

PERSONAL INFORMATION

Debora Nozza

✉ debora.nozza@unibocconi.it

🌐 <https://www.deboranozza.com/>

EMPLOYMENT

- Sep 2022 - present **Assistant Professor in Computing Sciences**, *Bocconi University*, Milan.
My research focuses on advancing fairness and the societal relevance of Natural Language Processing models. Primarily, I have played a pivotal role in detecting and mitigating hate speech within multilingual contexts.
- Feb 2020 - Aug 2022 **Research Fellow**, *Bocconi University*, Milan.
Supervisor: Prof. Dirk Hovy.
Working for the ERC Starting Grant INTEGRATOR, Incorporating Demographic Factors into Natural Language Processing Models.
- Jan 2018 - Jan 2020 **Research Fellow in Computer Science**, *University of Milano-Bicocca*, Milan.
Title: "Deep Learning Representation for Large Attributed Graphs".
Supervisors: Prof. Enza Messina and Prof. Elisabetta Fersini.
My research project aimed at designing and developing novel Deep Learning unsupervised models for learning an embedding representation from large heterogeneous attributed graphs, considering both the structure of the graph and the attributes associated with each node.

EDUCATION

- Dec 2014 - Oct 2018 **PhD in Computer Science**, *University of Milano-Bicocca*, Milan.
Thesis Title: "Deep Learning for Feature Representation in Natural Language Processing".
Supervisors: Prof. Enza Messina and Prof. Elisabetta Fersini.
My research activity focused on the exploitation and improvements of data representation obtained by Deep Learning techniques in order to address Natural Language Processing tasks, i.e. Sentiment Analysis, Named Entity Recognition and Linking, Hate Speech detection.
- Feb 2017 - May 2017 **Visiting PhD Student**, *INRIA*, Sophia Antipolis – Méditerranée.
Visiting PhD student in the WIMMICS group under the supervision of Dr. Fabien Gandon.
- 2011 - 2014 **Master's Degree in Computer Science**, *University of Milano-Bicocca*, Milan.
110/110 summa cum laude.
Thesis Title: "Learning representation for sentiment classification".
Supervisors: Prof. Enza Messina and Prof. Patrick Gallinari (University Pierre and Marie Curie, Paris).
- July 2013 - Feb 2014 **Visiting Master Student**, *University Pierre and Marie Curie*, Paris.
Internship Student at Laboratoire d'Informatique de Paris 6 (LIP6) in the Machine Learning and Information Access research group, under the supervision of Prof. Patrick Gallinari.
- 2008 - 2011 **Bachelor's Degree in Computer Science**, *University of Milano-Bicocca*, Milan.
Thesis Title: "Experimenting techniques of constrained inference for Information Extraction".
Supervisors: Prof. Enza Messina and Prof. Elisabetta Fersini.

RESEARCH ACTIVITIES

Awards and Scholarships

- **ERC Starting Grant** PERSONAE (Personalized and Subjective approaches to Natural Language Processing), **PI** (1.5M EUR).
- **Grant** from Fondazione Cariplo for MONICA (MONItoring Coverage, Attitudes and Accessibility of Italian measures in response to COVID-19), **PI** (120,000 EUR).
- Google Cloud Platform credits (5,000 EUR).

Talks and Seminars

- **Workshop Speaker** at *NCRM/Exeter Computational Communication Methods Spring School* (Exeter, UK, Apr 2023).
- **Invited Talk** at *GESIS Leibniz Institute for the Social Sciences* (Cologne, Germany, Dec 2022).
- **Invited Talk** at *INRIA Paris* (Paris, France, Oct 2022).
- **Invited Talk** at *Fondazione Bruno Kessler* (Trento, Italy, Oct 2022).
- **Invited Talk** “Roadmap to universal hate speech detection” at *Fairness in AI - Bocconi University* (Milan, Italy, June 2022).
- **Invited Talk** “Gli algoritmi hanno pregiudizi? Le conseguenze dell’intelligenza artificiale” at *University of Milano Bicocca* organized by the Department of Business and Law (Online, December 3, 2021)
- **Podcast** participation “The Potential Pitfalls of Algorithms” in the *Bocconi Clarity in a Messy World* podcast. (Mar 10, 2021).
- **Tutorial** “JAX for Natural Language Processing” at the *M2L Summer School in collaboration with DeepMind*. (Online, January 4, 2021).
- **Podcast** participation “Etica e bias nell’intelligenza artificiale” in the *Smarter Podcast: the Italian podcast on Artificial Intelligence*. (Dec 10, 2020).
- **Invited Talk** “Directions and Challenges for Misogyny Detection: from corpus construction to prediction models” at the *Alan Turing Institute* organized by the Social Data Science Interest Group (Online, November 20, 2020).
- **Invited Talk** “Making Sense of Language-Specific Contextual Representation Models” at the *Workshop #1 on Named Entity Recognition in the Context of Machine Translation* organized by the Contracting Authority DG CNECT of European Commission and the company Cross-Lang (Online, June 5, 2020).
- **Invited talk** “Deep learning in natural language processing” at the conference of *AI, Machine Learning and Sentiment Analysis Applied to Finance* (London, UK, 2016).

