

DR. GIACOMO FRISONI



I'm a PhD Student in Computer Science and Engineering at the University of Bologna. I investigate how to combine language models and structured knowledge for Natural Language Processing and Understanding in Health domains.

// CONTACT

- ✉ giacomo.frisoni@unibo.it
- 📍 Via dell'Università, 50
47522 Cesena (FC), Italy
- ☎ +39 0547 338820 (office)
+39 331 2718059 (personal)
- 🏠 Personal website

Research and academy

- 📄 0000-0003-1104-2014
- 📄 DBLP
- 📄 Google Scholar
- 📄 Scopus
- 🏛 University of Bologna
- 👥 DISI UniBo NLP Research Group

Development

- 🐙 @GiacomoFrisoni
- 📺 @gfrisoni

Social and misc

- 🌐 Giacomo Frisoni

// CITIZENSHIP

- 🇮🇹 Italy

// LANGUAGES

- 🇮🇹 Italian Mother language
- 🇬🇧 English B2 level

// RESEARCH INTERESTS

- Neuro-Symbolic Learning
- Natural Language Processing
- Natural Language Understanding
- Text Mining
- Semantic Parsing
- Explainable AI
- Knowledge Graphs
- Knowledge Representation and Reasoning
- Graph Neural Networks
- Deep Representation Learning
- AI for Health informatics

// 🎓 EDUCATION

- 📅 10/2020 - 10/2023
- 📍 CESENA (FC), ITALY

Ph.D., Computer Science and Engineering
Department of Computer Science and Engineering,
University of Bologna

- Ministerial scholarship
- Supervisor: Gianluca Moro
- Tutor: Antonella Carbonaro
- Area of study: Natural Language Understanding

- 📅 09/2017 - 03/2020
- 📍 CESENA (FC), ITALY

M.S., Computer Science and Engineering
University of Bologna

- Cumulative GPA: 4.0 (transcript)
- Graduation score: 110/110 *cum Laude*
- Graduation class: Computer Engineering
- Thesis topic: *A new unsupervised methodology of descriptive text mining for knowledge graph learning*
- Supervisor: Gianluca Moro
- Co-Supervisor: Antonella Carbonaro
- Area of study: Text Mining

- 📅 09/2014 - 10/2017
- 📍 CESENA (FC), ITALY

B.S., Computer Science and Engineering
University of Bologna

- Cumulative GPA: 4.0 (transcript)
- Graduation score: 110/110 *cum Laude*
- Graduation class: Computer Engineering
- Thesis topic: *Design and development of a software system for studying and researching rare diseases*
- Supervisor: Dario Maio
- Area of study: Databases

- 📅 09/2009 - 07/2014
- 📍 RIMINI (RN), ITALY

Scientific High School
ITIS Leonardo Da Vinci

- Final score: 100/100 *cum Laude*
- Member of the National Register of Excellence
- Focus on Cryptography (Bletchley Park Visitor)

// ⚙️ EXPERIENCE

- 📅 03/2020 - 10/2020
- 📍 CESENA (FC), ITALY

Pre-doctoral Researcher

- ▶ University of Bologna

- 📅 08/2020 - 08/2020
- 📍 VIRTUAL

Program Attendee at The Cornell, Maryland,
Max-Planck Pre-doctoral Research School

- ▶ Max Planck Institute for Software Systems

I was granted the opportunity to be one of ≈ 100 students internationally selected to participate in the Cornell, Maryland, Max Planck Pre-doctoral Research School in 2020, Saarbrücken, Germany (moved online due to COVID-19 pandemic). During this period, I attended lectures conducted by the top scientists and faculty members from participating institutions on various cutting-edge topics, including scalable machine learning and deep learning adversarial attacks. Certificate.

01/2017 - 01/2020
SMART WORKING
ROME (RM), ITALY

Software Developer

▶ CSEN, National Educational Sports Center
Designed, developed, and deployed the software system used in Italy by CSEN for judging taekwondo Poomsae during national competitions. Main use of C# and Kotlin.

03/2017 - 05/2017
CESENA (FC), ITALY

Trainee Student

▶ Smart City Laboratory
Designed, developed, and deployed a Microsoft Azure SQL database for supporting the investigation of rare diseases on the national territory.

06/2013 - 08/2013
02/2013 - 02/2013
06/2012 - 08/2012
RIMINI (RN), ITALY

Trainee Student

▶ Esa Software (now Team System)
Advanced use of .NET, Windows Presentation Foundation, and C#.

// ACHIEVEMENTS, HONOURS AND AWARDS

🏆 Con.Scienza 2020 Award Winner (02/2021)

National Conference of the Presidents and Directors of Science and Technology
National award—with only one nomination per university department—for having written one of the ten best scientific research works during the master's thesis.

