Sara Papi

✉ sara95papi@gmail.com   𝕏 @sarapapi   in sara-papi
🌐 https://sarapapi.github.io/   🎓 Google Scholar (h-index 12)

Experience

2024 – now 🔖 **AI Researcher** in Speech Processing and Multimodal LLMs (*Fondazione Bruno Kessler, Italy*) – WP Technical Lead in the MEETWEEN Horizon European Project and Task Lead in the DVPS Horizon European Project

2023 – 2023 🔖 **AI Research Intern** in Streaming Speech Recognition and Translation (*Microsoft, USA*) – Supervised by Jinyu Li

2019 – 2019 🔖 **AI Research Intern** in Automatic Speech Scoring (*Fondazione Bruno Kessler, Italy*)

Education

2020 – 2024 🔖 **Ph.D. cum Laude** in Information Engineering and Computer Science (Fondazione Bruno Kessler & University of Trento, Italy)

2017 – 2020 🔖 **M.Sc. cum Laude** in Computer and Automation Engineering (University of Siena, Italy)

2014 – 2017 🔖 **B.Sc.** in Information Engineering (University of Siena, Italy)

Awards and Achievements

2025 🔖 **Best PhD Graduate 2024 in Information and Communication Technology** by University of Trento

🔖 **PhD Excellence Award** by Fondazione Bruno Kessler

🔖 **Highly Commended 2024 Anthony C Clarke Award** for the **EAMT Best Thesis** by EAMT

2024 🔖 **Special Mention for the Best Italian PhD Thesis** by AIxIA 2024.

🔖 **Outstanding Paper and SAC Award** by ACL 2024

🔖 **Social Impact Paper Award** by EMNLP 2024

Invited Talks/Keynotes

2026 🔖 **Cohere Labs Open Science Community**. Talk Title: *"Towards Crosslingual Evaluation of General-Purpose Instruction-Following Models Across Text, Speech, and Vision"*.

2025 🔖 **MT Marathon**. Talk Title: *"Speech Translation in the LLM Era: Between Old and New Challenges"*.

🔖 **SlatorCon Remote**. Talk Title: *"How "Real" is your Real-Time Simultaneous Speech-to-Text System?"*.

**2024** 🔖 **IWSLT 2024**. Panel Title: *"Emerging trends in Speech Translation"*.

Research Publications

Journal Articles

1 **S. Papi**, P. Polák, D. Macháček, and O. Bojar, "How "real" is your real-time simultaneous speech-to-text translation system?" *Transactions of the Association for Computational Linguistics*, vol. 13, pp. 281–313, Apr. 2025, ISSN: 2307-387X. 🔗 DOI: 10.1162/tacl_a_00740 eprint: https://direct.mit.edu/tacl/article-pdf/doi/10.1162/tacl_a_00740/2513610/tacl_a_00740.pdf.

2   **S. Papi**, M. Gaido, A. Karakanta, M. Cettolo, M. Negri, and M. Turchi, "Direct speech translation for automatic subtitling," *Transactions of the Association for Computational Linguistics*, vol. 11, pp. 1355–1376, Nov. 2023, ISSN: 2307-387X. 🔗 DOI: `10.1162/tacl_a_00607`

Conference Proceedings

1   I. Abdulmumin et al., "Findings of the IWSLT 2025 evaluation campaign," in *Proceedings of the 22nd International Conference on Spoken Language Translation (IWSLT 2025)*, Jul. 2025, ISBN: 979-8-89176-272-5. 🔗 DOI: `10.18653/v1/2025.iwslt-1.44`

2   M. Gaido*, **S. Papi***, L. Bentivogli, A. Brutti, M. Cettolo, and other, "The warmup dilemma: How learning rate strategies impact speech-to-text model convergence," in *Proceedings of the 22nd International Conference on Spoken Language Translation (IWSLT 2025)*, Vienna, Austria (in-person and online), Jul. 2025, ISBN: 979-8-89176-272-5. 🔗 DOI: `10.18653/v1/2025.iwslt-1.4`

