Federico Ruggeri

Statement of Truth

I hereby certify that all content reported in this document is true and digitally signed by the undersigned.

Personal Data

Citizenship:	Italian
Affiliation:	DISI, University of Bologna, Italy
Address:	Viale del Risorgimento, 2, 40136, Bologna (BO), Italy
Institutional email:	federico.ruggeri6@unibo.it
ORCID:	0000-0002-1697-8586
Institutional website:	https://www.unibo.it/sitoweb/federico.ruggeri6/en
Research group website:	https://site.unibo.it/nlp/en
Research group GitHub:	https://github.com/nlp-unibo
Personal GitHub	https://github.com/federicoruggeri

Short Bio

My main research area is integrating unstructured knowledge into deep learning models. I'm currently focusing on Natural Language Processing (NLP) and Argument Mining, a branch of NLP that aims at extracting arguments from unstructured texts.

I obtained my PhD from the University of Bologna in 2022 with the thesis *"Towards Integrating Unstructured Knowledge in Natural Language Processing"*. During my PhD I've defined the notion of Unstructured Knowledge Integration (UKI) and investigated its applications in the fields of Legal Analytics and Argument Mining.

I hold the position of Post-doc Research Fellow at the Computer Science and Engineering Department (DISI) of the University of Bologna and I'm mainly researching knowledge extraction and Neuro-symbolic solutions for UKI. I'm currently supervising an internal project concerning the development of interpretable solutions for legal analytics. I have participated in several national and international projects focusing on the definition of interpretable, efficient NLP systems, and the integration of LLMs in industrial use cases.

Highlights

Education	PhD in Computer Science & Engineering 4-year Post-doc Research fellow
Reference	Section Academic Experience Section(s) Education
Teaching	292 cumulative hours as TA in two master degree courses 2-times Lecturer for PhD program (10 hours, 16 hours).
Reference(s)	Section Teaching
Projects	Involved in **five international** projects and **five** national projects.
Reference(s)	Section International Projects Section National Projects
Visiting	Open University (England)

	UKP Lab (Germany)
	Alan Turing Institute (England)
Reference(s)	Section International Experience
Publications	Author or co-author of 35 publications
Scopus	216 citations; h-index 8
Google Scholar	443 citations; h-index 12
Journals	4 Scimago Q1 International Journals
Conferences	4 GGS A++; 3 GGS A+; 4 GGS A; 1 GGS A-; 2 GGS B-; 2 GGS C
Workshops	9 publications
Reference(s)	Section Bibliometric Indexes
	Section Selected Publications for Analytical Research Evaluation
	Section Other Publications
Software Products	*cinnamon* Python library
Reference(s)	Section Software Products

Academic Experience

2024 - Current	**Post-doc Research Fellow** at University of Bologna
Project	*BI-REX GeMEB Project*
Supervisor(s)	*Prof. Paolo Torroni*
Description	Developing LLM-based solutions for industrial use cases that speed up existing user assistance systems while guaranteeing privacy.
2022-2024	**Post-doc Research Fellow** at University of Bologna
Project	*AI4EU Project*
Supervisor(s)	*Prof. Paolo Torroni, Prof. Michela Milano*
Description	Developed horizontal matchmaking solutions for ranking textual sources. These solutions should be compliant with AI-on-demand online services. Critical requirements concern textual sources maintenance.
2021-2022	**Post-doc Research Fellow** at University of Bologna
Project	*StairwAI ICT-49 Project*
Supervisor(s)	*Prof. Paolo Torroni, Prof. Michela Milano*
Description	My research fellowship is centered on developing NLP solutions for large-scale heterogeneous document matching. In particular, I've researched strategies for combining unstructured knowledge (e.g., textual documents, researcher profiles, user requests) and ontology knowledge graphs.
2021	**External Research Collaborator** with University of Modena and Reggio Emilia
Project	*FISR-COVID 2020 AMICA Project*
Supervisor(s)	*Prof. Marco Lippi*
Description	I've developed a ranking system for scientific literature based on Argument Mining features.
2018-2021	**PhD student** at University of Bologna
Curriculum	*Computer science and Engineering*
Supervisor(s)	*Prof. Paolo Torroni, Prof. Marco Lippi*

Description	I've defined the concept of unstructured Knowledge Integration (UKI) in NLP: the task of integrating textual knowledge into deep learning models. I developed solutions for application fields like Legal Analytics and Argument Mining. UKI requires advanced language understanding techniques and is subject to tight requirements like model transparency and interpretability. To address these issues, I've researched scalable and efficient data-driven neuro-symbolic solutions. I took inspiration from several research fields, including Reinforcement Learning, Active Learning, and Informed Machine Learning.

Apr-Oct 2018	**Research Scholarship** in collaboration with Magneti Marelli
Project	*Anomaly detection in the automotive research field*
Supervisor(s)	*Prof. Paolo Torroni, Prof. Marco Lippi*
Description	I developed anomaly detection solutions based on forecasting neural networks and encoder-decoder architectures.

