Simone **Balloccu**



Profile

I'm an AI researchers with 9 years of experience, and 24 peer-reviewed publications. My research focuses on how AI can support domain experts, and how effective it can be in real-world scenarios. I have extensive experience in the human evaluation of AI, and tend to prioritize it in my work.

Contact details

@ simone.balloccu@tu-darmstadt.de
+39 345 595 1827
uccollab.github.io
Room D205, S2|02
Robert-Piloty-Building,
Hochschulstraße 10, 64289
Darmstadt (DE)

Personal information

Citizenship: **Italian**

Languages: **Italian** (Native), **English** (IELTS-C1)

Skills

- Machine learning applied to NLP: training, analysis, bias, and fairness assessment.
- Extrinsic and Intrinsic NLP system evaluation, including Human-in-the-loop and Expert-in-the-loop.
- Dataset collection and annotation in NLP.

Research experience

Junior Lab leader **2024.10–present**
TU Darmstadt (DE)
› **NLP for expert domains (ExpNLP) group**
◇ **Topics:** Expert-AI dynamics

Postdoc **2023.07–2024.09**
Charles University (CZ)
› **ERC-2021-STG, NG-NLG project**
◇ **Topics:** Neuro-symbolic NLG; LLMs controllability.

Visiting researcher **2022.05–2022.10**
Philips Research (NL)
› **Horizon 2020, Philhumans project**
◇ **Topics:** LLMs for counselling; AI-driven health activity tracking.

Visiting researcher **2020.07–2021.05**
Philips Research (NL)
› **Horizon 2020, Philhumans project**
◇ **Topics:** LLMs for counselling.

Assistant researcher **2018.12–2019.09**
University of Cagliari (IT).
› **Sardegna Ricerche, DoUtDes project**
◇ **Topics:** AI-driven marketing campaigns for companies

Assistant researcher **2017.01–2019.09**
University of Cagliari (IT).
› **Sardegna Ricerche, OKGraph project**
◇ **Topics:** Unsupervised NLP algorithm for free-text knowledge extraction.

Education

PhD in Artificial Intelligence **2019.10–2023.05**
University of Aberdeen (UK).
› **Marie Skłodowska-Curie ESR - PHilhumans project (ESR-7)**
› Thesis title: *Emotion-aware and tailored diet coaching with Natural Language Generation.*

M.Sc in Computer Science **2017.10–2019.09**
University of Cagliari (IT).
› Thesis title: *Unsupervised Hypernym Discovery from plain text.*

B.Sc in Computer Science **2013.10–2017.09**
University of Cagliari (IT).
› Thesis title: *Unsupervised learning of IS-A relations from plain text.*

Teaching

Course co-ordinator **2025 - 2026**
TU Darmstadt (DE).
› **Deep Learning for Natural Language Processing**
› **Responsibilities:** *Lectures, course coordination, marking.*

Teaching assistant **2023 - 2024**
Charles University (CZ).
› **NPFL099: Statistical Dialogue Systems**
› **Responsibilities:** *Lectures, marking.*

Demonstrator **2021 - 2022**
University of Aberdeen (UK).
› **CS551H: Natural Language Generation**
› **Responsibilities:** *Labs, demonstrations.*

Invited lecturer **2020 - 2021**
University of Cagliari (IT).
› **SC/0132/EN: Natural Language Processing**
› **Responsibilities:** *Invited lectures, labs.*

Mentoring/Supervision

Across my career, I supervised and co-supervised more than 10 students (both B.Sc. and M.Sc.) at various institutions, including Charles University (CZ), University of Aberdeen (UK), and the University of Utrecht (NL). At TU Darmstadt (DE), besides B.Sc. and M.Sc., I currently supervise three PhD students within my research lab.

Conference organisation and Reviewing

I served as reviewer for various top-tier conferences (including EACL, ACL, and EMNLP) and as Area Chair for the EMNLP and EACL Demo track; I served in the organizing committee for various workshops (including HumEval, and Practical D2T) and conferences (including INLG and SIGDial). In 2026, I was recognized as "Outstanding Reviewer" by EACL.

Research and publications

Updated as per 16/10/2026. Up-to-date information on **Google Scholar**

Peer-reviewed publications: 25
Google h-index: 11
Google i-10 index: 14
Citations: 958

› Selected peer-reviewed publications

- **"Leak, Cheat, Repeat: Data Contamination and Evaluation Malpractices in Closed-Source LLMs"**
  Simone Balloccu, Patrícia Schmidtová, Mateusz Lango, Ondřej Dušek.
  In Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers), Mar 2024
  CORE Rank: A
  🏆 **Best Non Publicized Paper Award**

- **"Automatic Metrics in Natural Language Generation: A survey of Current Evaluation Practices"**
  Patricia Schmidtova, Saad Mahamood, **Simone Balloccu**, Ondrej Dusek, Albert Gatt, Dimitra Gkatzia, David M. Howcroft, Ondrej Platek, Adarsa Sivaprasad.
  In Proceedings of the 17th International Natural Language Generation Conference, Sept 2024
  CORE Rank: B
  🏆 **Best Evaluation Paper Award**

- **"Ask the experts: sourcing a high-quality nutrition counseling dataset through Human-AI collaboration"**
  Simone Balloccu, Ehud Reiter, Karen Jia-Hui Li, Rafael Sargsyan, Vivek Kumar, Diego Reforgiato, Daniele Riboni, Ondrej Dusek.
  In Findings of the Association for Computational Linguistics: EMNLP 2024

- **Are Experts Needed? On Human Evaluation of Counselling Reflection Generation.**
  Zixiu Wu, **Simone Balloccu**, Ehud Reiter, Rim Helaoui, Diego Reforgiato Recupero, Daniele Riboni.
  In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Jul 2023.
  CORE Rank: A*

- **Anno-MI: A Dataset of Expert-Annotated Counselling Dialogues.**
  Zixiu Wu, **Simone Balloccu**, Vivek Kumar, Rim Helaoui, Ehud Reiter, Diego Reforgiato Recupero, Daniele Riboni.
  In ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), May 2022.
  CORE Rank: B