3   T. K. Lam*, M. Gaido*, **S. Papi**, L. Bentivogli, and B. Haddow, "Prepending or cross-attention for speech-to-text? an empirical comparison," in *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, Apr. 2025, ISBN: 979-8-89176-189-6. 🔗 URL: `https://aclanthology.org/2025.naacl-long.153/`

4   N. Rao Koluguri et al., "Granary: Speech Recognition and Translation Dataset in 25 European Languages," in *Interspeech 2025*, 2025. 🔗 DOI: `10.21437/Interspeech.2025-540`

5   F. Verdini et al., "How to Connect Speech Foundation Models and Large Language Models? What Matters and What Does Not," in *Interspeech 2025*, 2025. 🔗 DOI: `10.21437/Interspeech.2025-2245`

6   M. Züfle, **S. Papi**, B. Savoldi, M. Gaido, L. Bentivogli, and J. Niehues, "NUTSHELL: A dataset for abstract generation from scientific talks," in *Proceedings of the 22nd International Conference on Spoken Language Translation (IWSLT 2025)*, Jul. 2025, ISBN: 979-8-89176-272-5. 🔗 DOI: `10.18653/v1/2025.iwslt-1.2`

7   I. S. Ahmad et al., "FINDINGS OF THE IWSLT 2024 EVALUATION CAMPAIGN," in *Proceedings of the 21st International Conference on Spoken Language Translation (IWSLT 2024)*, Aug. 2024. 🔗 DOI: `10.18653/v1/2024.iwslt-1.1`

8   M. Cettolo, A. Piergentili, **S. Papi**, M. Gaido, M. Negri, L. Bentivogli, et al., "Magnet-machines generating translations: A calamita challenge," in *Proceedings of the Tenth Italian Conference on Computational Linguistics (CLiC-it 2024)*, 2024.

9   M. Gaido, **S. Papi**, M. Negri, and L. Bentivogli, "Speech translation with speech foundation models and large language models: What is there and what is missing?" In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Aug. 2024. 🔗 DOI: `10.18653/v1/2024.acl-long.789`

10   M. Gaido, **S. Papi**, M. Negri, M. Cettolo, and L. Bentivogli, "SBAAM! eliminating transcript dependency in automatic subtitling," in *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Aug. 2024. 🔗 DOI: `10.18653/v1/2024.acl-long.201`

11   M. Gaido et al., "Automatic subtitling and subtitle compression: FBK at the IWSLT 2024 subtitling track," in *Proceedings of the 21st International Conference on Spoken Language Translation (IWSLT 2024)*, Aug. 2024. 🔗 DOI: `10.18653/v1/2024.iwslt-1.13`

12   M. Gaido*, **S. Papi***, L. Bentivogli, A. Brutti, M. Cettolo, and other, "MOSEL: 950,000 hours of speech data for open-source speech foundation model training on EU languages," in *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, Nov. 2024. 🔗 DOI: `10.18653/v1/2024.emnlp-main.771`

13 M. Gaido*, **S. Papi***, M. Negri, and L. Bentivogli, "How do hyenas deal with human speech? speech recognition and translation with ConfHyena," in *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, May 2024. 🔗 URL: https://aclanthology.org/2024.lrec-main.717/

14 **S. Papi**, M. Gaido, M. Negri, and L. Bentivogli, "SimulSeamless: FBK at IWSLT 2024 simultaneous speech translation," in *Proceedings of the 21st International Conference on Spoken Language Translation (IWSLT 2024)*, Aug. 2024. 🔗 DOI: 10.18653/v1/2024.iwslt-1.11

15 **S. Papi**, M. Gaido, M. Negri, and L. Bentivogli, "StreamAtt: Direct streaming speech-to-text translation with attention-based audio history selection," in *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Aug. 2024. 🔗 DOI: 10.18653/v1/2024.acl-long.202

16 **S. Papi** et al., "Leveraging timestamp information for serialized joint streaming recognition and translation," in *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2024. 🔗 DOI: 10.1109/ICASSP48485.2024.10447565

