Giuseppe **Attanasio**

Postdoctoral Researcher · Bocconi University

Via Roberto Sarfatti, 25, 20100 Milan, Italy

gattanasio.cc | giuseppeattanasio6@gmail.com

g8a9 | giuseppe-attanasio | @peppeatta | Scholar

"Computers aren't the thing. They are the thing that gets us to the thing." - Halt And Catch Fire

Bio

I am a Postdoctoral Researcher in the MilaNLP Group at the Department of Computing Sciences, Bocconi University. My research interests lay at the intersection of interpretability, fairness, and safety.

In my spare time, I like reading and watching Sci-Fi, playing basketball. I am also into DIY and 3D printing. I am a failed guitarist and pizzaiolo.

I update my website more frequently than this document, with recent professional news, projects, and publications.

Education

Politecnico di Torino — *Torino, Italy*

Ph.D. at Department of Control and Computer Engineering — *Nov. 2018 - Oct. 2022*

- Advisor: Elena Baralis

Politecnico di Torino — *Torino, Italy*

MD in Computer Engineering, Data Science track — *Oct. 2016 - Oct. 2018*

- Grade: 110/110 *cum Laude*

Politecnico di Torino — *Torino, Italy*

BD in Computer Engineering — *Oct. 2013 - Oct. 2016*

- Grade: 110/110

Publications

I believe in open-access research. Most of my papers are free to read, and the experimental code is open-sourced whenever possible. All the pointers are collected on my website.

Selected

Giuseppe Attanasio et al. "Entropy-based Attention Regularization Frees Unintended Bias Mitigation from Lists". In: *Findings of the Association for Computational Linguistics: ACL 2022*. Dublin, Ireland: Association for Computational Linguistics, May 2022, pp. 1105–1119. DOI: 10.18653/v1/2022.findings-acl.88. URL: https://aclanthology.org/2022.findings-acl.88

Giuseppe Attanasio et al. "ferret: a Framework for Benchmarking Explainers on Transformers". In: *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*. Dubrovnik, Croatia: Association for Computational Linguistics, May 2023, pp. 256–266. URL: https://aclanthology.org/2023.eacl-demo.29

Federico Bianchi et al. *Safety-Tuned LLaMAs: Lessons From Improving the Safety of Large Language Models that Follow Instructions*. arXiv:2309.07875 [cs]. Sept. 2023. DOI: 10.48550/arXiv.2309.07875. URL: http://arxiv.org/abs/2309.07875 (visited on 09/15/2023)

Alkis Koudounas et al. "Exploring subgroup performance in End-to-End speech models". In: *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. 2023

Chronologically

Federico Bianchi et al. *Safety-Tuned LLaMAs: Lessons From Improving the Safety of Large Language Models that Follow Instructions*. arXiv:2309.07875 [cs]. Sept. 2023. DOI: 10.48550/arXiv.2309.07875. URL: http://arxiv.org/abs/2309.07875 (visited on 09/15/2023)

Eliana Pastor et al. *Explaining Speech Classification Models via Word-Level Audio Segments and Paralinguistic Features*. arXiv:2309.07733 [cs, eess]. Sept. 2023. DOI: 10.48550/arXiv.2309.07733. URL: http://arxiv.org/abs/2309.07733 (visited on 09/15/2023)

Paul Röttger et al. *XSTest: A Test Suite for Identifying Exaggerated Safety Behaviours in Large Language Models*. arXiv:2308.01263 [cs]. Aug. 2023. DOI: 10.48550/arXiv.2308.01263. URL: http://arxiv.org/abs/2308.01263 (visited on 09/15/2023)

Helena Bonaldi et al. *Weigh Your Own Words: Improving Hate Speech Counter Narrative Generation via Attention Regularization*. arXiv:2309.02311 [cs]. Sept. 2023. DOI: 10.48550/arXiv.2309.02311. URL: http://arxiv.org/abs/2309.02311 (visited on 09/15/2023)

Alkis Koudounas et al. "ITALIC: An ITALian Intent Classification Dataset". In: *INTERSPEECH 2023*. IEEE. 2023

Amanda Cercas Curry et al. "MilaNLP at SemEval-2023 Task 10: Ensembling Domain-Adapted and Regularized Pretrained Language Models for Robust Sexism Detection". In: *Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval-2023)*. Toronto, Canada: Association for Computational Linguistics, July 2023, pp. 2067–2074. DOI: 10.18653/v1/2023.semeval-1.285. URL: https://aclanthology.org/2023.semeval-1.285 (visited on 09/15/2023)

Patrick John Chia et al. *E Pluribus Unum: Guidelines on Multi-Objective Evaluation of Recommender Systems*. arXiv:2304.10621 [cs]. Apr. 2023. DOI: 10.48550/arXiv.2304.10621. URL: http://arxiv.org/abs/2304.10621 (visited on 09/15/2023)

