modern Philology*, Sofia, Bulgaria.

2000–2004   **BSc in Mathematics**, *Sofia University, Faculty of Mathematics and Informatics*, Sofia, Bulgaria.