# Manuela Sanguinetti

https://orcid.org/0000-0002-0147-2208

Main page

Google Scholar

DBLP

## **Employment**

o1/03/2023 – Current Assistant Professor. University of Cagliari. Department of Mathematics and Computer Science

01/06/2020 - 28/02/2023

**Post-Doc Researcher.** University of Cagliari, Department of Mathematics and Computer Science

21/02/2020 - 31/05/2020

**Research intern.** University of Turin, Department of Computer Science

01/08/2018 - 31/12/2020

**Post-Doc Researcher.** University of Turin, Department of Computer Science

01/07/2017 - 30/06/2018

**Post-Doc Researcher.** University of Turin, Department of Computer Science

01/08/2016 - 31/03/2017

**Research intern.** University of Pisa, Department of Computer Science

## **Teaching**

#### Courses

08/05/2024 - 30/05/2024 Computer skills (Instructor), Post-graduate School of Hospital Pharmacy, University of Cagliari

14/04/2023 - 10/05/2023

Computer skills (Instructor), Post-graduate School of Hospital Pharmacy, University of Cagliari

01/03/2022 - 26/05/2022

■ Programming 2 (Lab Tutor), Undergraduate course on Computer Science, University of Cagliari

### **Theses**

A.A. 2023-2024

Student: Luca Faccin. Thesis title: "Automatic detection of persuasion techniques in memes: comparing models and configurations". Master's Degree in Computer Science, University of Cagliari (Co-advisor)

A.Y. 2021-2022

Student: Matteo Pusceddu. Thesis title: "Entity Set Expansion – Tecniche e Modelli". Bachelor's Degree in Computer Science, University of Cagliari (Co-advisor)

Student: Michela Lecca. Thesis title: "Set Expansion and Labeling Using Masked Language Models". Bachelor's Degree in Computer Science, University of Cagliari (Co-advisor)

A.Y. 2020-2021

Students: Matteo Gioele Collu and Federico Sanna (joint thesis). Thesis title: "A Data-to-Text Conversational Agent for Dieting Support". Bachelor's Degree in Computer Science, University of Cagliari (Co-advisor)

A.Y. 2019/2020

Student: Marco Scalerandi. Thesis title: "Strumenti per l'analisi e il miglioramento di conversazioni utente-chatbot". Bachelor's Degree in Computer Science, University of Turin (Co-advisor)

## **Education**

2012 – 2016	<b>PhD in Computer Science.</b> University of Turin. Thesis title: Combining Approaches for
	a Parallel Treebank Development and Alignment.

2008 – 2011 **Master's Degree in Translation**. University of Turin.

2002 – 2008 **Bachelor's Degree in Foreign Languages**. University of Cagliari

## **Publications**

#### **Journals**

- P. F. Balestrucci, E. D. Nuovo, M. Sanguinetti, L. Anselma, C. Bernareggi, and A. Mazzei, "An educational dialogue system for visually impaired people," *IEEE Access*, vol. 12, pp. 150 502–150 519, 2024. 
  DOI: 10.1109/ACCESS.2024.3479883.
- M. Sanguinetti and M. Atzori, "Conversational agents for energy awareness and efficiency: A survey," *Electronics*, vol. 13, no. 2, 2024, ISSN: 2079-9292. ODI: 10.3390/electronics13020401.
- E. Di Nuovo, M. Sanguinetti, A. Mazzei, E. Corino, and C. Bosco, "VALICO-UD: Treebanking an Italian Learner Corpus in Universal Dependencies," *Italian Journal of Computational Linguistics*, vol. 8, no. 1, 2022. ODI: 10.4000/ijcol.1007.
- A. Mazzei, L. Anselma, M. Sanguinetti, *et al.*, "Anticipating User Intentions in Customer Care Dialogue Systems," *IEEE Transactions on Human-Machine Systems*, no. November, 2022, ISSN: 21682305. ODI: 10.1109/THMS.2022.3184400.
- M. Sanguinetti, C. Bosco, L. Cassidy, et al., "Treebanking user-generated content: a UD based overview of guidelines, corpora and unified recommendations," *Language Resources and Evaluation*, 2022, ISSN: 15728412. ODI: 10.1007/s10579-022-09581-9.
- F. Poletto, V. Basile, M. Sanguinetti, C. Bosco, and V. Patti, "Resources and benchmark corpora for hate speech detection: a systematic review," *Language Resources and Evaluation*, vol. 55, no. 2, pp. 477–523, 2021, ISSN: 15728412. ODI: 10.1007/s10579-020-09502-8.
- A. T. E. Capozzi, M. Lai, V. Basile, *et al.*, ""Contro L'Odio": A Platform for Detecting, Monitoring and Visualizing Hate Speech against Immigrants in Italian Social Media," *Italian Journal of Computational Linguistics*, vol. 6, no. 1, pp. 77–97, 2020. ODI: 10.4000/ijcol.659.

