

Pier Felice Balestrucci

Curriculum Vitae

✉ pierbale97@gmail.com
🌐 [pierbale.github.io](https://github.com/pierbale)

NLP researcher specializing in LLM-based systems, conversational AI, and human-centered evaluation. Builds end-to-end workflows for multilingual data annotation, prompting, fine-tuning, retrieval, and model assessment using automatic metrics and human evaluation. Research focuses on dialogue personalization, pragmatic phenomena such as irony and hate speech, and accessible interaction.

Personal Information

First Name Pier Felice
Last Name Balestrucci
Nationality Italian
Year of Birth 1997

Work Experience

- Since Nov 2025 **Research Fellow**
Department of Computer Science - University of Turin, Turin, Italy
(ongoing) Project: *Understanding, Generation and Management of Pragmatically Constrained Utterances in Dialogue Systems through Perspective-Aware Language Modeling*. Develop end-to-end NLP pipelines in Python, PyTorch, and Hugging Face for multilingual annotation and LLM evaluation in irony detection and generation. Compare prompting, fine-tuning, and perspectivist approaches using automatic metrics and human evaluation, producing reusable datasets and evaluation assets.
- Oct 2025 - Sep 2026 **Contract Lecturer**
Department of Computer Science - University of Turin, Turin, Italy
Design and delivery of the Master's-level course *Intelligent Systems*, covering AI and machine learning for Strategic and Military Sciences.
- Nov 2022 - Oct 2025 **Doctoral Researcher in Computer Science**
Department of Computer Science - University of Turin, Turin, Italy
Built modular dialogue systems integrating transformer-based NLU, retrieval, and LLM components for accessible, non-visual exploration of STEM content. Converted evaluation findings into accessibility requirements and successive system iterations supporting screen-reader-compatible interaction. Designed reproducible experiment and evaluation pipelines, with results published at international NLP venues and in a Q1 scientific journal. National doctoral scholarship funded by the Italian Ministry of University and Research.
- Jun - Oct 2022 **Data Scientist**
MA Computer S.r.l., Turin, Italy
Prepared and validated multinational procurement data and delivered a statistical module supporting supply-chain optimization.

- Jan - Jun 2022 **Data Analyst (internship)**
Accenture, Milan, Italy
Developed supervised machine-learning models for suspicious-transaction detection; prepared datasets and validated outputs with a banking data analytics team.
- Dec 2020 - Aug 2021 **Student Collaborator**
Department of Computer Science - University of Turin, Turin, Italy
Prepared accessible educational materials for students with disabilities through a part-time university scholarship.

Education

- Nov 2022 - Sep 2026 (expected) **PhD in Computer Science**, University of Turin, Turin, Italy
Thesis title: Personalisation in Dialogue Systems: From Accessible Interaction to Pragmatic Awareness
Scope: investigates dialogue-system personalization through two complementary areas, accessible interaction and pragmatic awareness, using multilingual data, LLM-based methods, and human-centered evaluation.
- Nov 2019 - Jun 2022 **MSc in Artificial Intelligence - 110/110 with honors**, University of Turin, Turin, Italy
Thesis title: Voice-based interaction for navigating mathematical expressions
Scope: designed and developed a voice interface for non-visual navigation of mathematical expressions using dialogue and structured representations.
- Sep 2016 - Oct 2019 **BSc in Computer Science - 110/110 with honors**, University of Bari, Bari, Italy
Thesis title: "Pepper for Tourism" - Design and development of an information system for tourism in Apulia using the social robot Pepper
Scope: developed a conversational and multimodal tourism information system for the Pepper social robot.

Research and Training Experiences

- Sep 2024-Feb 2025 **Visiting Researcher at the Institute of Formal and Applied Linguistics, Charles University (Prague, Czechia)**
Created P-MultiWoZ and investigated whether LLMs reproduce generational language styles while preserving dialogue goals; the work was published in Findings of EMNLP 2025.
<https://ufal.mff.cuni.cz/home-page>
- Jun 2024 **Lectures on Computational Linguistics 2024, organized by AIXIA and AILC**
Topics: Cognitive Evaluation of LLMs, Lexical Semantic Change, Digital Humanities, Prompting
<https://www.ai-lc.it/en/lectures-2/lectures-2024/>
- Jul 2023 **European Summer School in Logic, Language and Information (ESSLLI)**
Topics: Formal Language Theory, LLMs, Computational Creativity, Semantic Representations
<https://2023.esslli.eu>
- Mar 2023 **Bertinoro International Spring School on AI (BISS)**
Courses: Hybrid Quantum Computing, AI for Health, Neural MT, Programming Language Design
<https://cs.unibo.it/projects/BISS/2023/program/>

