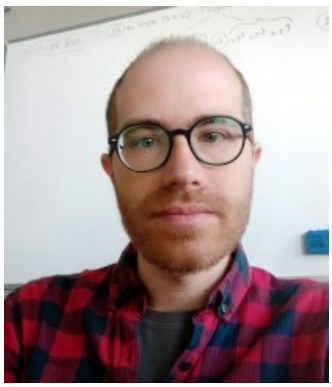
PERSONAL INFORMATION

Pierpaolo Basile



Università degli Studi di Bari Aldo Moro
CAMPUS - Dipartimento di Informatica
Via E. Orabona 4, 70125, Bari
+39 0805442301

pierpaolo.basile@uniba.it

https://swap.di.uniba.it/members/basile.pierpaolo/

H-index: 32 total citations: 4342 (source: Google Scholar)

H-index: 20 total citations: 1914 (source: Scopus)

Enterprise	University	EPR
☒ Management Level	☐ Full professor	☐ Research Director and 1st level Technologist / First Researcher and 2nd level Technologist
☐ Mid-Management Level	☒Associate Professor	☐ Level III Researcher and Technologist
☐ Employee / worker level	☐ Researcher and Technologist of IV, V, VI and VII level / Technical collaborator	☐ Researcher and Technologist of IV, V, VI and VII level / Technical collaborator

WORK EXPERIENCE

Apr. 2024 - Present

Co-founder and CEO

AI2B srl – Accredited spin-off of the University of Bari Aldo Moro, Italy

Dec. 2021 - Present

Associate Professor (IINF-05/A - Information Processing Systems)

University of Bari Aldo Moro, Department of Computer Science, Italy

- Research topics: natural language processing, information retrieval, recommender systems

Research

Dec. 2018 - Dec. 2021

Assistant Professor – RTD-B (ING-INF/05 - Information Processing Systems)

University of Bari Aldo Moro, Department of Computer Science, Italy

- Research topics: natural language processing, information retrieval, recommender systems

Research

Jun. 2016 – Dec. 2018

Assistant Professor – RTD-A (INF/01 - Computer Science)

University of Bari Aldo Moro, Department of Computer Science, Italy

- Research topics: natural language processing, multilingual entity linking

Research

Jul. 2013 – Jul. 2016

Research Fellow (Post-Doc Researcher)

University of Bari Aldo Moro, Department of Computer Science, Italy

Research topics: natural language processing, compositional operators in distributional semantic models

Research

Jun. 2009 – Jun. 2013

Research Fellow (Post-Doc Researcher)

University of Bari Aldo Moro, Department of Computer Science, Italy

Research topics: natural language processing, semantic indexing, information retrieval

Research

EDUCATION AND TRAINING

2009

Ph.D in Computer Science

University of Bari Aldo Moro, Italy

- Word Sense Disambiguation and Intelligent Information Access

2005

M.Sc. in Computer Science

University of Bari Aldo Moro, Italy

- JIGSAW: an algorithm for Word Sense Disambiguation, 108/110

PROJECTS (last five years)

With leadership roles

Sep. 2017 – Sep. 2018 Research contract between the Digital Transformation Team and the Department of Computer Science, University of Bari Aldo Moro. (project leader)

Jun. 2016 – Dec. 2018 Future in Research – project: “Multilingual Entity Linking” (cod. IAVIN27) funded by the Apulia Region (project leader)

Project reviewer

Hasler Foundation, PRIN 2017

COMMUNITY SERVICE (last 5y)

Program co-chair: IEEE International Conference on Humanized Computing and Communication with Artificial Intelligence (HCCAI 2020)

Workshops co-chair: 3rd, 2nd, 1st Knowledge-aware and Conversational Recommender Systems Workshop (KaRS 2021, KaRS 2019, KaRS 2018), 4th, 2nd, 1st Workshop on Natural Language for Artificial Intelligence (NL4AI 2020, NL4AI 2018, NL4AI 2017), 1st Workshop on Temporal Dynamics in Digital Libraries (TDDL 2017), co-located at TPDL 2017