#### **Conferences and Workshops**

- A. T. Cignarella, M. Sanguinetti, S. Frenda, A. Marra, C. Bosco, and V. Basile, "QUEEREOTYPES: A multi-source Italian corpus of stereotypes towards LGBTQIA+ community members," in *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, N. Calzolari, M.-Y. Kan, V. Hoste, A. Lenci, S. Sakti, and N. Xue, Eds., Torino, Italia: ELRA and ICCL, May 2024, pp. 13 429–13 441. URL: https://aclanthology.org/2024.lrec-main.1176.
- E. D. Nuovo, M. Sanguinetti, P. F. Balestrucci, L. Anselma, C. Bernareggi, and A. Mazzei, "Educational dialogue systems for visually impaired students: Introducing a task-oriented user-agent corpus," in Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation, LREC/COLING 2024, 20-25 May, 2024, Torino, Italy, N. Calzolari, M. Kan, V. Hoste, A. Lenci, S. Sakti, and N. Xue, Eds., ELRA and ICCL, 2024, pp. 5507–5519. URL: https://aclanthology.org/2024.lrec-main.489.

- L. Zedda, A. Perniciano, A. Loddo, C. Di Ruberto, M. Sanguinetti, and M. Atzori, "Snarci at SemEval-2024 task 4: Themis model for binary classification of memes," in *Proceedings of the 18th International Workshop on Semantic Evaluation (SemEval-2024)*, A. K. Ojha, A. S. Doğruöz, H. Tayyar Madabushi, G. Da San Martino, S. Rosenthal, and A. Rosá, Eds., Mexico City, Mexico: Association for Computational Linguistics, Jun. 2024, pp. 853–858. URL: https://aclanthology.org/2024.semeval-1.122.
- M. Sanguinetti, M. Atzori, and N. Puddu, "Rewordqald9: A bilingual benchmark with alternative rewordings of QALD questions," in *Proceedings of Poster and Demo Track and Workshop Track of the 18th International Conference on Semantic Systems co-located with 18th International Conference on Semantic Systems (SEMANTiCS 2022), Vienna, Austria, September 13th to 15th, 2022*, U. Simsek, D. Chaves-Fraga, T. Pellegrini, and S. Vahdat, Eds., ser. CEUR Workshop Proceedings, vol. 3235, CEUR-WS.org, 2022.

  © URL: http://ceur-ws.org/Vol-3235/paper7.pdf.
- M. Atzori, A. Barenghi, S. Comai, et al., "Improving work life conditions via portable knowledge-driven recommender system (discussion paper)," in Proceedings of the 29th Italian Symposium on Advanced Database Systems, SEBD 2021, Pizzo Calabro (VV), Italy, September 5-9, 2021, S. Greco, M. Lenzerini, E. Masciari, and A. Tagarelli, Eds., ser. CEUR Workshop Proceedings, vol. 2994, CEUR-WS.org, 2021, pp. 498–505. 