🏆 PhD Call First Position (07/2020)

First position in the ranking out of 132 participants for the PhD call in Computer Science and Engineering, University of Bologna.

🏆 Best Paper Award (03/2020)

9th International Conference on Data Science, Technology and Applications (DATA 2020)
The DATA conference series is a reference venue for researches in data/text mining. In 2020, my first co-authored paper "Learning Interpretable and Statistically Significant Knowledge from Unlabeled Corpora of Social Text Messages: A Novel Methodology of Descriptive Text Mining" has been selected as the best contribution among the 70 papers that got past the peer review (14% acceptance rate). The award included an invitation to submit an extended version on Springer Volume. Certificate.

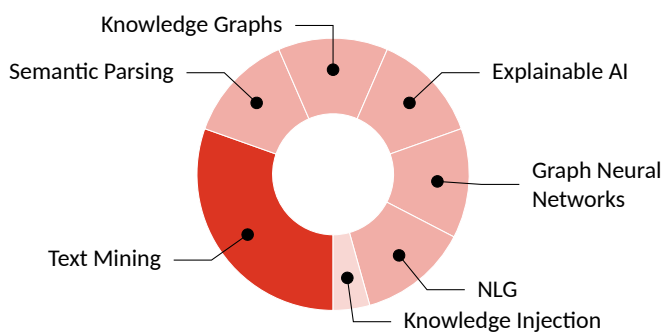
🏆 High School Awards (2014)

Awards as the best student of ITIS Leonardo Da Vinci Scientific High School according to career results:

- winner of the Guido Paolucci Scholarship, BCC of Gradara;
- certificate of Merit and Scholarship, Banca Malatestiana;
- winner of the "Talent Search" project, a training course concerning the programming of microcontrollers, Confindustria Rimini;
- winner of the "ITIS-CNA CAR competition: young inventors", best industrial project, educational software about Cryptography.

// PUBLICATIONS

Author of 7 papers. Citations: 10, h-Index: 3 (Google Scholar metrics as of 2022-03-23). Main keywords and research areas are reported below.



// PRE-SKILLS*

Skill	Proficiency (0-5)
Leadership	5
Team Work	4
Communication	4
Organizing	4
Creativity	4
Motivation	4
Problem Solving	4

// PRO-SKILLS*,†

Programming Languages

Language	Proficiency (0-5)
Python	5
R	5
Scala	5
Java	5
C#	5
Bash	4
Prolog	4

Other Languages

Language	Proficiency (0-5)
LaTeX	5
Markdown	5
HTML, (S)CSS	4

ML Frameworks

Framework	Proficiency (0-5)
PyTorch	4
Jax/Flax	4
TensorFlow	3
Keras	3

Operating Systems

OS	Proficiency (0-5)
Windows	5
Linux	4
MacOS	4

Software & Tools

Tool	Proficiency (0-5)
Visualization (e.g. matplotlib, ggplot, plotly, ...)	4
Data handling/analysis (e.g. numpy, scipy, pandas, scikit-learn, ...)	5
ML monitoring (e.g. W&B, TensorBoard, ...)	5
NLP libraries (e.g. HuggingFace, spaCy, ...)	4
Docker	4

* Average on anonymous scores (from 0 to 5, rounded down) requested from the people I collaborated with (colleagues, teachers, supervisors, and professionals outside the IT context)

† The proficiency skill evaluation scale has the following meaning

Proficiency	Meaning
0 (5 grey dots)	No Experience
1 (1 red, 4 grey dots)	Can read and make small changes to existing programs
2 (2 red, 3 grey dots)	Can utilize basic features without much help
3 (3 red, 2 grey dots)	Can develop medium programs and do nontrivial troubleshooting
4 (4 red, 1 grey dot)	Can develop large programs using all basic and advanced features
5 (5 red dots)	Understanding and (appropriate) usage of most lesser-known features

// CERTIFICATES

🌟 Natural Language Processing Specialization

Coursea, 10/2020 - /

🔗 ID NV29J2BMGADP

🌟 Natural Language Processing with Attention Models

Coursea, 10/2020 - /

🔗 ID VMQSUQBQJPM

🌟 Natural Language Processing with Classification and Vector Spaces

Coursea, 08/2020 - /

🔗 ID QK8EQ2GF87YN

🌟 Natural Language Processing with Probabilistic Models

Coursea, 08/2020 - /

🔗 ID QPALRXRLYWW5

🌟 Natural Language Processing with Sequence Models

Coursea, 08/2020 - /

🔗 ID 276YSYDTCSLH

// REFEREES

List of references available to contact:

Prof. Gianluca Moro

✉ gianluca.moro@unibo.it

👤 M.S. and Ph.D. Thesis Supervisor

Prof. Antonella Carbonaro

✉ antonella.carbonaro@unibo.it

👤 M.S. Thesis Co-Supervisor and Ph.D. Tutor

// VOLUNTERRING

🏠 AMAE Onlus - National Association for Esophageal Achalasia, 2019 - today

In September 2016 I discovered I have a rare disease called "Esophageal Achalasia". Since then I have dedicated myself to merging my NLP skills and patient-centered experiences for creating tools in medical and biomedical domains finalized to research advancement.