Publication co-chair: ACL 2025 - The 63rd Annual Meeting of the Association for Computational Linguistics

Sponsor chair: 6th Italian Conference on Computational Linguistics (CLiC-it 2019)

TPC member: ACL Rolling Review; ACL 2018-2020; EMNLP 2017, 2018, 2020, 2021; EACL 2021; RecSys 2021-2019; COLING 2020, 2022; SAC 2019-2021; ECAI 2020; ESWC 2017-2021, ISWC 2017-2021; NLDB 2017-2020; CIKM 2020-2021; IJCAI 2021, 2019; AAAI 2020

Regular reviewer: Italian Journal of Computational Linguistics (IJCoL), Language Resources and Evaluation, Journal of Information Processing and Management, Semantic Web Journal, Data Mining and Knowledge Discovery, Journal of Transactions on Games, Progress in Artificial Intelligence, AI Communications

Tutorial: Lexical Semantic Change: Models, Data and Evaluation - LREC 2022

Linked Open Data for Semantics-aware Recommender Systems - ESWC 2017

Association board: Association board of AI*IA – Italian Association for Artificial Intelligence (2023-2025)

Association board of AILC (secretary) – Italian Association of Computational Linguistics (2025 - current)

EDITORIAL ACTIVITY

Co-editor: Special Issue on Gaming Applications for Cultural Heritage, IEEE Transactions on Games

Editorial Board: Language Resources and Evaluation Journal

Co-editor: Special Issue "Knowledge Graphs for Search and Recommendation", Information (ISSN 2078-2489).

PHD SUPERVISION

2024 - Present: PhD supervisor at the University of Bari: Eleonora, Angelica Ghizzota. Development and application of Generative Artificial Intelligence models based on the Symbiotic AI paradigm to support the Public Administration. PhD program in Computer Science and Mathematics, XL Cycle.

2023 - Present: PhD co-supervisor at the University of Pisa: Elio Musacchio. *Understanding language-based human signals: taking into account multimodality and adaptivity through Large Language Understanding Models*. National Ph.D. in Artificial Intelligence for Society.

2021 - 2023: PhD co-tutor at the University of Bari: Pierluigi Cassotti. *Computational Approaches for Lexical Semantic Change Detection*. PhD program in Computer Science and Mathematics, XXXVI Cycle.

According to law 679/2016 of the Regulation of the European Parliament of 27th April 2016, I hereby express my consent to process and use my data provided in this CV

2017 - 2020 PhD co-tutor at the University of Bari: Lucia Siciliani. *Question Answering over Knowledge Graph*. PhD program in Computer Science and Mathematics, XXXIII Cycle.

TEACHING

2024 – 2025, 2025-2026 "Informatics Laboratory", Bachelor Degree in Computer Science, University of Bari, Italy.

2021 - 2022 "Deep Learning for Natural Language Processing", PhD Programme in Computer Science and Mathematics, University of Bari, Italy.

2021 – 2022, 2022 – 2023, 2023 – 2024, 2024-2025, 2025-2026 "Natural Language Processing", Master Degree in Computer Science, University of Bari, Italy

2021 – 2022, 2022-2023, 2023-2024, 2024-2025, 2025-2026 "Videogame development", Bachelor Degree in Computer Science, University of Bari, Italy.

2019-2020, 2020-2021, 2021-2022, 2022-2023, 2023-2024, 2024-2025, 2025-2026 "Advanced Programming Methods", Bachelor Degree in Computer Science, University of Bari, Italy.

2019-2020 "Advanced Natural Language Processing", PhD Programme in Computer Science and Mathematics, University of Bari, Italy.

2018-2019 "Algorithms and Data Structures", Bachelor Degree in Computer Science, University of Bari, Italy.

2022-2023, 2023-2024 "Information Retrieval", Master Degree in Digital Libraries, University of Bari, Italy.