Vittorio Pipoli et al. "Squeeze and Learn: Compressing Long Sequences with Fourier Transformers for Gene Expression Prediction". In: (2023)

Alkis Koudounas et al. "Exploring subgroup performance in End-to-End speech models". In: *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. 2023

Giuseppe Attanasio et al. "Is It Worth the (Environmental) Cost? Limited Evidence for the Benefits of Diachronic Continuous Training". In: *arXiv preprint arXiv:2210.07365* (2022)

Patrick John Chia et al. "Contrastive Language and Vision Learning of General Fashion Concepts". In: *Scientific reports* (2022)

Giuseppe Attanasio et al. "ferret: a Framework for Benchmarking Explainers on Transformers". In: *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*. Dubrovnik, Croatia: Association for Computational Linguistics, May 2023, pp. 256–266. URL: https://aclanthology.org/2023.eacl-demo.29

Jacopo Tagliabue et al. "EvalRS: a Rounded Evaluation of Recommender Systems". In: *arXiv preprint arXiv:2207.05772* (2022)

Debora Nozza, Federico Bianchi, and Giuseppe Attanasio. "HATE-ITA: Hate Speech Detection in Italian Social Media Text". In: *Proceedings of the Sixth Workshop on Online Abuse and Harms (WOAH)*. Seattle, Washington (Hybrid): Association for Computational Linguistics, July 2022, pp. 252–260. DOI: 10.18653/v1/2022.woah-1.24. URL: https://aclanthology.org/2022.woah-1.24

Giuseppe Attanasio, Debora Nozza, and Federico Bianchi. "MilaNLP at SemEval-2022 Task 5: Using Perceiver IO for Detecting Misogynous Memes with Text and Image Modalities". In: *Proceedings of the 16th International Workshop on Semantic Evaluation (SemEval-2022)*. Seattle, United States: Association for Computational Linguistics, July 2022, pp. 654–662. DOI: 10.18653/v1/2022.semeval-1.90. URL: https://aclanthology.org/2022.semeval-1.90

Giuseppe Attanasio et al. "Benchmarking Post-Hoc Interpretability Approaches for Transformer-based Misogyny Detection". In: *Proceedings of NLP Power! The First Workshop on Efficient Benchmarking in NLP*. Dublin, Ireland: Association for Computational Linguistics, May 2022, pp. 100–112. DOI: 10.18653/v1/2022.nlppower-1.11. URL: https://aclanthology.org/2022.nlppower-1.11

Giuseppe Attanasio et al. "Entropy-based Attention Regularization Frees Unintended Bias Mitigation from Lists". In: *Findings of the Association for Computational Linguistics: ACL 2022*. Dublin, Ireland: Association for Computational Linguistics, May 2022, pp. 1105–1119. DOI: 10.18653/v1/2022.findings-acl.88. URL: https://aclanthology.org/2022.findings-acl.88

Patrick John Chia et al. "Fashionclip: Connecting language and images for product representations". In: *arXiv preprint arXiv:2204.03972* (2022)

Giuseppe Attanasio et al. “E-MIMIC: Empowering Multilingual Inclusive Communication”. In: *Proceedings of the First International Workshop on Data science for equality, inclusion and well-being challenges*. Dec. 2021. URL: https://iris.polito.it/handle/11583/2946252 (visited on 12/21/2021)

Federico Bianchi et al. “Contrastive Language-Image Pre-training for the Italian Language”. In: *arXiv:2108.08688 [cs]* (Aug. 2021). arXiv: 2108.08688. URL: http://arxiv.org/abs/2108.08688 (visited on 08/21/2021)

Giuseppe Attanasio and Eliana Pastor. “PoliTeam @ AMI: Improving Sentence Embedding Similarity with Misogyny Lexicons for Automatic Misogyny Identification in Italian Tweets”. en. In: *EVALITA Evaluation of NLP and Speech Tools for Italian - December 17th, 2020*. Ed. by Valerio Basile et al. Accademia University Press, 2020, pp. 48–54. ISBN: 9791280136329. DOI: 10.4000/books.aaccademia.6807. URL: http://books.openedition.org/aaccademia/6807 (visited on 08/16/2021)

Giuseppe Attanasio, Luca Cagliero, and Elena Baralis. “Leveraging the explainability of associative classifiers to support quantitative stock trading”. In: *Proceedings of the Sixth International Workshop on Data Science for Macro-Modeling*. DSMM ’20. New York, NY, USA: Association for Computing Machinery, June 2020, pp. 1–6. ISBN: 978-1-4503-8030-0. DOI: 10.1145/3401832.3402679. URL: https://doi.org/10.1145/3401832.3402679 (visited on 11/15/2020)

