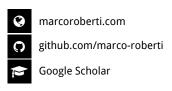


## Ph.D. Candidate





## **ABOUT ME**

I'm Marco Roberti, a Ph.D. candidate in the Department of Computer Science of the University of Turin. My main research interest is the application of Deep Neural Networks to the Data-to-Text Generation task. I collaborate with the Sorbonne University (France) and the Italian National Institute for Astrophysics (INAF). Research and personal projects are available on GitHub.

PROG. LANGUAGES HUMAN LANGUAGES

Python: expert (♦ PyTorch • HuggingFace • ♦ PyTorch-Lightning • ↑ Keras • OpenNMT)

Java: skilled (Android • JavaFX • Google Maps)

C: prior experience (UNIX system calls • teaching)

English - proficient
French - fluent
Italian - native

**EXPERIENCE** 

2018 – 2022 Ph.D, Computer Science University of Turin (IT)

I develop novel algorithms for Data-to-Text Generation, focusing on character-based representation and hallucination-free models. Besides, I investigate the use of machine learning techniques for anomaly detection in ESA's Gaia space mission data.

Deep Learning / Data-to-Text Generation / NLP

2019 Research intern LIP6 – Sorbonne University (FR)

I built a character-based RNN encoder-decoder model for data-to-text generation (E2E Challenge dataset). This internship resulted in the publication of *Copy Mechanism and Tailored Training for Character-Based Data-to-Text Generation*.

PyTorch / RNNs / Sequence-to-Sequence architectures

2018 – 2021 Adjunct lecturer University of Turin (IT)

M.Sc. in Artificial Intelligence: Deep Learning lectures, including the presentation of my Ph.D. research work.

M.Sc. in Military Strategy: C and UNIX programming lab sessions.

**EDUCATION** 

2016 – 2018 M.Sc, Computer Science – summa cum laude University of Turin (IT)

Machine Learning / Statistics / Artificial Intelligence

2016 M.Sc, Computer Science (Semester abroad) University of Montpellier (FR)

NLP / Image Processing / Distributed algorithms

2011 – 2015 B.Sc, Computer Science – summa cum laude University of Turin (IT)

Algorithms / Data structures / Design patterns

ACADEMIA & PROJECTS

First-authored publications (complete record available on Scholar):

Controlling Hallucinations at Word Level in Data-to-Text Generation

C. Rebuffel & M. Roberti (equal contribution), L. Soulier, G. Scoutheeten, R. Cancelliere, P. Gallinari — Springer's Data Mining

and Knowledge Discovery, ECML DKDD 2022 (Journal Track)

and Knowledge Discovery, ECML-PKDD 2022 (Journal Track)

2019 Copy Mechanism and Tailored Training for Character-Based Data-to-Text Generation

M. Roberti, G. Bonetta, R. Cancelliere, P. Gallinari — ECML-PKDD 2019

2021 Anomaly Detection Techniques in the Gaia Space Mission Data

M. Roberti, A.Druetto, D. Busonero, R. Cancelliere, D. Cavagnino, M. Gai — Journal of Signal Processing Systems

Personal projects:

2021 @Calend AI Twitter bot

An automated Italian Twitter profile backed by a small toy 😕 Language Model. It mimics the peculiar writing style and

content of an Italian public figure.

2018 **E2E DataSet for PyTorch**The E2E Challenge Dataset, packed as a PyTorch DataSet subclass.

SKILLS

**Curiosity** took me to computer science, still guiding my research inspiration.

I face my tasks with serious **responsibility**.

My approach when the play gets hard: problem-solving & having-fun!

EACL Conference Springer's Neural Processing Letters WebNLG@INLG2020 (also as a member of the PC)

 $\mathbf{C}$ 

FREE TIME

**Leading** the Silviadizenzero local association: I engage in environmental disclosure in public and on the radio, organize actions and social events – including the annual "Ginger Fest" –, and manage the volunteers. Keynote **speaker and moderator** in the Digital Ethics Forum (2019-2021). Taking care of my physical and mental health via **swimming** and bodyweight training.