2017-2018, 2016-2017 "Information Retrieval", Bachelor Degree in Computer Science, University of Bari, Italy.

2016-2017 "Geometry and Meaning", PhD Programme in Computer Science and Mathematics, University of Bari, Italy.

INSTITUTIONAL RESPONSIBILITIES

2023 - 2024 Coordinator of the II level Master in Data Science

2023 - Present Member of the technical and scientific committee of SiMA (Sistema Museale di Ateneo), University of Bari, Italy.

2021 - Present Academic Board of the PhD Programme in Computer Science and Mathematics, University of Bari, Italy.

2017 - 2020 Coordinator of the Admissions Tests of the Bachelor Program in Computer Science, University of Bari, Italy.

INVITED TALKS

2024 "LLaMAntino: Adapting LLaMA models to new languages and domains" - 1st Italian Workshop on FAIR Management of Linguistic Legal Data (LEGALITA 2024)

2022 "AI e Recommender System" – Meeting associazione AIxIA e TELT

2017 "Question Answering for a Language Game and Customer Support" - MOD 2017 - The Third International Conference on Machine Learning, Optimization and Big Data.

2016 Panelist: "Language Technologies at the crossroad of industrial, social and cultural innovation dynamics", Third Italian Conference on Computational Linguistics (CLIC-it 2016).

FELLOWSHIPS AND AWARDS

2018 Alan Turing Institute - Visiting Researcher Application - The project titled "Diachronic Analysis of Language" has been accepted as visiting research project at the Alan Turing Institute.

According to law 679/2016 of the Regulation of the European Parliament of 27th April 2016, I hereby express my consent to process and use my data provided in this CV

2015 Future in Research - Author of the project “Multilingual Entity Linking” (cod. IAVIN27) funded by the Apulia Region (Future In Research program) for the funding of 170 fixed-term research positions in the Apulia universities.

2014 Linked Open Data-Enabled Recommender System Challenge - ESWC 2014 first system in the Top-N recommendation from binary user feedback task.

2014 CLiC-it 2014 Best Young Paper Award “Analysing Word Mining over Time by Exploiting Temporal Random Indexing” assigned by the First Italian Conference on Computational Linguistics (CLiC-it 2014).

2009 Grant awarded for attending the IJCAI 2009 - 21st International Joint Conference on Artificial Intelligence, Pasadena, California, USA.

2009 Distinguished Ph.D. theses 2009-2010 assigned by the Special Interest Group on Natural Language Processing of the Italian Association for Artificial Intelligence.

ADDITIONAL INFORMATION

2024 ASN I Fascia – settore 09/H1

2018 ASN II Fascia – settore 01/B1

2017 ASN II Fascia – settore 09/H1

PUBLICATIONS

According to law 679/2016 of the Regulation of the European Parliament of 27th April 2016, I hereby express my consent to process and use my data provided in this CV

