Mae Sosto (they/them)

mae.sosto@cwi.nl
mae-sosto MaeSosto
https://maesosto.github.io



About Myself

I am a computer scientist and AI researcher working at the intersection of Natural Language Processing (NLP) and queer-related themes, with a particular focus on social justice, LGBTQIA+ representation, and fairness in AI.
I am currently a PhD Candidate in the Human-Centered Data Analytics (HCDA) group at Centrum Wiskunde & Informatica (CWI) in Amsterdam, in collaboration with the Responsible AI group at Utrecht University, supervised by Laura Hollink and Davide Ceolin.
My research investigates how language technologies, particularly large language models (LLMs), represent queer identities, reproduce heterocisnormative assumptions, and reflect and reinforce societal biases and inequalities. Combining computational methods with critical social science perspectives, I examine bias, discrimination, and the representation of LGBTQIA+ communities in AI systems, aiming to challenge systemic injustices and contribute to more equitable, inclusive language technologies.

Research Publications

Bibliography

1. C. Ferrando, L. Draetta, M. Madeddu, *et al.*, "Multipride at evalita 2026: Overview of the multilingual detection of slur reclamation in the lgbtq+ context task," in *Proceedings of the 9th Evaluation Campaign of Natural Language Processing and Speech Tools for Italian. Final Workshop (EVALITA 2026)*, Apr. 2026.
2. M. Sosto, D. S. M. Pandiani, and L. Hollink, "QueerGen: How LLMs reflect societal norms on gender and sexuality in sentence completion task," in *Findings of the Association for Computational Linguistics: EACL 2026*, V. Demberg, K. Inui, and L. Marquez, Eds., Rabat, Morocco: Association for Computational Linguistics, Mar. 2026, pp. 4305–4326, ISBN: 979-8-89176-386-9. DOI: 10.18653/v1/2026.findings-eacl.225.
3. S. Weber, A. Wang, A. Gupta, *et al.*, "Queer nlp: A critical survey on literature gaps, biases and trends," 2026. arXiv: 2602.16151 [cs.CY]. URL: https://arxiv.org/abs/2602.16151.
4. M. Sosto and A. Barrón-Cedeño, "Queerbench: Quantifying discrimination in language models toward queer identities," 2024. arXiv: 2406.12399 [cs.CL]. URL: https://arxiv.org/abs/2406.12399.

References

Dr. Laura Hollink
Tenured Researcher, Head of Human-Centered Data Analytics Group
Centrum Wiskunde & Informatica (CWI)
Science Park 123, 1098 XG Amsterdam, Netherlands
L.Hollink@cwi.nl
+31 20 592 4047
https://homepages.cwi.nl/~hollink/

Dr. Delfina Sol Martinez Pandiani
Assistant Professor of Cultural Data Analysis, Faculty of Humanities
University of Amsterdam
Turfdraagsterpad 1, Postbus 94550 1090 GN Amsterdam, Netherlands
d.s.martinezpandiani@uva.nl
https://www.uva.nl/en/profile/m/a/d.s.c.martinez-pandiani/d.s.c.martinez-pandiani.html

Employment History

11/2024 – Current

PhD Candidate
Centrum Wiskunde & Informatica (CWI) · Amsterdam, Netherlands
At Centrum Wiskunde & Informatica (CWI), within the Human-Centered Data Analytics (HCDA) group, my research investigates how language technologies, particularly large language models (LLMs), represent queer identities, reproduce heterocisnormative assumptions, and reflect and reinforce societal biases and inequalities.
My work has contributed to projects such as QueerGen, which investigates how language models reflect societal norms surrounding gender and sexuality through sentence completion tasks, and Rainbow Meter, which evaluates LLMs' factual accuracy and stance on LGBTI rights across different languages and countries.
Alongside my research, I actively engage with the research community by presenting my work at academic conferences and workshops, participating in panels and discussions, and delivering guest lectures on bias in AI and the representation of queer communities in language technologies. I also contribute to teaching as a teaching assistant, supporting students in exploring AI-related topics and their broader societal implications.
Website: https://www.cwi.nl/en/

10/2023 – 09/2024

Software Developer
Wospee Srl · Bologna, Italy
Focused on developing user-friendly HR platforms using Python, MySQL, and AWS. Led initiatives in HR solutions, managed client relations and internal projects, and contributed to the strategic development roadmap. Integrated new functionalities with the Cezanne HR system.
Website: https://wospee.com

10/2021 – 30/09/2022

Software Developer
Lavoropiù · Bologna, Italy
Full-stack development of the company's management system. Used C#, JavaScript, ASP, HTML, CSS, SQL, Bootstrap, and React. Built hybrid mobile apps with Cordova and managed SQL Server databases. Provided system administration and client support.
Website: https://www.lavoropiu.it/

Employment History (continued)

08/2021 – 12/2021
Software Developer
Istituto Medico Legale · Bologna, Italy
Developed a web platform to connect legal and medical professionals with clients. Full-stack development using JavaScript, ASP, HTML, CSS, SQL, Bootstrap, and React.
Website: `https://www.istitutomedicolegale.it`

10/2019 – 11/2019
Software Developer
CHEBUONI.IT S.R.L · Bologna, Italy
Gained practical experience with React, Next.js, Vue.js, and JSON for building interactive frontend components and small-scale projects.
Website: `https://goodiewebgroup.com/`

11/2016 – 12/2016
Network Specialist
Lepida SPA · Bologna, Italy
Configured routers and switches, maintained servers, installed new network equipment, and managed VLANs. Studied key communication protocols.
Website: `http://www.lepida.it`