- Member of the board of directors
- Representative of the IT sector
- Data Scientist

// MOTTOS

“While technology is important, it's what we do with it that truly matters.

- Muhammad Yunus, Nobel Peace Prize Winner

“Language is at the heart of human intelligence. It therefore is and must be at the heart of our efforts to build artificial intelligence. No sophisticated AI can exist without mastery of language.

- Rob Towes, Forbes

Contributions in Conference Proceedings Sorted By Time

Learning Interpretable and Statistically Significant Knowledge from Unlabeled Corpora of Social Text Messages: A Novel Methodology of Descriptive Text Mining [doi](#)

👤 G. Frisoni, G. Moro, A. Carbonaro

📅 2020 📖 Proceedings of the 9th International Conference on Data Science, Technology and Applications, DATA 2020, Lieusaint, Paris, France, July 7-9, 2020, pp. 121-134, SciTePress (14% acceptance rate)

Unsupervised Descriptive Text Mining for Knowledge Graph Learning [doi](#)

👤 G. Frisoni, G. Moro, A. Carbonaro

📅 2020 📖 Proceedings of the 12th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, IC3K 2020, Volume 1: KDIR, Budapest, Hungary, November 2-4, 2020, vol. 1, pp. 316-324, SciTePress (21% acceptance rate)

Phenomena explanation from text: Unsupervised learning of interpretable and statistically significant knowledge [doi](#)

👤 G. Frisoni, G. Moro

📅 2020 📖 DATA (Revised Selected Papers), Communications in Computer and Information Science, vol. 1446, pp. 293-318, Springer

Journal Publications Sorted By Time

A Survey on Event Extraction for Natural Language Understanding: Riding the Biomedical Literature Wave [doi](#)

👤 G. Frisoni, G. Moro, A. Carbonaro

📅 2021 📖 IEEE Access, vol. 9, pp. 160721-160757

Unsupervised event graph representation and similarity learning on biomedical literature [doi](#)

👤 G. Frisoni, G. Moro, G. Carlassare, A. Carbonaro

📅 2021 📖 Sensors, vol. 22 (1)

Human Being Detection from UWB NLOS Signals: Accuracy and Generality of Advanced Machine Learning Models [doi](#)

👤 G. Moro, F. Di Luca, D. Dardari, G. Frisoni

📅 2021 📖 Sensors, vol. 22 (4)

Contributions in Forums Sorted By Time

Towards rare disease knowledge graph learning from social posts of patients [doi](#)

👤 G. Frisoni, G. Moro, A. Carbonaro

📅 2020 📖 Research and Innovation Forum 2020 - Disruptive Technologies in Times of Change, RIIFORUM 2020, Athens, Greece, 15-17 April 2020, pp. 577-589, Springer

Planned

Text-to-Text Extraction and Verbalization of Biomedical Event Graphs

👤 G. Frisoni, G. Moro, L. Balzani

📖 Submitted to NAACL 2022 (under review; passed the first selection)

Bio-QA-GNN: Interpretable Biomedical Question Answering Combining Language Models and Knowledge Graphs

👤 G. Frisoni, G. Moro, A. Carbonaro

📖 Next to be submitted to Oxford Bioinformatics

Cogito Ergo Summ: Event-Augmented Abstractive Summarization of Biomedical Papers

👤 G. Frisoni, G. Moro, P. Italiani

📖 Next to be submitted to EMNLP 2022

// SCIENTIFIC ACTIVITIES

Participation in Research Groups

DISI UniBo NLP Research Group, 03/2020 – today,

The primary research group I worked with since my M.S. degree. The DISI UniBo NLP group—led by prof. Gianluca Moro—includes a team of Ph.D. students and researchers who are part of the Department of Computer Science and Engineering (DISI) of the University of Bologna, Italy. We pursue a vision focused on proposing original solutions for crucial NLP/NLU tasks, following innovative trends like XAI, memory-enhanced neural networks, graph neural networks, deep metric learning, cross-modal AI, and structured knowledge↔language model integration. Our papers have been accepted to top journals and conferences, including AAAI and ACL. We also have state-of-the-art hardware resources (e.g., +6 NVIDIA GeForce RTX 3090 Turbo 24GB) and powerful servers to support our projects.