Teaching experiences *Bocconi University, Milan.*

- Professor for the course of “*Deep Learning For Computer Vision*” for the master degree in Data Science and Business Analytics. Overall hours: 36. Period: Sep 2022 - Dec 2022.
- Professor for the course of “*Machine Learning And Artificial Intelligence*” for the bachelor degree in Mathematical and Computing Sciences for Artificial Intelligence. Overall hours: 30. Period: Feb 2023 - Apr 2023.
- Professor for the course of “*Computer Science*” for the bachelor degree. Overall hours: 40. Period: Apr 2023 - Jun 2023.

University of Trieste, Trieste.

- Adjunct Professor for the course of “*Natural Language Processing*” for the master degree in Data Science and Scientific Computing. Overall hours: 40. Period: May 2020 - Jul 2020; Mar 2021 - May 2021; Mar 2022 - Jul 2022.

University of Milano-Bicocca, Milan.

- Teaching Assistant for the course of “Data Analytics” for the master degree in Computer Science. Overall hours: 24. Period: Feb 2020 - Jul 2020; Feb 2019 - Jul 2019.
- Teaching Assistant for the course of “Data Analytics” for the master degree in Computer Science. Overall hours: 12. Period: Feb 2018 - Jul 2018; Feb 2017 - Jul 2017.
- Teaching Assistant for the course of “Probabilistic Models for Decision Making” for the master degree in Computer Science. Overall hours: 30. Period: Sep 2017 - Jan 2018.
- Teaching Assistant for the course of “Laboratory of programming in Java” for the bachelor degree in Mathematics. Overall hours: 30. Period: Sep 2016 - Jan 2017; Sep 2015 - Jan 2016.
- Co-Supervisor of 18 bachelor’s and master’s degree thesis in Computer Science.

Industrial Research Collaboration

- Collaboration with Rewire API for a multilingual hate speech detection resource (Feb 2022 - Apr 2022).
- Web service and Social Media Manager for DMI Unit at Bocconi University (May 2020 - Apr 2021).
- Collaboration with Opt Out Tools for creating a browser extension that hides misogynistic harassment from the Twitter feed (Mar 2020 - Mar 2021).
- Collaboration with Clinical Insights Ltd for the project “Patient profiling for drug compliance” (Jan 2018 - Feb 2018).
- Collaboration with Consorzio Milano Ricerche for the project regarding the development, exploration and writing of “Proof of Concept: Social Media Sentiment Analytics (Brexit)” (Nov 2016 - Mar 2017).
- Collaboration with Consorzio Milano Ricerche for the project regarding the analysis and presentation of “Methods for the analysis of longitudinal prescription data” (Nov 2014 - Dec 2014).

SERVICE ACTIVITIES

Editorial Activities

- *Organizer* of the 5th Workshop on Gender Bias in Natural Language Processing, ACL 2024.
- *Organizer* of the 8th Workshop on Online Abuse and Harms (WOAH), NAACL 2024.
- *Organizer* of the 7th Workshop on Online Abuse and Harms (WOAH), ACL 2023.
- *Organizer* of the Data Challenge Track Workshop at the 17th International Conference on Web and Social Media (AAAI ICWSM 2023).
- *Organizer* of the 6th Workshop on Natural Language for Artificial Intelligence (NL4AI), AI*IA. Udine, Italy, 30 Nov, 2022.
- *Area Chair* at ACL 2023, EACL 2023, EMNLP 2023.
- *Area Chair* at the 14th Women in Machine Learning Workshop (WiML) at NeurIPS 2019.
- *Reviewer* for Natural Language Engineering - Cambridge University Press, Computers & Operations Research - Elsevier, Expert Systems with Applications - Elsevier, Information Processing & Management - Elsevier, Journal of Web Semantics - Elsevier, Online Social Networks and Media - Elsevier, EPJ Data Science - Springer, and Information - MDPI.
- *Reviewer* for ACL, EACL, EMNLP, AACL.