17 **S. Papi***, M. Gaido*, A. Pilzer, and M. Negri, "When good and reproducible results are a giant with feet of clay: The importance of software quality in NLP," in *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Aug. 2024. 🔗 DOI: 10.18653/v1/2024.acl-long.200

18 B. Savoldi, **S. Papi**, M. Negri, A. Guerberof-Arenas, and L. Bentivogli, "What the harm? quantifying the tangible impact of gender bias in machine translation with a human-centered study," in *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, Nov. 2024. 🔗 DOI: 10.18653/v1/2024.emnlp-main.1002

19 D. Fucci, M. Gaido, **S. Papi**, M. Cettolo, M. Negri, and L. Bentivogli, "Integrating language models into direct speech translation: An inference-time solution to control gender inflection," in *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, Dec. 2023. 🔗 DOI: 10.18653/v1/2023.emnlp-main.705

20 M. Gaido, **S. Papi**, M. Negri, and M. Turchi, "Joint speech translation and named entity recognition," in *Interspeech 2023*, 2023. 🔗 DOI: 10.21437/Interspeech.2023-1767

21 **S. Papi**, M. Gaido, and M. Negri, "Direct models for simultaneous translation and automatic subtitling: FBK@IWSLT2023," in *Proceedings of the 20th International Conference on Spoken Language Translation (IWSLT 2023)*, Jul. 2023. 🔗 DOI: 10.18653/v1/2023.iwslt-1.11

22 **S. Papi**, M. Negri, and M. Turchi, "Attention as a guide for simultaneous speech translation," in *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Jul. 2023. 🔗 DOI: 10.18653/v1/2023.acl-long.745

23 **S. Papi**, M. Turchi, and M. Negri, "Alignatt: Using attention-based audio-translation alignments as a guide for simultaneous speech translation," in *Interspeech 2023*, 2023. 🔗 DOI: 10.21437/Interspeech.2023-170

24 **S. Papi**, P. Wang, J. Chen, J. Xue, J. Li, and Y. Gaur, "Token-level serialized output training for joint streaming asr and st leveraging textual alignments," in *2023 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, 2023. 🔗 DOI: 10.1109/ASRU57964.2023.10389715

25 M. Gaido*, **S. Papi***, D. Fucci, G. Fiameni, M. Negri, and M. Turchi, "Efficient yet competitive speech translation: FBK@IWSLT2022," in *Proceedings of the 19th International Conference on Spoken Language Translation (IWSLT 2022)*, May 2022. 🔗 DOI: 10.18653/v1/2022.iwslt-1.13

26 **S. Papi**, M. Gaido, M. Negri, and M. Turchi, "Does simultaneous speech translation need simultaneous models?" In *Findings of the Association for Computational Linguistics: EMNLP 2022*, Dec. 2022. 🔗 DOI: 10.18653/v1/2022.findings-emnlp.11

27 **S. Papi**, M. Gaido, M. Negri, and M. Turchi, "Over-generation cannot be rewarded: Length-adaptive average lagging for simultaneous speech translation," in *Proceedings of the Third Workshop on Automatic Simultaneous Translation*, Jul. 2022. 🔗 DOI: `10.18653/v1/2022.autosimtrans-1.2`

28 **S. Papi**, A. Karakanta, M. Negri, and M. Turchi, "Dodging the data bottleneck: Automatic subtitling with automatically segmented ST corpora," in *Proceedings of the 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, Nov. 2022. 🔗 DOI: `10.18653/v1/2022.aacl-short.59`

29 A. Karakanta*, **S. Papi***, M. Negri, and M. Turchi, "Simultaneous speech translation for live subtitling: From delay to display," in *Proceedings of the 1st Workshop on Automatic Spoken Language Translation in Real-World Settings (ASLTRW)*, Aug. 2021. 🔗 URL: `https://aclanthology.org/2021.mtsummit-asltrw.4/`

30 **S. Papi**, M. Gaido, M. Negri, and M. Turchi, "Dealing with training and test segmentation mismatch: FBK@IWSLT2021," in *Proceedings of the 18th International Conference on Spoken Language Translation (IWSLT 2021)*, Aug. 2021. 🔗 DOI: `10.18653/v1/2021.iwslt-1.8`

31 **S. Papi**, M. Negri, and M. Turchi, "Visualization: The missing factor in simultaneous speech translation," in *Proceedings of the Seventh Italian Conference on Computational Linguistics (CLiC-it 2021)*, 2021.