Publications
best and most relevant
in the last 10 years

1. Musacchio, E., Siciliani, L., Basile, P., & Semeraro, G. (2026). Do you understand Italian? Evaluating LVLMs on Italian visual question-answering. Language Resources and Evaluation, 60(3), 49.
2. Musacchio, E., Siciliani, L., Basile, P., & Semeraro, G. (2026). MAMMa: adaptation of multimodal LLM embeddings to multilinguality. Neural Computing and Applications, 38(15), 638.
3. Polignano, M., Basile, P., & Semeraro, G. (2026). Advanced natural-based interaction for the italian language: Llamantino-3-anita. Scientific Reports.
4. Musacchio, E., Siciliani, L., Basile, P., & Semeraro, G. (2025). Assessing and Improving the Multilingual Visual Word Sense Disambiguation Ability of Vision-Language Models. Frontiers in Artificial Intelligence and Applications, 10.3233/FAIA251306.
5. Musacchio, E., Siciliani, L., Basile, P., & Semeraro, G. (2025). Extending Large Language Models to multimodality for non-English languages. Computer Vision and Image Understanding, 104618.
6. Di Bonaventura, C., Siciliani, L., Basile, P., Penuela, A. M., & McGillivray, B. (2025). From Detection to Explanation: Effective Learning Strategies for LLMs in Online Abusive Language Research. In Proceedings of the 2025 International Conference on Computational Linguistics (COLING 2025).
7. Siciliani, L., Ghizzota, E., Basile, P., & Lops, P. (2024). OIE4PA: open information extraction for the public administration. Journal of Intelligent Information Systems, 62(1), 273-294.
8. Siciliani, L., Taccardi, V., Basile, P., Di Ciano, M., & Lops, P. (2023). AI-based decision support system for public procurement. Information systems, 119, 102284.
9. Cassotti, P., Siciliani, L., de Gemmis, M., Semeraro, G., & Basile, P. (2023). XL-LEXEME: WiC pretrained model for cross-lingual LEXical sEMantic changE. In PROCEEDINGS OF THE CONFERENCE-ASSOCIATION FOR COMPUTATIONAL LINGUISTICS. MEETING (Vol. 2, pp. 1577-1585). Association for Computational Linguistics (ACL).
10. Siciliani, Lucia, Basile, Pierpaolo, Lops, Pasquale, Semeraro, Giovanni. MQALD: Evaluating the impact of modifiers in question answering over knowledge graphs. SEMANTIC WEB, 13 (2), p. 215 – 231, 2022. DOI: 10.3233/SW-210440
11. Tsakalidis, Adam, Basile, Pierpaolo, Bazzi, Marya, Cucuringu, Mihai, McGillivray, Barbara. DUKweb, diachronic word representations from the UK Web Archive corpus. SCIENTIFIC DATA, 8 (1), art. no. 269, 2021. DOI: 10.1038/s41597-021-01047-x
12. Florio, Komal, Basile, Valerio, Polignano, Marco, Basile, Pierpaolo, Patti, Viviana. Time of your hate: The challenge of time in hate speech detection on social media. APPLIED SCIENCE, 10 (12), art. no. 4180, 2020. DOI: 10.3390/APP10124180
13. Narducci, Fedelucio, Basile, Pierpaolo, de Gemmis, Marco, Lops, Pasquale, Semeraro, Giovanni. An investigation on the user interaction modes of conversational recommender systems for the music domain. USER MODELING AND USER-ADAPTED INTERACTION, vol. 30, p. 251-284, 2020. DOI: 10.1007/s11257-019-09250-7.
14. Polignano, M., Basile, P., De Gemmis, M., Semeraro, G., & Basile, V. (2019). Alberto: Italian BERT language understanding model for NLP challenging tasks based on tweets. In CEUR workshop proceedings (Vol. 2481, pp. 1-6). CEUR.
15. Narducci, Fedelucio, Basile, Pierpaolo, Musto, Cataldo, Lops, Pasquale, Caputo, Annalina, de Gemmis, Marco, Iaquinta, Leo, Semeraro, Giovanni (2016). Concept-based item representations for a cross-lingual content-based recommendation process. INFORMATION SCIENCES, vol. 374, p. 15-31, ISSN: 0020-0255, DOI: 10.1016/j.ins.2016.09.022

Consapevole delle sanzioni penali, nel caso di dichiarazioni non veritiere, di formazione o uso di atti falsi, richiamate dall'art. 76 del D.P.R. 445/2000, dichiaro, altresì, di essere informato che i dati personali raccolti saranno trattati, anche con strumenti informatici, esclusivamente nell'ambito del procedimento per il quale la presente dichiarazione viene resa e che al riguardo competono al sottoscritto tutti i diritti previsti all'art. 7 della medesima legge".- "Si autorizza il trattamento dei dati personali ai sensi della vigente normativa sulla privacy – D.lgs. n. 196/2003 e ss.mm.ii. ed ex articolo 13 del Regolamento Europeo 2016/679 (GDPR)

Bari, September 2026.