Selected Publications

- [1] Pier Felice Balestrucci, Ondřej Dušek, Luca Anselma, and Alessandro Mazzei. Can Large Language Models Personalize Dialogues to Generational Styles? In *Findings of the Association for Computational Linguistics: EMNLP 2025*, pages 64–77, Suzhou, China, 2025. Association for Computational Linguistics. doi: 10.18653/v1/2025.findings-emnlp.5. URL <https://aclanthology.org/2025.findings-emnlp.5/>.
- [2] Michael Oliverio, Pier Felice Balestrucci, Alessandro Mazzei, and Valerio Basile. WebNLG-IT: Construction of an aligned rdf-italian corpus through machine translation techniques. In *Findings of the Association for Computational Linguistics: ACL 2025*, pages 12073–12083, Vienna, Austria, 2025. Association for Computational Linguistics. doi: 10.18653/v1/2025.findings-acl.625. URL <https://aclanthology.org/2025.findings-acl.625/>.
- [3] Michael Oliverio, Pier Felice Balestrucci, Luca Anselma, and Alessandro Mazzei. AIML+: Enhancing AIML for the educational domain through frames and large language models. In *Artificial Intelligence in Education*, Lecture Notes in Computer Science, pages 176–189. Springer Nature Switzerland, 2025. doi: 10.1007/978-3-031-98459-4_13. URL https://doi.org/10.1007/978-3-031-98459-4_13.
- [4] Pier Felice Balestrucci, Silvia Casola, Soda Marem Lo, Valerio Basile, and Alessandro Mazzei. I'm Sure You're a Real Scholar Yourself: Exploring Ironic Content Generation by Large Language Models. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 14480–14494, Miami, Florida, USA, 2024. Association for Computational Linguistics. doi: 10.18653/v1/2024.findings-emnlp.847. URL <https://aclanthology.org/2024.findings-emnlp.847/>.
- [5] Pier Felice Balestrucci, Elisa Di Nuovo, Manuela Sanguinetti, Luca Anselma, Cristian Bernareggi, and Alessandro Mazzei. An Educational Dialogue System for Visually Impaired People. *IEEE Access*, 12: 150502–150519, 2024. doi: 10.1109/ACCESS.2024.3479883. URL <https://doi.org/10.1109/ACCESS.2024.3479883>.

The complete and up-to-date publication list is available on [Google Scholar](#).

Conference Presentations and Talks

- Feb 2025 Speaker at Mathematics Education and Visual Impairments Workshop, Laboratorio Polin
<https://www.sema.unito.it/>
- Sep 2024 Poster presentation at iNLG 2024, GenChal Session, presenting *DipInfo-UniTo at the GEM'24 Data-to-Text Task: Augmenting LLMs with the Split-Generate-Aggregate Pipeline*
<https://2024.inlgmeeting.org>
- May 2024 Poster presentation at LREC-COLING 2024, presenting *Educational Dialogue Systems for Visually Impaired Students: Introducing a Task-Oriented User-Agent Corpus*
<https://lrec-coling-2024.org>
- Nov 2023 Poster presentation at CLiC-it 2023, presenting *Building a Spoken Dialogue System for Supporting Blind People in Accessing Mathematical Expressions*
<https://cllc2023.ilc.cnr.it>
- Dec 2022 Speaker at Making Science Inclusive, Laboratorio Polin
https://www.integr-abile.unito.it/workshop2022_2/

Research Projects

2024 - 2026 **Perspective-Aware NLP - University of Turin/Amazon Alexa research collaboration**

Project investigating how individual and group perspectives affect the interpretation and generation of pragmatic phenomena. Contributed to multilingual corpora, annotation guidelines, and evaluation pipelines for irony and rhetorical figures, comparing zero-shot, few-shot, and fine-tuned LLMs under aggregate and perspective-specific settings.