32 **S. Papi***, M. Gaido*, M. Negri, and M. Turchi, "Speechformer: Reducing information loss in direct speech translation," in *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, Nov. 2021. 🔗 DOI: `10.18653/v1/2021.emnlp-main.127`

33 **S. Papi**, E. Trentin, R. Gretter, M. Matassoni, and D. Falavigna, "Mixtures of deep neural experts for automated speech scoring," in *Interspeech 2020*, 2020. 🔗 DOI: `10.21437/Interspeech.2020-1055`

Under Review

1 M. Cettolo, M. Gaido, M. Negri, **S. Papi**, and L. Bentivogli, *How to evaluate speech translation with source-aware neural mt metrics*, 2025. arXiv: `2511.03295 [cs.CL]`. 🔗 URL: `https://arxiv.org/abs/2511.03295`

2 M. Gaido, **S. Papi**, M. Cettolo, M. Negri, and L. Bentivogli, *Simulstream: Open-source toolkit for evaluation and demonstration of streaming speech-to-text translation systems*, 2025. arXiv: `2512.17648 [cs.CL]`. 🔗 URL: `https://arxiv.org/abs/2512.17648`

3 M. Nissim et al., *Challenging the abilities of large language models in italian: A community initiative*, 2025. arXiv: `2512.04759 [cs.CL]`. 🔗 URL: `https://arxiv.org/abs/2512.04759`

4 **S. Papi**, D. Fucci, M. Gaido, M. Negri, and L. Bentivogli, *Cross-attention is half explanation in speech-to-text models*, 2025. arXiv: `2509.18010 [cs.CL]`. 🔗 URL: `https://arxiv.org/abs/2509.18010`

5 **S. Papi** et al., *Hearing to translate: The effectiveness of speech modality integration into llms*, 2025. arXiv: `2512.16378 [cs.CL]`. 🔗 URL: `https://arxiv.org/abs/2512.16378`

6 **S. Papi** et al., *MCIF: Multimodal crosslingual instruction-following benchmark from scientific talks*, 2025. arXiv: `2507.19634 [cs.CL]`. 🔗 URL: `https://arxiv.org/abs/2507.19634`

7 P. Polák, **S. Papi**, L. Bentivogli, and O. Bojar, *Better late than never: Evaluation of latency metrics for simultaneous speech-to-text translation*, 2025. arXiv: `2509.17349 [cs.CL]`. 🔗 URL: `https://arxiv.org/abs/2509.17349`

Community Service

Organizer/Committee Member

2026   **IWSLT 2026 - Evaluation Campaigns**, Instruction Following (Chair) and Simultaneous Translation (Organizer).

2025   **IWSLT 2025 - Evaluation Campaigns**, Instruction Following (Chair) and Simultaneous Translation (Organizer).

2024   **EACL 2024 Student Research Workshop** (Student Program Chair). Conference.

       IWSLT 2024 - Evaluation Campaigns, Subtitling (Organizer). Conference.

Area Chair/Reviewer

2026   **LANS-TTS - IEEE Access** (Journal)

2025   **ACL, EMNLP** (Conferences) – Area Chair

       NeurIPS, CLIC-it, Interspeech, IWSLT, ICASSP 2025 (SPADE Workshop) (Conferences) – Reviewer

       LANS-TTS - IEEE Access (Journal)

2024   **ACL, EMNLP, ICASSP, IWSLT, Interspeech** (Conferences) – Reviewer

2023   **HITLAML** (Conference) – Reviewer

2022   **AMTA 2022** (Conference) – Reviewer