



# Mae Sosto

📍 **Home** : Via Ferrarese 156/30, 40131, Bologna, Italy

✉ **Email**: [mae.sosto@gmail.com](mailto:mae.sosto@gmail.com) 📞 **Phone**: (+39) 3312291798

🌐 **Website**: <https://maesosto.github.io>

🌐 **LinkedIn**: [www.linkedin.com/in/mae-sosto-8533b416b](https://www.linkedin.com/in/mae-sosto-8533b416b)

👤 **GitHub**: <https://github.com/MaeSosto>

**Gender**: Other **Date of birth**: 28/01/1998 **Nationality**: Italian

## ABOUT ME

Passionate about AI, I'm deeply intrigued by Natural Language Processing (NLP), fairness, diversity and inclusion. My academic journey culminated in a master's thesis project "QueerBench: Quantifying Discrimination in Language Models towards Queer Identities" which focus on assessing potential harm caused by sentence completions generated by Large Language Models (LLMs) concerning LGBTQIA+ individuals. This research has been adapted, accepted and presented at the Queer in AI workshop at NAACL 2024 conference.

Currently, I'm working as a full-time Python software developer at Wospee, where I contribute to innovative solutions and collaborate within dynamic teams.

Recently graduated with top honours from a master's program in Computer Science (LM) at the University of Bologna (Italy), with a double degree from Bielefeld University (Germany), I've consistently pursued excellence in both academia and practical application.

My professional experience spans across esteemed Italian IT enterprises, including Lavoropiù S.p.A., GPI S.p.A., Lepida S.p.A., and VAR Group, where I've honed my skills and gained invaluable industry insights.

## PUBLICATIONS

[ 2024 ] [QueerBench: Quantifying Discrimination in Language Models Toward Queer Identities](#)

**Reference**: Mae Sosto, Alberto Barrón-Cedeño

With the increasing role of Natural Language Processing (NLP) in various applications, challenges concerning bias and stereotype perpetuation are accentuated, which often leads to hate speech and harm. Despite existing studies on sexism and misogyny, issues like homophobia and transphobia remain underexplored and often adopt binary perspectives, putting the safety of LGBTQIA+ individuals at high risk in online spaces. In this paper, we assess the potential harm caused by sentence completions generated by English large language models (LLMs) concerning LGBTQIA+ individuals. This is achieved using QueerBench, our new assessment framework, which employs a template-based approach and a Masked Language Modeling (MLM) task. The analysis indicates that large language models tend to exhibit discriminatory behaviour more frequently towards individuals within the LGBTQIA+ community, reaching a difference gap of 7.2% in the QueerBench score of harmfulness.

## EDUCATION AND TRAINING

[ 09/2021 – 14/12/2023 ] **Master degree in Computer Science**

**Alma Mater Studiorum - Univeristy of Bologna** <https://www.unibo.it>

**Address:** Via Zamboni 33, 40126 , Bologna, Italy | **Field(s) of study:** Computer Sc |  
**Final grade:** 100L | **Thesis:** QueerBench: Quantifying Discrimination in Language Models towards Queer Identities

[ 09/2022 – 03/2023 ] **Double Degree in Computer Science**

**Bielefeld University** [https://www.uni-bielefeld.de/\(en\)/](https://www.uni-bielefeld.de/(en)/)

**City:** Bielefeld | **Country:** Germany |

[ 10/01/2022 ] **B2 level linguistic suitability of the CEFR at the university language center**

**Alma Mater Studiorum - Univeristy of Bologna** <https://centri.unibo.it/cla/it>

**Address:** Via Zamboni, 33, 40126 , Bologna, Italy |

[ 09/2017 – 03/2021 ] **Bachelor degree in Computer Science**

**Alma Mater Studiorum - Univeristy of Bologna** <https://www.unibo.it>

**Address:** Via Zamboni, 33, 40126 , Bologna, Italy | | **Final grade:** 97 | **Thesis:** Design and development of an application for clinical data management oriented to people affected by Thalassemia

[ 2016 – 2016 ] **IT Essentials course**

**Cisco Networking Academy**

- Define information technology (IT) and describe the components of a personal computer
- Describe how to protect people, equipment, and environments from accidents, damage, and contamination
- Perform a step-by-step assembly of a desktop computer
- Explain the purpose of preventive maintenance and identify the elements of the troubleshooting process
- Install and navigate an operating system
- Configure computers to attach to an existing network
- Describe the features, characteristics, and operating systems of mobile devices
- Implement basic hardware and software security principles
- Apply good communication skills and professional behavior while working with customers
- Perform preventive maintenance and advanced troubleshooting
- Assess customer needs, analyze possible configurations, and provide solutions or recommendations for hardware, operating systems, networking, and security