2021 - 2024 **NoVAGraphS - Non-Visual Access to Graphical Structures**

<http://www.integr-abile.unito.it/progetto-novagraphs/>

Interdisciplinary project on assistive dialogue systems for blind and visually impaired students. Contributed to modular architectures integrating NLU, dialogue management, RAG, and LLM components. The project produced an extensible conversational platform for non-visual exploration of mathematical expressions, functional diagrams, and finite-state automata, supporting screen-reader workflows and controlled accessibility evaluation.

Reviewing and Program Committees

- Reviewer
- ACM Journal on Computing and Cultural Heritage (ACM JOCCH)
<https://dl.acm.org/journal/jocch>
 - Artificial Intelligence (AIJ)
<https://www.sciencedirect.com/journal/artificial-intelligence>
 - The 28th European Conference on Artificial Intelligence (ECAI 2025)
25 - 30 Oct 2025, Bologna, Italy
<https://ecai2025.org/>
 - The 11th Italian Conference on Computational Linguistics (CLiC-it 2025)
24 - 26 Sep 2025, Cagliari, Italy
<https://clic2025.unica.it/>
 - ACL Rolling Review (ARR) 2024 - for EMNLP 2024, COLING 2025, and NAACL 2025
<https://aclrollingreview.org/>
 - The 25th Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL 2024)
18 - 20 Sep 2024, Kyoto, Japan
<https://2024.sigdial.org>
 - The 24th Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL 2023)
11 - 15 Sep 2023, Prague, Czechia
<https://2023.sigdial.org>
 - The 28th International Conference on Natural Language & Information Systems (NLDB 2023)
21 - 23 Jun 2023, Derby, United Kingdom
<http://www.wikicfp.com/cfp/servlet/event.showcfp?eventid=170418>

- The 16th International Conference on Health Informatics (HealthInf 2023)
16 - 18 Feb 2023, Lisbon, Portugal
<https://healthinf.scitevents.org/?y=2023>
- Program Committees
 - The 30th Annual International Conference on Natural Language & Information Systems (NLDB 2025)
4 - 6 Jul 2025, Kanazawa, Japan
<https://www.jaist.ac.jp/event/nldb2025/about.html>
 - Workshop on Natural Language for Artificial Intelligence (NL4AI, 8th edition)
25 - 28 Nov 2024, Bolzano, Italy
<http://sag.art.uniroma2.it/NL4AI/>
 - The 29th International Conference on Natural Language & Information Systems (NLDB 2024)
25 - 27 Jun 2024, Turin, Italy
<https://nldb2024.di.unito.it>

Public Engagement

- 2023 **FundsTOgether – Crowdfunding for Research and Innovation**
Participation in a university-wide alternative finance program for research and innovation, implemented through crowdfunding. The initiative aims to fund projects with strong social and environmental impact by engaging a wide audience.
- 2023 **MatematlcA senza barriere – Crowdfunding Campaign**
Participation in the “MatematlcA senza barriere” crowdfunding campaign, part of the FundsTOgether program. Activities included training, project selection, and the launch and management of the campaign.
- 2022 **SpeechMate – POV Call, Laboratorio Polin**
Participation in the POV Call with the SpeechMate project, involving collaborations with associations and companies interested in accessibility technologies.
- 2022 **Making Science Inclusive – Workshop, Laboratorio Polin**
Contributed to the organization of a workshop addressing problems, solutions, and proposals for improving the accessibility of scientific texts and publications.

Teaching

- 2025 - 2026 Contract Lecturer
Course: *Intelligent Systems*, MSc in Strategic and Military Sciences, University of Turin
- 2025 - 2026 Tutoring - Tecnologie Web Approcci Avanzati, 40 hours
MSc in Communication, ICT and Media, University of Turin
- 2023 - 2024 Tutoring - Programmazione II A, B, C (course code inf0330), 40 hours
BSc in Computer Science, University of Turin
- 2023 - 2024 Tutoring - Tecnologie Web Approcci Avanzati, 40 hours
MSc in Communication, ICT and Media, University of Turin

Thesis Supervision

- 2023-2024 Co-supervisor of Davide Candela's Bachelor's thesis:
"Translation of Input into Controlled Natural Language for a Chatbot Specialized in Finite State Automata Using a Large Language Model"
- 2024-2025 Co-supervisor of Stefano Porta's Master's thesis:
"Design, ingegnerizzazione e realizzazione di un sistema di dialogo basato su LLM nel dominio delle tecnologie assistive"

Language Skills

- Italian Native speaker
- English Professional working proficiency in reading, writing, and conversation