G. Attanasio et al. “DSLE: A Smart Platform for Designing Data Science Competitions”. In: *2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*. ISSN: 0730-3157. July 2020, pp. 133–142. DOI: 10.1109/COMPSAC48688.2020.00026

Luca Cagliero et al. “Training ensembles of faceted classification models for quantitative stock trading”. en. In: *Computing* 102.5 (May 2020), pp. 1213–1225. ISSN: 1436-5057. DOI: 10.1007/s00607-019-00776-7. URL: https://doi.org/10.1007/s00607-019-00776-7 (visited on 11/15/2020)

G. Attanasio et al. “Combining News Sentiment and Technical Analysis to Predict Stock Trend Reversal”. In: *2019 International Conference on Data Mining Workshops (ICDMW)*. ISSN: 2375-9259. Nov. 2019, pp. 514–521. DOI: 10.1109/ICDMW.2019.00079

Giuseppe Attanasio et al. “Quantitative cryptocurrency trading: exploring the use of machine learning techniques”. In: *Proceedings of the 5th Workshop on Data Science for Macro-modeling with Financial and Economic Datasets*. DSMM’19. New York, NY, USA: Association for Computing Machinery, June 2019, pp. 1–6. ISBN: 978-1-4503-6823-0. DOI: 10.1145/3336499.3338003. URL: https://doi.org/10.1145/3336499.3338003 (visited on 11/15/2020)

G. Attanasio et al. “HOT: Hold your own tools for AR-based constructive art”. In: *2017 IEEE Symposium on 3D User Interfaces (3DUI)*. Mar. 2017, pp. 256–257. DOI: 10.1109/3DUI.2017.7893369

Research projects

The following is a list of research projects I carried out or where I have contributed to the conceptualization of the work. Projects include collaborations with both academic peers and industry.

Fairness in End-to-End Speech Representation Models In this ongoing collaboration, we aim to study how modern neural end-to-end speech models (e.g., Wav2Vec 2.0 or HuBERT) behave on specific subgroups of the dataset. The end goal is to detect bias toward under-represented social groups or recording conditions where models under-perform.

Entities involved: Amazon Alexa AI.

Gender Inclusive Communication in Italian. E-MIMIC (Empowering Multilingual Inclusive comMunICation) is a joint effort of classical and computational linguistics communities in fighting against non-inclusive, prejudiced language forms. We use modern transformer-based language models to learn inclusive language in written text. Currently, we focus on Academic and Public Administration Italian documents with the goal of non-inclusive content identification and rephrasing.

Entities involved: University of Turin, University of Bologna, University of Bergamo.

Contrastive Language-Image Pre-training for the Italian Language. During the HuggingFace JAX Community Week, we had the chance to specialize OpenAI’s CLIP for the Italian language. CLIP is one of the most recent multi-modal models that connect images and text. The original model is limited to English, so we extended its capabilities. We hence presented the first CLIP model for the Italian Language, trained on more than 1.5 million high-quality image-text pairs. The released model outperforms the multilingual CLIP model on image retrieval and zero-shot classification tasks.
The project was selected among the 15 finalist projects - out of 100 - of the Flax/JAX Community Week

organized by HuggingFace in partnership with Google and received a special nominee in the final evaluation round.

Entities involved: Bocconi University, University of Milano-Bicocca, University of Groningen.

EvalRS: Rounded Evaluation of Recommender Systems Standard evaluation metrics for Recommender Systems fail to capture important aspects, such as behavioral testing and fairness on subgroups of users. In this project, we proposed EvalRS, an open-source framework, a data challenge, and a workshop (CIKM 2022) to foster more rounded evaluations of RecSys.

Entities involved: Coveo, Coveo Labs, Stanford University, Microsoft, NVIDIA.

Vision and Language Multi-Modality for General Fashion Concepts Learning The cost of training tailored supervised models can be prevented by learning general representations of concepts and reusing them. We show CLIP (Contrastive Language-Italian Pre-Training) can learn general fashion concepts (e.g., long/short sleeves, high/low heels) efficiently and effectively.

Entities involved: Coveo, Coveo Labs, Stanford University, Farfetch.

MALTO: Machine Learning at Polito. I co-founded MAchine Learning At poliTO (MALTO), a university-funded student team at Politecnico di Torino.

The team's main goal is to participate in various data science competitions. These competitions typically involve achieving high performance on machine learning problems (e.g., classification or regression) and proposing novel, cutting-edge approaches.

Entities involved: Politecnico di Torino.

Unsupervised Learning for Smart Fleet Management. I joined a joint research project between my research group and the company KMaster (Telepass SpA). We designed and implemented an end-to-end machine learning and clustering-based pipeline to characterize driving behaviors and fleet management.