Organization of International Shared Tasks

- HODI: Homotransphobia Detection in Italian (HODI) at the 8th evaluation campaign of Natural Language Processing and Speech tools for Italian (EVALITA’23). Italy, September 7-8, 2023.
- EMit: Categorical Emotion Detection in Italian Social Media at the 8th evaluation campaign of Natural Language Processing and Speech tools for Italian (EVALITA’23). Italy, September 7-8, 2023.
- AMI: Automatic Misogyny Identification (AMI) at the 7th evaluation campaign of Natural Language Processing and Speech tools for Italian (EVALITA’20). Italy, December 2-3, 2020.
- SemEval-2019 Task 5: HatEval: Multilingual Detection of Hate Speech Against Immigrants and Women in Twitter at the 13th International Workshop on Semantic Evaluation (SemEval-2019). Minneapolis, USA, June 6-7, 2019.
- AMI: Automatic Misogyny Identification (AMI) at the 6th evaluation campaign of Natural Language Processing and Speech tools for Italian (EVALITA’18). Italy, December 10-12, 2018.

SELECTED PUBLICATIONS

Please refer to the following websites for the complete publication lists: www.deboranozza.com/publication/ and <https://scholar.google.it/citations?user=AKi-UWQAAAAJ>

- G. Attanasio, F.M. Plaza-del-Arco, , **D. Nozza**, and A. Lauscher. “*A Tale of Pronouns: Interpretability Informs Gender Bias Mitigation for Fairer Instruction-Tuned Machine Translation*”. In Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (Long Papers). Singapore, Dec, 2023.
- A. Lauscher, **D. Nozza**, E. Miltersen, A. Crowley, and D. Hovy. “*What about “em”? How Commercial Machine Translation Fails to Handle (Neo-)Pronouns*”. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Toronto, Canada, Jul, 2023.
- **D. Nozza** and D. Hovy. “*The State of Profanity Obfuscation in Natural Language Processing Scientific Publications*”. In Findings of the Association for Computational Linguistics: ACL 2023. Toronto, Canada, Jul, 2023.
- G. Ruggeri and **D. Nozza**. “*A Multi-dimensional study on Bias in Vision-Language models*”. In Findings of the Association for Computational Linguistics: ACL 2023. Toronto, Canada, Jul, 2023.
- F. Bianchi, P. Kalluri, E. Durmus, F. Ladhak, M. Cheng, **D. Nozza**, T. Hashimoto, D. Jurafsky, J. Zou, A. Caliskan. “*Easily Accessible Text-to-Image Generation Amplifies Demographic Stereotypes at Large Scale*”. In Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency: FAccT '23. Chicago, USA, June, 2023.
- G. Attanasio, E. Pastor, C. Di Bonaventura, **D. Nozza**. “*ferret: a Framework for Benchmarking Explainers on Transformers*”. In Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations (EACL). Dubrovnik, Croatia, May, 2023.
- P. Röttger, **D. Nozza**, F. Bianchi, D. Hovy. “*Data-Efficient Strategies for Expanding Hate Speech Detection into Under-Resourced Languages*”. In Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing: EMNLP 2022. Abu Dhabi, United Arab Emirates, Dec, 2022.
- G. Attanasio, **D. Nozza**, D. Hovy, E. Baralis. “*Entropy-based Attention Regularization Frees Unintended Bias Mitigation from Lists*”. In Findings of the Association for Computational Linguistics: ACL 2022. Dublin, Ireland, May, 2022.
- F. Bianchi, S. Terragni, D. Hovy, **D. Nozza**, E. Fersini. “*Cross-lingual Contextualized Topic Models with Zero-shot Learning*”. Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume. Online, 2021.
- **D. Nozza**, “*Exposing the limits of Zero-shot Cross-lingual Hate Speech Detection*”. Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers). Online, 2021.
- **D. Nozza**, F. Bianchi, D. Hovy. “*HONEST: Measuring Hurtful Sentence Completion in Language Models*”. Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers). Online, 2021.
- **D. Nozza**, C. Volpetti, E. Fersini. “*Unintended Bias in Misogyny Detection*”. In Proceedings of the 2019 IEEE/WIC/ACM International Conference on Web Intelligence (WI 2019). Thessaloniki, Greece, October 14-17, 2019. **Runner up for the Best Paper**.