

Emanuele La Malfa

personal page

Work Experience

Post Doctoral Researcher

Benchmarking Large Language Models.

Jul. 2023 – Current
The Alan Turing Institute & University of Oxford, UK

- Principal Investigators: Prof. [Michael Wooldridge](#), [Nigel Shadbolt](#), and [Anthony Cohn](#).

Research Assistant

Enabling rapid adoption of artificial intelligence through an anonymized data protocol and explainable models.

Oct. 2019 – Mar. 2021
University of Oxford, UK

- Principal Investigator: Prof. [Marta Kwiatkowska](#).
- Collaboration with [GenieAI](#) and funded through the [InnovateUK](#) scheme.
- The collaboration led to a paper published at [EMNLP'20](#).

Consultant

IBM, Global Business Unit

Mar. 2018 – Aug. 2019
IBM, Milan, Italy

- Design, development, and deploy of a search engine for a public financial institution.
- Design, development, and deploy of a cloud document management system for a public financial institution.

Education

University of Oxford

PhD (DPhil) in Computer Science, supervised by Prof. [Marta Kwiatkowska](#).

Oct. 2019 – Nov. 2023
Oxford, UK

Polytechnic University of Milan

Master's Degree in Computer Science and Engineering.

Feb. 2015 – Sept. 2017
Milan, Italy

Polytechnic University of Milan

Bachelor's Degree in Computer Engineering.

Sept. 2011 – Sept. 2014
Milan, Italy

Publications

NLP and Linguistics

Language-Models-as-a-Service: Overview of a New Paradigm and its Challenges

Pre-print (long paper) - media coverage [here](#) and [here](#)

Emanuele La Malfa, Aleks Petrov, Frieder Simon, Christoph Weinhuber, Raza Nazar, Anthony Cohn, Nigel Shadbolt and Michael Wooldridge

Language Models Tokenizers Introduce Unfairness Between Languages

NeurIPS'23 (long paper, main track) - [website](#)

Aleksandar Petrov, **Emanuele La Malfa**, Philip Torr, Adel Bibi

Emergent Linguistic Structures in Neural Networks are Fragile

Working paper (long paper)

Emanuele La Malfa, Matthew Wicker, Marta Kwiatkowska

The King is Naked: on the Notion of Robustness for Natural Language Processing

AAAI'22 (long paper, main track) – **oral presentation** –

Emanuele La Malfa, Marta Kwiatkowska

On Guaranteed Optimal Robust Explanations for NLP Models

IJCAI'21 (long paper, main track)

Emanuele La Malfa, Rhiannon Micheltore, Agnieszka Zbrzeny, Nicola Paoletti, Marta Kwiatkowska

Assessing Robustness of Text Classification through Maximal Safe Radius Computation

EMNLP'20 (long paper)

Emanuele La Malfa, Min Wu, Luca Laurenti, Benjie Wang, Anthony Hartshorn, Marta Kwiatkowska

Complex Networks Theory and Deep Learning

Deep Neural Networks as Complex Networks

Pre-print (long paper)

Emanuele La Malfa, Gabriele La Malfa, Claudio Caprioli, Giuseppe Nicosia, Vito Latora

Characterizing Learning Dynamics of Deep Neural Networks via Complex Networks

IEEE-ICTAI'21 (long paper, main track)

Emanuele La Malfa, Gabriele La Malfa, Giuseppe Nicosia, Vito Latora

Teaching Experience

Machine Learning

Classes

2023 (MT)
University of Oxford, UK

Probabilistic Model Checking

Practical sessions

2023 (MT)
University of Oxford, UK

Ethical Computing in Practice

Practical sessions

2023 (HT)
University of Oxford, UK

Deep Learning in Healthcare

Practical sessions

2023 (HT)
University of Oxford, UK

Probabilistic Model Checking

Practical sessions

2022 (MT)
University of Oxford, UK

Machine Learning

Classes

2021 (MT)
University of Oxford, UK

Probabilistic Model Checking

Practical sessions

2021 (MT)
University of Oxford, UK

Fundamentals of Computer Science

Practical sessions

2016 (Oct.-Dec.)
Polytechnic University of Milan, Italy

Conferences and Workshops Organization

Benchmarking Large Language Models

Workshop Organizer

28/11/2023
The Alan Turing Institute, London, UK

LOD2020, LOD2021, LOD2022, LOD2023

Conference Chair

Lake District/Siena

Invited Lectures, Talks, and Presentations

On Robustness for Natural Language Processing

Group Talk

19/04/2023
ICREA, Barcelona

On the Notion of Robustness for Natural Language Processing

Departmental Talk

17/01/2023
King's College University of London, UK

Robustness for Natural Language Processing

Lecture – Deep Fridays

22/04/2022
University of Bologna, Italy

Explainable AI

Lecture – Advanced Artificial Intelligence Course

04/03/2022
Royal Holloway University of London, UK

Tutoring and Mentoring

Williams-Exeter Exchange Programme

Tutoring Saad Waheed and Alisa Kanganis, (Williams-Exeter Programme exchange students in machine learning).

2023

University of Oxford

Tutoring Edward Kusel and Aleksandar Radoslavov for their part-B projects.

2022

Lead the Future - Mentor

Lead the Future is a nonprofit organization that helps Italian talents in STEM disciplines to find their path to brilliant careers.

2022-current

Academic Reviewing and Volunteering

The Alan Turing Institute - Reviewer

Reviewers for the Turing Fellow Program - Panel "Fundamental Research in Data Science and AI"

October - December 2023

Ukrainian Global University - Interviewer

I interviewed Ukrainian students who want to study in a partner university abroad.

April-June 2023

The Kharkiv and Przemyśl Project

I spent a week in Przemyśl (Poland) as a volunteer to help refugees who arrived (returned) from (to) Ukraine.

August 2022

Ukrainian Global University - Interviewer

I interviewed a dozen of prospective undergraduate Ukrainian students who want to study in a partner university abroad.

June-May 2022

Eutanasia Legale - Volunteer

I have collected signatures for a referendum to decriminalize euthanasia. The overall campaign gathered 1.2 million valid signatures.

July 2021