  Our URL: http://ceur-ws.org/vol-2994/paper57.pdf.
- S. Balloccu, E. Reiter, M. G. Collu, F. Sanna, M. Sanguinetti, and M. Atzori, "Unaddressed challenges in persuasive dieting chatbots," in *UMAP 2021 Adjunct Publication of the 29th ACM Conference on User Modeling, Adaptation and Personalization*, 2021, pp. 392–395, ISBN: 9781450383677. ODI: 10.1145/3450614.3463602.
- A. T. Cignarella, M. Lai, A. Marra, and M. Sanguinetti, ""La ministro è incinta": A Twitter account of women's job titles in Italian," in *Proceedings of the 8th Italian Conference on Computational Linguistics* (CLiC-it 2021), 2021. ODI: https://doi.org/10.4000/books.aaccademia.10525.
- L. Anselma, M. Di Lascio, D. Mana, A. Mazzei, and M. Sanguinetti, "Content Selection for Explanation Requests in Customer-Care Domain," in 2nd Workshop on Interactive Natural Language Technology for Explainable Artificial Intelligence, 2020, pp. 5–10. URL: https://aclanthology.org/2020.nl4xai-1.3.
- A. T. Cignarella, V. Basile, M. Sanguinetti, C. Bosco, P. Rosso, and F. Benamara, "Multilingual Irony Detection with Dependency Syntax and Neural Models," in *Proceedings of the 28th International Conference on Computational Linguistics*, 2020, pp. 1346–1358. ODI: 10.18653/v1/2020.coling-main.116.
- A. T. Cignarella, M. Sanguinetti, C. Bosco, and P. Rosso, "Marking Irony Activators in a Universal Dependencies Treebank: The Case of an Italian Twitter Corpus," in *Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020)*, 2020, pp. 5098–5105, ISBN: 979-10-95546-34-4.
- M. Di Lascio, M. Sanguinetti, L. Anselma, et al., "Natural language generation in dialogue systems for customer care," in *Proceedings of the Seventh Italian Conference on Computational Linguistics (CliC-it 2021)*, vol. 2769, 2020. ODI: 10.4000/books.aaccademia.8450.
- M. Sanguinetti, C. Bosco, L. Cassidy, et al., "Treebanking User-Generated Content: A Proposal for a Unified Representation in Universal Dependencies," in Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020), 2020, pp. 5240–5250. URL: https://www.aclweb.org/anthology/2020.lrec-1.645.
- M. Sanguinetti, G. Comandini, E. D. Nuovo, et al., "Haspeede 2 @ EVALITA2020: overview of the EVALITA 2020 hate speech detection task," in Proceedings of the Seventh Evaluation Campaign of Natural Language Processing and Speech Tools for Italian. Final Workshop (EVALITA 2020), V. Basile,