Selected papers from the group

- G. Moro and L. Valgimigli, "Efficient self-supervised metric information retrieval: A bibliography based method applied to COVID literature," *Sensors*, vol. 21, no. 19, 2021.
- G. Moro and L. Ragazzi, "Semantic Self-Segmentation for Abstractive Summarization of Long Documents in Low-Resource Regimes," In Proceedings of the 36th AAAI Conference on Artificial Intelligence, Vancouver, BC, Canada, 22 February–1 March 2022, pp. 1–9. AAAI Press, 2022.
- G. Moro, L. Ragazzi, L. Valgimigli, and Davide Freddi, "Discriminative Marginalized Probabilistic Neural Method for Multi-Document Summarization of Medical Literature," In Proceeding of the 60th Annual Meeting of the Association for Computational Linguistics. 2022.
- G. Moro, L. Ragazzi, and L. Valgimigli, "Large-sized Multi-document Summarization of Biomedical Studies with End-to-end Selective Marginalization Learning," Passed the first selection of IJCAI 2022.
- G. Moro and S. Salvatori, "Cross-Modal Retrieval in Fashion Domain Decoupling Caption and Image Embeddings," Next to be submitted to Neurocomputing.

Research Projects

Social media analysis centered on rare bone disease patients

Project selected for financing by the Department of Rare Skeletal Disorders, Rizzoli Orthopaedic Institute, Bologna.

// TEACHING ACTIVITIES

Seminars

Introduction to Azure Cosmos DB

"web Services and Applications" M.S. course, Computer Science and Engineering, University of Bologna, March, 2018

Knowledge Graph Learning from Text

"Semantic Web" M.S. course, Computer Science and Engineering, University of Bologna, March 26, 2020

POIROT: Phenomena Explanation from Text

"Text Mining" M.S. course, Computer Science and Engineering, University of Bologna, December 2, 2020

A look at Knowledge Graphs, Ontologies and Semantic Similarity

"Semantic Web" M.S. course, Computer Science and Engineering, University of Bologna, March 22, 2021

A gentle introduction to Natural Language Understanding from Text: from Phenomena Explanation to Event Extraction and Event Graph Embedding

"Data Mining, Text Mining and Big Data Analytics" M.S. course, Artificial Intelligence, University of Bologna, December 10, 2021

Co-supervision of Bachelor Students

Extraction of medical correlations from unlabeled social posts with neural language models and data clustering

Candidate: Alessandro Lombardini, Supervisor: Gianluca Moro. October, 2020

Automatic translation of social documents shared by rare patients

Candidate: Anna Fabris, Supervisor: Antonella Carbonaro. October, 2020

Semantic similarity and clustering of concepts from the medical literature represented with language models and event-based knowledge graphs

Candidate: Giulio Carlassare, Supervisor: Antonella Carbonaro, Other Co-supervisor: Gianluca Moro. March, 2021

Time-evolving knowledge graphs based on Poirot: dynamic representation of patients' voices

Candidate: Samuele Ceroni, Supervisor: Antonella Carbonaro. March, 2021

Verbalization of biomedical events expressed in the scientific literature: controlled generation of natural language from semantic graphs by means of a text-to-text transformer

Candidate: Lorenzo Balzani, Supervisor: Gianluca Moro. October, 2021

Study and experimentation of advanced metrics for the evaluation of natural language generation models

Candidate: Marco Avagnano, Supervisor: Antonella Carbonaro. December, 2021

Unsupervised representation and similarity learning for event graphs mentioned in the biomedical literature

Candidate: Eleonora Bertoni, Supervisor: Gianluca Moro. December, 2021

Co-supervision of Master Students

🎓 Study and implementation of the graphic interface for a health data application

Candidate: Matteo Sertori, Supervisor: Antonella Carbonaro, Other Co-supervisor: Gianluca Moro. October, 2020

🎓 Generate explanations of medical concept sets made up of correlated terms extracted from patient social posts with linear transformers

Candidate: Alessia Ventani, Supervisor: Gianluca Moro. March, 2021

// SELECTED EXTRA-RESEARCH PROJECTS

📦 Chess Multiplayer

Chess application with multiplayer features | Scala, Prolog, Akka, MongoDB. Use of Scrum as Agile software development methodology (with Product Owner role inside the team). Adoption of advanced software quality techniques and Continuous Integration.

📅 08/2018 - 10/2018 , 🧑‍🤝‍🧑 , x4 contributors

📦 Drowsiness Detection System

Driver drowsiness detection system with a behavioral measure based on eyes closure | Raspberry Pi, Python, OpenGL, Computer Vision models.

📅 02/2019 - 04/2019 , x2 contributors

📦 Big Data and NLP models for Esophageal Achalasia Social Media Analysis

Topic modeling and phenomena explanation for food or treatment low opinion score on top of $\approx 1.500.000$ unlabeled posts shared by patients and caregivers in social media communities | Named Entity Recognition, Named Entity Linking, Sentiment Analysis | HDFS, MapReduce, Spark, SparkSQL YARN, Hive.

📅 08/2019 - 10/2019

March 25, 2022