[ 04/06/2021 ] **B2 English language certificate**

**British School Group** <https://britishschool.it/>

**Address:** Via Zamboni 1, Bologna, Italy |

[ 09/2012 – 07/2017 ] **Secondary School Diploma in IT**

**Istituto Tecnico Aldini Valeriani Sirani** <http://www.iav.it/home>

**Address:** Via bassanelli 9/11, Bologna, Italy | | **Final grade:** 91

In-depth study of the subjects: Systems and networks, Technologies, Design of information systems, Telecommunications and Information Technology.

Basic knowledge of Telecommunications and Electronics.

During the five years, two internships abroad (with, therefore, related work experience in various companies).

[ 03/2016 – 03/2016 ] **English language certificate B1**

[ 2015 – 2015 ] **High Risk Safety Course COD. 4 / 15A**

**Istituto Tecnico Aldini Valeriani Sirani** <http://www.iav.it/home>

**Address:** Via Bassanelli 9/11,, Bologna, Italy |

Occupational health and safety training: Specific training - High risk

Contents As required by Article 4 of the State-Regions Agreement of 21/12/2011, the following have been addressed issues in relation to the risks related to the duties of the workers involved in the course, the characteristics of the company and al its sector / sector to which it belongs: - Work environments Labeling Work-related stress - Video terminals Risks physical - PPE Work organization - Microclimate and lighting - Noise Signs - General electrical - Vibration Accident risks - General mechanics - Radiation - Accidents and near accidents - Mist machines - Oils - Fumes - Vapors - Dust - Safety procedures with reference to the specific risk profile - Equipment - Explosion risks - Emergencies - Chemical risks - Falls from height - Exodus and fire procedures - Carcinogenic risks Manual handling loads - Organizational procedures for first aid - Biological risks Goods handling (equipment for - lifting, means of transport)

Duration 12 hours

[ 2015 – 2015 ] **English language course A2**

**British International School**

**Address:** Chiswick High Road 22, Londra, United Kingdom |

[ 2014 – 2014 ] **General training course for workers**

**Istituto Tecnico Aldini Valeriani Sirani** <http://www.iav.it/home>

**Address:** Via Bassanelli 9/11, Bologna, Italy |

## WORK EXPERIENCE

---

[ 10/10/2023 – Current ] **Software developer**

**Wospee Srl**

**City:** Bologna | **Country:** Italy

The role encompasses overseeing three key areas:

- HR Solutions Development: I lead initiatives aimed at developing innovative HR solutions tailored to meet our clients' needs effectively.
- Client Management and Internal Project Management: I actively engage in client interactions and oversee internal project management activities to ensure seamless project delivery and client satisfaction.
- Contribution to Future Development Roadmap: I play a pivotal role in shaping our future development roadmap, aligning our strategies with long-term objectives.

My primary focus revolves around collaborating within small teams to develop a user-friendly platform for clients to manage their HR processes seamlessly. This involves extensive work in Python, interfacing with MySQL databases, and leveraging AWS cloud infrastructure. Additionally, we are involved in integrating various new functionalities with the existing Cezanne HR system, enhancing its capabilities and performance.

[ 10/2021 – 30/09/2022 ] **Software Developer**

**Lavoropiù** <https://www.lavoropiu.it/>

**Address:** P.zza San Martino 1, Bologna, Italy

Analysis, development and programming of company management system both backend and frontend side.

Languages used: C #, Javascript, ASP, HTML, CSS, SQL with the support of Bootstrap and React frameworks.

Hybrid mobile development with Cordova for Android and IOS devices.

Management of the company database (SQL server).

System administrator and customer support if necessary.

[ 08/2021 – 12/2021 ]

### **Software Developer**

**Istituto Medico Legale**    <https://www.istitutomedicolegale.it>

**Address:** Via Basilicata 12, 64026 , Bologna, Italy

Analysis, development and programming of the company management system both backend and frontend side.

The goal was to build a web platform which allowed to create a connection between legal figures (such as lawyers), health professionals (doctors and specialists) and clients.