- D. Croce, M. D. Maro, and L. C. Passaro, Eds., ser. CEUR Workshop Proceedings, vol. 2765, CEUR-WS.org, 2020. ODI: https://doi.org/10.4000/books.aaccademia.6897.
- M. Sanguinetti, A. Mazzei, V. Patti, M. Scalerandi, D. Mana, and R. Simeoni, "Annotating Errors and Emotions in Human-Chatbot Interactions in Italian," in *Proceedings of the 14th Linguistic Annotation Workshop*, 2020, pp. 148–159. URL: https://www.aclweb.org/anthology/2020.law-1.14.
- V. Basile, C. Bosco, E. Fersini, et al., "Multilingual Detection of Hate Speech Against Immigrants and Women in Twitter," in *Proceedings of the 13th International Workshop on Semantic Evaluation* (SemEval-2019), 2019, pp. 54–63. ODI: 10.18653/v1/s19-2081.
- A. T. Capozzi, M. Lai, V. Basile, et al., "Computational linguistics against hate: Hate speech detection and visualization on social media in the "Contro L'Odio" project," in *Proceedings of the Sixth Italian Conference on Computational Linguistics* (CliC-it 2019), vol. 2481, 2019.
- A. T. Cignarella, M. Sanguinetti, C. Bosco, P. Rosso, and D. Informatica, "Is This an Effective Way to Annotate Irony Activators?" In *Proceedings of the Sixth Italian Conference on Computational Linguistics* (CliC-it 2019), 2019.
- E. Di Nuovo, C. Bosco, A. Mazzei, and M. Sanguinetti, "Towards an Italian learner treebank in universal dependencies," in *Proceedings of the Sixth Italian Conference on Computational Linguistics* (CliC-it 2019), vol. 2481, 2019.
- C. Francesconi, C. Bosco, F. Poletto, and M. Sanguinetti, "Error analysis in a hate speech detection task: The case of Haspeede-TW at Evalita 2018," in *Proceedings of the Sixth Italian Conference on Computational Linguistics (CliC-it 2019)*, vol. 2481, 2019.
- V. Basile, M. Lai, and M. Sanguinetti, "Long-term social media data collection at the University of Turin," in *Proceedings of the Fifth Italian Conference on Computational Linguistics (CLiC-it 2018)*, Accademia University Press, 2018. ODI: 10.4000/books.aaccademia.3075.
- C. Bosco, F. Dell'Orletta, F. Poletto, M. Sanguinetti, and M. Tesconi, "Overview of the Evalita 2018 Hate Speech Detection Task," in *EVALITA Evaluation of NLP and Speech Tools for Italian: Proceedings of the Final Workshop*, 2018. ODDI: 10.4000/books.aaccademia.4503.
- M. Sanguinetti, C. Bosco, A. Lavelli, A. Mazzei, O. Antonelli, and F. Tamburini, "Postwita-UD: An Italian Twitter treebank in Universal Dependencies," in *Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC 2018)*, 2018, pp. 1768–1775, ISBN: 9791095546009. URL: https://aclanthology.org/L18-1279/.
- M. Sanguinetti, F. Poletto, V. Patti, C. Bosco, and M. Stranisci, "An Italian Twitter Corpus of Hate Speech Against Immigrants," in *Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC 2018)*, vol. 2006, 2018, pp. 2798–2805, ISBN: 979-10-95546-00-9. URL: https://aclanthology.org/L18-1443/.
- F. Poletto, M. Stranisci, M. Sanguinetti, V. Patti, and C. Bosco, "Hate Speech Annotation: Analysis of an Italian Twitter Corpus," in *Proceedings of the Fourth Italian Conference on Computational Linguistics* (CLiC-it 2017), Accademia University Press, 2017. ODI: 10.4000/books.aaccademia.2448.
- M. Sanguinetti, C. Bosco, A. Mazzei, A. Lavelli, and F. Tamburini, "Annotating Italian Social Media Texts in Universal Dependencies," in *Proceedings of the Fourth International Conference on Dependency Linguistics (Depling 2017)*, 2017, pp. 229–239. URL: https://www.aclweb.org/anthology/W17-6526.
- D. Zeman, M. Popel, M. Straka, et al., "CoNLL 2017 shared task: Multilingual parsing from raw text to universal dependencies," in CoNLL 2017 SIGNLL Conference on Computational Natural Language Learning, Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies, 2017, ISBN: 9781945626708. ODI: 10.18653/v1/K17-3001.

- M. Sanguinetti, "Experimenting the use of catenae in Phrase-Based SMT," in *Proceedings of the Second Italian Conference on Computational Linguistics (CLiC-it 2015)*, 2015, pp. 246–251. ODI: 10.4000/books.aaccademia.1548.
- E. Sulis, M. Lai, M. Vinai, and M. Sanguinetti, "Exploring sentiment in social media and official statistics: A general framework," in *Proceedings of the 2nd International Conference on Emotion and Sentiment in Social and Expressive Media: Opportunities and Challenges for Emotion-Aware Multiagent Systems*, vol. 1351, CEUR-WS.org, 2015, pp. 96–105.
- C. Bosco, L. Allisio, V. Mussa, et al., "Detecting happiness in Italian tweets: Towards an evaluation dataset for sentiment analysis in Felicittà," in Proceedings of the 5th International Workshop on EMOTION, SOCIAL SIGNALS, SENTIMENT & LINKED OPEN DATA, ES³LOD 2014, 2014, pp. 56–63.