06/2016 – 07/2016
IT Support Technician
GPI · Cona (FE), Italy
Provided OS configuration, system backups, hardware replacement, and Help Desk support. Troubleshot issues with various devices including computers, printers, and tablets.
Website: `http://www.gpi.it`

04/2016 – 04/2016
Helpdesk
VAR Group · Funo (BO), Italy
Conducted quality control and provided Help Desk support. Completed CRM training and analyzed critical business processes.
Website: `http://www.vargroup.it`

03/2016 – 03/2016
IT Trainer
Cosworth Electronic LTD · Cambridge, United Kingdom
Gained insights into motorsport and automotive electronics R&D and production. Explored smartphone data for promoting wellbeing.
Website: `http://www.cosworth.co`

Education

09/2021 – 14/12/2023
Master's Degree in Computer Science
Alma Mater Studiorum – University of Bologna · Bologna, Italy
Final grade: 100 with honors
Thesis: *QueerBench: Quantifying Discrimination in Language Models towards Queer Identities*
Website: `https://www.unibo.it`

Education (continued)

09/2022 – 03/2023

Double Degree in Computer Science
Final grade: 1 out of 5 (1 is the best)
Bielefeld University · Bielefeld, Germany
Part of the double degree program in collaboration with the University of Bologna.

09/2017 – 03/2021

Bachelor's Degree in Computer Science
Alma Mater Studiorum – University of Bologna · Bologna, Italy
Final grade: 97 out of 110
Thesis: *Design and development of an application for clinical data management oriented to people affected by Thalassemia*
Website: `https://www.unibo.it`

09/2012 – 07/2017

Secondary School Diploma in IT
Istituto Tecnico Aldini Valeriani Sirani · Bologna, Italy
Final grade: 91 out of 100
Focused on Systems and Networks, Technologies, Information Systems Design, Telecommunications, and IT.
Basic knowledge of Telecommunications and Electronics.
Completed two internships abroad with practical experience in various companies.
Website: `http://www.iav.it/home`

Skills

Languages
Mother tongue: Italian.
Strong reading, writing and speaking competencies for English.
Basic knowledge and understanding of Spanish.

Coding
NLP, Machine Learning, Deep Learning, C, C++, Java, Python (excellent), PHP, R, SQL, Bash, Node.js, Ruby, JavaScript, ReactJS, HTML5, CSS3, ASP.NET Web Forms, ASP.NET Web API, REST/SOAP Web Services, JSON, XML, Android (basic).

Web Dev
HTML, CSS, JavaScript, React, Bootstrap, RESTful APIs, Web Services (SOAP/REST), ASP.NET, Android Studio.

Databases
MySQL, PostgreSQL, MongoDB, SQLite, SQL Server.

Misc.
Academic research, team collaboration, problem solving, social media communication tools, Natural Language Processing (NLP), AI, system networking, digital platform usability.

Driving License
B

Miscellaneous Experience

Certification

2022 — **B2 CEFR Language Certification (English)**
University Language Center, Alma Mater Studiorum – University of Bologna · Bologna, Italy
Website: `https://centri.unibo.it/cla/it`

2021 — **B2 English Language Certificate**
British School Group · Bologna, Italy
Website: `https://britishschool.it`

2016 — **B1 English Language Certificate**
Select English Cambridge · Cambridge, United Kingdom

2015 — **High Risk Safety Course (COD. 4 / 15A)**
Istituto Tecnico Aldini Valeriani Sirani · Bologna, Italy
Occupational health and safety training: Specific training for high-risk environments (12 hours).
Covered topics: workplace labeling, stress, PPE, microclimate and lighting, noise, vibrations, accident risks, radiation, chemical and biological hazards, emergency procedures, fall prevention, load handling, and more.
Website: `http://www.iav.it/home`

A2 English Language Course
British International School · London, United Kingdom
Address: Chiswick High Road 22

2014 — **General Training Course for Workers**
Istituto Tecnico Aldini Valeriani Sirani · Bologna, Italy
Occupational safety and health fundamentals as per Italian regulations.
Website: `http://www.iav.it/home`

Volunteering

2022 – Current — **LGBTQIA+ Activist**
Il Cassero LGBTQIA+ Center · Bologna, Italy
Active member of Il Cassero, contributing to community-building, advocacy, and awareness campaigns focused on LGBTQIA+ rights and intersectional justice.
Website: `https://cassero.it`

Some Prefer Cake Festival
Bologna, Italy
Ongoing volunteer for the international lesbian film festival "Some Prefer Cake," supporting event organization, audience engagement, and queer feminist cultural programming.
Website: `https://someprefercakefestival.com/en/`

Miscellaneous Experience (continued)

Projects

08/2023 – 12/2023

Master Thesis: QueerBench: Quantifying Discrimination in Language Models towards Queer Identities
University of Bologna · Bologna, Italy
This thesis investigates societal biases in NLP, focusing on the representation of LGBTQIA+ identities in large language models (LLMs). Using a template-based masked language modelling (MLM) approach, it assesses harmfulness in sentence completions based on subject identity (queer vs. non-queer) and pronoun usage (binary, neutral, neo). Results reveal significantly higher harmfulness in completions involving queer subjects. The work contributes to bias mitigation strategies in LLMs and supports more inclusive language processing.

12/2020 – 03/2021

Bachelor Thesis: TalApp – A Mobile App for People with Thalassemia
University of Bologna · Bologna, Italy
This thesis details the design, development, and testing of TalApp, an Android app supporting thalassemia patients in managing clinical data, transfusions, exams, and therapies. The app includes notification systems to assist with ongoing treatment compliance. Informed by patient and doctor consultation, TalApp was usability tested and is available on the Play Store.
Available at: `https://play.google.com/store/apps/details?id=com.MartinaSosto.talapp&gl=IT`