Languages used: Javascript, ASP, HTML, CSS, SQL with the support of Bootstrap and React frameworks.

[ 15/10/2019 – 04/11/2019 ]

### **Software Developer**

**CHEBUONI.IT S.R.L**    <https://goodiewebgroup.com/>

**Address:** Via Marsala 3, 40126, Bologna, Italy

My experience includes delving into the basics of React and its framework, Next.js, followed by the creation of mini-projects to apply these acquired skills. Additionally, I've studied data exchange management through JSON formats and applied this knowledge to develop mini-projects. Furthermore, I have experience in crafting frontend pages with interactive features using the JavaScript framework, Vue.js.

[ 11/2016 – 12/2016 ]

### **Network Specialist**

**Lepida SPA.**    <http://www.lepida.it>

**Address:** Viale Aldo Moro 64, 40127, Bologna, Italy

I specialize in configuring routers (access points) and switches, with a focus on managing extensive networks. My expertise includes server maintenance, data updates for new equipment installations, and in-depth knowledge of VLAN management. Additionally, I have studied key communication protocols.

[ 06/2016 – 07/2016 ]

### **Information Technology Support Technician**

**GPI**    <http://www.gpi.it>

**Address:** Via Aldo Moro 8, 44124, Cona (FE), Italy

I specialize in configuring, backing up, and restoring operating systems. Additionally, I have experience swapping out internal computer components, replacing printer components, and providing Help Desk support through remote connections and ticketing systems. My expertise extends to troubleshooting and resolving various issues with desktops, laptops, printers (including scanners), mobile devices, and tablets.

[ 04/2016 – 04/2016 ]

### **Helpdesk**

**VAR Group**    <http://www.vargroup.it>

**Address:** Via Gramsci 43,, Funo (BO), Italy

I performed quality control tasks and provided support to the Help Desk through remote connections and ticketing systems. Additionally, I completed a course on CRM (Customer Relationship Management) and conducted analyses of critical company processes.

[ 03/2016 – 03/2016 ]

### **Information technology trainers**

**Cosworth Electronic LTD**    [www.cosworth.com](http://www.cosworth.com)

**Address:** Brookfield Technology Centre, Twenty Pence Rd,, CB24 8PS, Cambridge, United Kingdom

Learn about:

- Motorsport and Automotive electronics R&D and Production
- Happier and Healthier with Smartphone Data

[ 2017 – 2019 ]

## **Babysitter**

**City:** Poggio Renatico | **Country:** Italy

Main professional skills acquired: sense of responsibility.

## **LANGUAGE SKILLS**

---

**Mother tongue(s):** Italian

**Other language(s):**

**English**

**LISTENING** B2 **READING** B2 **WRITING** B2

**SPOKEN PRODUCTION** B2 **SPOKEN INTERACTION** B2

**Spanish**

**LISTENING** A1 **READING** A1 **WRITING** A1

**SPOKEN PRODUCTION** A1 **SPOKEN INTERACTION** A1

*Levels: A1 and A2: Basic user; B1 and B2: Independent user; C1 and C2: Proficient user*

## **SOCIAL AND POLITICAL ACTIVITIES**

---

[ 2021 – Current ]

**Member of Global Cassero group at APS Arcigay Il Cassero** APS Arcigay Il Cassero, Bologna

I attend Global Cassero which is the the Cassero LGBT Center (Arcigay Center) international exchange group. The group has the goal to create and carry out initiatives relating to LGBTQIA + issues, involving the community and citizens interconnecting different groups and association through non formal education and Erasmus+ projects.

**Link:** <https://www.instagram.com/globalcassero/>

[ 05/05/2015 ]

### **TIM Girls Hackathon**

An initiative that aims to bring high school students closer to programming and reduce the

gender gap in technology.

The event - free and with tablet up for grabs for the winners - falls within the STEM month of the We are project

Peer of the Ministry of Education, University and Research. The aim of the day will be to create an app against Cyberbullying

[ 03/2017 – 04/2017 ]

### **Tutor of the course "The technology that helps you .. The pc and the mobile phone" with the CNA Pensionati Bologna association**

It is never too late to learn how to use the PC, at any age: this is demonstrated by the elderly of CNA Pensionati Bologna, around thirty, engaged in the course for computer literacy and the use of tablets and smartphones at the Aldini Valeriani.