  OURL: https://iris.unito.it/handle/2318/146318#.XG2EZf0vNTw.mendeley.
- C. Bosco, F. D. Orletta, S. Montemagni, M. Sanguinetti, and M. Simi, "The Evalita 2014 Dependency Parsing task," in *Proceedings of the First Italian Conference on Computational Linguistics CLiC-it 2014 & the Fourth International Workshop EVALITA 2014*, pp. 1–8, ISBN: 9788867414727.
- C. Bosco and M. Sanguinetti, "Towards a Universal Stanford Dependencies parallel treebank," in *Proceedings of the 13th International Workshop on Treebanks and Linguistic Theories (TLT13)*, 2014. OURL: https://hdl.handle.net/11022/0000-0000-2E8C-0.
- M. Sanguinetti and C. Bosco, "Converting the parallel treebank ParTUT in Universal Stanford Dependencies," in *Proceedings of the First Italian Conference on Computational Linguistics CLiC-it 2014 and of the Fourth International Workshop EVALITA 2014 9-11 December 2014, Pisa, 2014*, pp. 316–321.

  DOI: 10.12871/clicit2014161.
- M. Sanguinetti, C. Bosco, and L. Cupi, "Exploiting catenae in a parallel treebank alignment," in *Proceedings of the 9th International Conference on Language Resources and Evaluation, LREC 2014*, 2014, pp. 1824–1831, ISBN: 9782951740884.
- M. Sanguinetti, E. Sulis, V. Patti, et al., "Developing corpora and tools for sentiment analysis: the experience of the University of Turin group," in Proceedings of the First Italian Conference on Computational Linguistics CLiC-it 2014 & the Fourth International Workshop EVALITA 2014, 2014, pp. 322–327.
- M. Sanguinetti, C. Bosco, and L. Lesmo, "Dependency and constituency in translation shift analysis," in *Proceedings of the Second International Conference on Dependency Linguistics (DepLing 2013)*, 2013, pp. 282–291.
- C. Bosco, M. Sanguinetti, and L. Lesmo, "The Parallel-TUT: A multilingual and multiformat treebank," in *Proceedings of the 8th International Conference on Language Resources and Evaluation, LREC 2012*, 2012, pp. 1932–1938, ISBN: 9782951740877.
- M. Sanguinetti and C. Bosco, "Translational divergences and their alignment in a parallel multilingual treebank," in *Proceedings of the 11th Workshop on Treebanks and Linguistic Theories (TLT11)*, 2012.
- M. Sanguinetti and C. Bosco, "Building the multilingual TUT parallel treebank," in *Proceedings of the International Workshop on Annotation and Exploitation of Parallel Corpora*, 2011, pp. 19–28, ISBN: 9789544520212.

#### Chapters

M. Sanguinetti and C. Bosco, "PartTUT: The Turin University Parallel Treebank," in Harmonization and Development of Resources and Tools for Italian Natural Language Processing within the PARLI Project. Studies in Computational Intelligence, R. Basili, C. Bosco, R. Delmonte, A. Moschitti, and M. Simi, Eds., vol. 589, Springer International Publishing Switzerland, 2015, pp. 51–70, ISBN: 9783319142050. ODI: 10.1007/978-3-319-14206-7\_8.