Entities involved: KMaster (Telepass Group).

Public Activities & Teaching

Research Community

Workshops. I co-organized the first edition of EvalRS: a Rounded Evaluation of Recommender Systems which has been a data challenge and workshop during CIKM 2022. The event fosters a more rounded evaluation of recommender systems, encompassing aspects beyond standard evaluation, such as behavioral testing and group fairness.

University

M2L School. I co-tutored the second edition of the Mediterranean Machine Learning (M2L) Summer School, with ownership of the Natural Language Processing part. We prepared a shared notebook to guide students through 1) implementing a Transformer from scratch, 2) training the model for language modeling and machine translation, and 3) testing on gender bias.

Teaching Assistant. I have been a teaching assistant for several courses at Politecnico di Torino. In total, I have held 171 hours of complementary teaching.

- (2021) **Data science lab: processes and methods**, 60h, MD in Data Science and Engineering
- (2020) **Data science lab: processes and methods**, 39h, MD in Data Science and Engineering
- (2020) **Business Intelligence for Big Data**, 21h, MD in Industrial engineering and management
- (2019) *** Data science lab: processes and methods**, 39h, MD in Data Science and Engineering
- (2019) **Business Intelligence for Big Data**, 21h, MD in Industrial engineering and management
- (2018) **Basi di dati**, 21h, BD in Computer Engineering

- (2018) **The Fourth Industrial Revolution: Promises and Pitfalls in Blending New and Traditional Approaches in Manufacturing and Service Sectors**, 30h, Alta Scuola Politecnica School

* While Introduction to Databases and Business Intelligence were pre-existent and consolidated courses, I have been a major contributor in shaping *Data Science Lab: process and methods*, launched in September 2019 and currently one of the central courses in the Data Science and Engineering master degree at Politecnico.

The course is the first introduction to the Python programming language and the basics of Data Science and Machine Learning libraries for MD students. We worked hard to provide students with comprehensive exercises and solutions (10 laboratories, for a total of 60+ pages of lab exercises and 250+ pages of solutions). All the material is freely available on the course website.

Research Bites. I have devised and contributed to the launch of Research Bites, a series of short research talks and seminars held by Ph.D. students for students of the course Data Science Lab: process and methods. The goal of RBs is to disseminate cutting-edge research topics, in short, high-level pills. The series is now in its second edition.

Master Thesis Supervisor. I have supervised the work of 9 master students.

Industry

I have held consulting hours and courses twice.

- (2021) **Data Theory: Data Visualization with Python**
  Reply: Digital Services, Technology and Consulting, Turin (online)
- (2020) **Python technologies for Data Analytics**
  Applied Mechatronic Engineering & Technologies, Turin

Peer reviewing

I reviewed at least one work submitted to the following venues:

- ACL Rolling Review. Editions: 2021, 2022
- ACM KDD SIGKDD Conference On Knowledge Discovery And Data Mining. Editions: 2020, 2021
- ACM SIGMOD/PODS International Conference on Management of Data. Editions: 2021
- IEEE ICDE: IEEE International Conference on Data Engineering. Editions: 2020
- IEEE ICDM: IEEE International Conference on Data Mining. Editions: 2021
- ACM SAC: ACM/SIGAPP Symposium On Applied Computing. Editions: 2021
- EDBT: International Conference on Extending Database Technology. Editions: 2020, 2021
- DaWaK: International Conference on Big Data Analytics and Knowledge Discovery. Editions: 2019, 2021

I reviewed at least one work submitted to the following journals:

- Expert Systems With Applications, Elsevier
- Future Generation Computer Systems, Elsevier
- Machine Learning With Applications, Elsevier

Work Experience

Kupata S.r.l. *Torino, Italy*

Founder *Nov. 2016 - 2020*

- Kupata's main goal is to streamline the Lost and Found process. It brings innovation with a solution that helps people in returning items in a simple, secure, and undisclosed way. The business involves a physical object, the Kupa, and a social community that encourages members to act in the right way.

Consoft Sistemi S.p.A. *Torino, Italy*

Curricular intern *Mar. 2016 - Jul. 2016*

- I participated to the bootstrap phase of the Consoft's proprietary Knowledge Base platform. We built our solution upon Orange HRM, an open source PHP-based platform.

Skills

I am familiar with *italicized* entries.

Programming	Python, C, C++, Java, *C#, JavaScript, PHP, Matlab, SQL*
Deep Learning	PyTorch, Tensorflow, JAX, Keras
Framework	*Apache Hadoop, Apache Spark*
Scripting	Bash, *Awk*
DevOps	*AWS*, *GCP*, *Docker*
Front-end	*Hugo, React, Dash*
Graphics	Inkscape, GIMP, *Blender, Unity 3D*
Languages	Italian, English, *Spanish, French*