They are tutored by the students of the Informatics section of the Aldini high school and this intergenerational exchange gives a special value to the initiative promoted by CNA Pensionati.

[ 08/2023 – 12/2023 ]

### QueerBench: Quantifying Discrimination in Language Models towards Queer Identities

Authors: Mae Sosto ([mae.sosto@gmail.com](mailto:mae.sosto@gmail.com)), Alberto Barrón-Cedeño ([a.barron@unibo.it](mailto:a.barron@unibo.it)), Beatrice Spallaccia ([beatrice.spallaccia@unibo.it](mailto:beatrice.spallaccia@unibo.it))

Abstract:

This thesis explores the evolving landscape of Natural Language Processing (NLP) and its intersection with societal biases, focusing on the LGBTQIA+ community. With the rise of computers in language comprehension, interpretation, and generation, NLP has become integral to various applications, posing challenges related to bias and stereotype perpetuation. As technology advances, there is a parallel emphasis on fostering inclusivity, particularly in digital spaces critical for the safety of LGBTQIA+ individuals.

Acknowledging the transformative influence of language on identity, this research underscores NLP's role in countering hate speech and bias online. Despite existing studies on sexism and misogyny, issues like homophobia and transphobia remain underexplored, often adopting binary perspectives. This behaviour not only marginalizes gender-diverse individuals but also perpetuates harmful behaviors.

The primary focus of this study is to assess the potential harm caused by sentence completions generated by large language models (LLMs) concerning LGBTQIA+ individuals. Employing a template-based approach, the investigation centres on the Masked Language Modelling (MLM) task and categorizes subjects into queer and non-queer terms, as well as neo-pronouns, neutral pronouns, and binary pronouns. The analysis reveals similarities in the assessment of pronouns by LLMs, with harmfulness rates around 6.1% for binary pronouns and approximately 5.4% and 4.9% for neo- and neutral pronouns. Sentences with queer terms (words that refer to a queer identity) as subjects peak at 16.4% harmfulness, surpassing non-queer subjects by 7.4%. This research contributes valuable insights into mitigating harm in language model outputs and promoting equitable language processing for the queer community.

[ 12/2020 – 03/2021 ]

### TalApp

<https://play.google.com/store/apps/details?id=com.MartinaSosto.talapp&gl=IT>

In a world increasingly inclined to the use of new technologies, we can observe the advantages obtained from the digitalization of healthcare and the spread of mHealth, "mobile health", a practice for which mobile devices act as an intermediary to manage data relating to medicine and health.

This paper analyzes the design, development, release and testing phases of TalApp, an Italian application for Android devices, dedicated to meeting the needs of people with thalassemia. Since the latter is a disease that requires a strict regime of analysis and controls, the pathology is best studied through consultation with doctors and patients, to find the appropriate way to provide help to their needs.

TalApp allows to annotate clinical data, allowing the patient to create a history of previous and future transfusions, exams and therapies. In addition, it provides a monitoring system based on notifications and alarms that urges users to better follow their ongoing therapies and scheduled health commitments.

A usability test was carried out through a survey, the functionalities and the degree of user satisfaction were analyzed. The app was then released on the Play Store.

## DIGITAL SKILLS

---

### My Digital Skills

#### Programming Languages

Python | C++ | C# | C | JavaScript

#### Frontend development

HTML | CSS | Framework Web ReactJS | Bootstrap | XML

#### Backend development

PHP | MySQL | ASP .NET Web form | JSON | ASP .NET Web API | SQL

#### Tools and platforms

Visual Studio Code | Linux | Microsoft Office | Github | Git

#### AI and NLP

Framework Deep Learning: Pytorch | Natural Language Processing | NLP | AI | Deep learning | Language models | • Machine Learning | NLP, Machine Learning Basics | Java

#### Android development

Android Studio | Android

## DRIVING LICENCE

---

**Cars:** B

## VOLUNTEERING

---

[ 2022 – Current ] **LGBTQIA+ Activist** Il Cassero LGBTQIA+ Center

**Link:** <https://cassero.it>

[ 2022 – Current ] **Volunteer at Some Prefer Cake, the international festival of lesbian** Bologna

Write here the description...

**Link:** <https://someprefercakefestival.com/en/>

[ 2019 – 2020 ] **Bartender** Lab DERIVA

In 2019 I sometimes worked as bartender and event organizer with the Uni LGBTQ association of the University of Bologna.