### **Proceedings**

- M. Atzori, P. Ciaccia, M. Ceci, et al., Eds., Proceedings of the 32nd Symposium of Advanced Database Systems, Villasimius, Italy, June 23rd to 26th, 2024, vol. 3741, CEUR Workshop Proceedings, CEUR-WS.org, 2024. URL: https://ceur-ws.org/vol-3741.
- A. Lieto, C. Battaglino, D. P. Radicioni, and M. Sanguinetti, Eds., Proceedings of the 3rd International Workshop on Artificial Intelligence and Cognition, Turin, Italy, September 28-29, 2015, vol. 1510, CEUR Workshop Proceedings, CEUR-WS.org, 2015. URL: http://ceur-ws.org/vol-1510.

## **Projects**

- e.INS Ecosystem of Innovation for Next Generation Sardinia. SPOKE 7 Low Carbon Technologies for efficient energy system (WP6.1: Digitalisation and flexibility services in energy communities). Funded by the National Reform and Resilience Plan (PNRR). Duration: 2022-2025. Link: https://www.einssardinia.it/
- DEtect and Evaluate Manipulation of ONline information (DEMON) funded by MIUR under the PRIN 2022. Duration: 2023–2025. Link: http://demon.unica.it/
- High quality Open data Publishing and Enrichment (HOPE). Funded by the Italian Ministry of Higher Education. Duration: 2020-2023. Link: http://hope-prin.org/
- IHatePrejudice. Immigrants, Hate and Prejudice in Social Media. Funded by Fondazione Compagnia San Paolo, Turin. Duration: 2017-2019. Link: http://hatespeech.di.unito.it/ihateprejudice.html
- Hate Speech and Social Media. Analysis of Hate Speech Contents in Piedmont Region. Funded by Fondazione CRT, Turino. Duration: 2016-2018. Link: http://hatespeech.di.unito.it/crt.html

## **Committees**

## Management committees

Management Committee member of the COST Action Universality, Diversity and Idiosyncrasy in Language Technology (UniDive). Link: https://www.cost.eu/actions/CA21167/#tabs+Name:Management%20Committee

#### Scientific committees

2024	Area Chair of the Language Resources and Evaluation track of the 31st International Conference on Computational Linguistics (COLING 2025) – Co-chair of the Doctoral Consortium at the 32nd Symposium on Advanced Database Systems (SEBD 2024) – Member of the Program Committee of the 14th Italian Information Retrieval Workshop (IIR 2024)
2023	Program Committee dello <i>Student Research Workshop</i> della 17th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2023 SWR)
2022 - 2023	Workshop on Perspectivist Approaches to Disagreement in NLP (NLPerspectives 2022, 2023)
2020 - 2024	Member of the Program Committe of the Conference on Language Resources and Evaluation (LREC)
2020 - 2022	Member of the Program Committe of the Workshop on Resources and Techniques for User and Author Profiling in Abusive Language (ResT UP, ResT-UP2)
2018	Member of the Program Committe of the IEEE DSAA Special Session on Sentiment,

Emotion, and Credibility of Information in Social Data (SeCredISData) @DSAA 2018: In-

since 2024 Member of the Editorial Board of the Journal of Intelligent Information Systems

ternational Conference on Data Science and Advanced Analytics

## **Committees (continued)**

since 2017 Member of the Program Committe of the *Italian Conference of Computational Linguistics (CliC-it)* 

## Organizing committees

- Hate Speech Detection (HaSpeeDe2) task at EVALITA 2020. Link: http://www.di.unito.it/~tutreeb/haspeede-evalita20/index.html
- Multilingual detection of hate speech against immigrants and women in Twitter (HatE-val) at the International Workshop on Semantic Evaluation (SemEval). Link: https://competitions.codalab.org/competitions/19935
- Hate Speech Detection (HaSpeeDe) task at EVALITA 2018. Link: http://www.di.unito.it/~tutreeb/haspeede-evalita18/index.html
- Third edition of the International Workshop on Artifical Intelligence and Cognition (AIC 2015), Turin. Link: http://www.di.unito.it/~lieto/AIC2015/program\_committee.html
- Dependency Parsing task at EVALITA 2014. Link:https://www.evalita.it/campaigns/evalita-2014/tasks/dependency-parsing/