Teaching

2025-2026	
Role	Teaching Assistant
Course	91258 - *Natural Language Processing*
Supervisor	Prof. Paolo Torroni
Info	6 CFUs, 48 hours
Curriculum	Master degree in Artificial Intelligence, University of Bologna, Italy
Role	Teaching Assistant
Course	91251 - *Introduction to Languages for Artificial Intelligence Module 1*
Supervisor	Prof. Michael Lodi
Info	6 CFUs, 24 hours
Curriculum	Master degree in Artificial Intelligence, University of Bologna, Italy
2024-2025	
Role	Teaching Assistant
Course	91258 - *Natural Language Processing*
Supervisor	Prof. Paolo Torroni
Info	6 CFUs, 48 hours
Curriculum	Master degree in Artificial Intelligence, University of Bologna, Italy
Role	Teaching Assistant
Course	91251 - *Introduction to Languages for Artificial Intelligence Module 1*
Supervisor	Prof. Michael Lodi
Info	6 CFUs, 24 hours
Curriculum	Master degree in Artificial Intelligence, University of Bologna, Italy
Role	PhD Course Lecturer
Course	*Robust and Reproducible Research*
Info	3 CFUs, 16 hours (frontal teaching)
Student Rating	93.00%
2023-2024	
Role	Teaching Assistant
Course	91258 - *Natural Language Processing*
Supervisor	Prof. Paolo Torroni
Info	6 CFUs, 40 hours

Curriculum	Master degree in Artificial Intelligence, University of Bologna, Italy
Role	Teaching Assistant
Course	91251 - *Languages and Algorithms for Artificial Intelligence Module 1*
Supervisor	Prof. Simone Martini
Info	12 CFUs, 20 hours
Curriculum	Master degree in Artificial Intelligence, University of Bologna, Italy
Role	PhD Course Lecturer
Course	*Robust and Reproducible Experimental Deep Learning Setting*
Info	2 CFUs, 10 hours (frontal teaching)
Student Rating	87.60%
2022-2023	
Role	Teaching Assistant
Course	91258 - *Natural Language Processing*
Supervisor	Prof. Paolo Torroni
Info	6 CFUs, 40 hours
Curriculum	Master degree in Artificial Intelligence, University of Bologna, Italy
Role	Teaching Assistant
Course	91251 - *Languages and Algorithms for Artificial Intelligence Module 1*
Supervisor	Prof. Simone Martini
Info	12 CFUs, 20 hours
Curriculum	Master degree in Artificial Intelligence, University of Bologna, Italy
2021-2022	
Role	Teaching Assistant
Course	91258 - *Natural Language Processing*
Supervisor	Prof. Paolo Torroni
Info	6 CFUs, 40 hours
Curriculum	Master degree in Artificial Intelligence, University of Bologna, Italy
Role	Teaching Assistant
Course	91251 - *Languages and Algorithms for Artificial Intelligence Module 1*
Supervisor	Prof. Simone Martini
Info	12 CFUs, 20 hours
Curriculum	Master degree in Artificial Intelligence, University of Bologna, Italy
2020-2021	
Role	Teaching Assistant
Course	91258 - *Natural Language Processing*
Supervisor	Prof. Paolo Torroni
Info	6 CFUs, 40 hours
Curriculum	Master degree in Artificial Intelligence, University of Bologna, Italy

Academic Supervisor

Available at: https://amslaurea.unibo.it/view/relatore/Ruggeri=3AFederico=3A=3A/.

Master Student(s)	
P.I. Clotan	Selective rationalization via genetic-based learning: extending the Gen-SPP architecture with the CMA-ES algorithm [LM-DM270]
2024	
Master Student(s)	
E. Muneer	An Artificial Intelligence System for the Autonomous Pagination of Newspapers [LM-DM270]
G. Tanzi	Automating Test Case Generation for Automotive Industry using Large Language Models [LM-DM270]
Y. Noviello	Enhancing Chatbot Efficacy in Italian Language through Retrieval Augmented Generation and LoRA Fine-Tuning [LM-DM270]
S. Fantazzini	Efficient Knowledge Distillation for Green NLP Models: Bridging the Gap with Large Language Models [LM-DM270]
S. Di Giacomo	Generazione e valutazione di test automatici tramite modelli di NLP avanzati [LM-DM270]
2023	
Master Student(s)	
G. Pappacoda	From text to knowledge: Large Language Models-based methods for knowledge extraction [LM-DM270]
R. Saleem Butt	Design and Implementation of a Neural Machine Translation Engine for Computer-Assisted Translations [LM-DM270]
A. Proia	Neural Clustering on Tree Structured Data: A case study on Argument Mining [LM-DM270]
F. Pieroni	Developing an Automated ESG Data Extraction and Analysis Tool with NLP Techniques and Large Language Models [LM-DM270]
D. Koluh	Graph-Based Keyword Extraction from Scientific Paper Abstracts using Word Embeddings [LM-DM270]
J. M. Hartsuiker	Finetuning commercial Large Language Models with LoRA for enhanced Italian language understanding [LM-DM270]
2022	
Master Student(s)	
G. Murro	SynBA: A contextualized Synonim-Based adversarial Attack for text classification [LM-DM270]
A. Fabris	Tree regularized neural networks for key point analysis [LM-DM270]
A. Falai	Conditioning Text-to-Speech synthesis on dialect accent: a case study [LM-DM270]
H. Zhang	Quality and Aspect based Argument Generation [LM-DM270]
2021	
Master Student(s)	
E. Mancini	Disruptive Situations Detection on Public Transports through Speech Emotion Recognition [LM-DM270]
M. Rondina	Realizzazione di un sistema per l'estrazione di articoli scientifici in base a query utente in linguaggio naturale [LM-DM270]

Academic Service

2025	
Role	Committee member for research position call

Description	Avviso pubblico di selezione per titoli e colloquio per l'affidamento di un incarico libero-professionale (machine translation) con titolare di P.Iva. per le esigenze del dipartimento di interpretazione e traduzione - DIT
Affiliation	University of Bologna, Dipartimento di Interpretazione e Traduzione (DIT)
Source	https://bandi.unibo.it/s/dit/avviso-di-selezione-per-incarico-libero-professionale-bando-machine-translation-dipartimento-di-interpretazione-e-traduzione-dit

Education

2022

Type	**PhD**
Title	*Towards Unstructured Knowledge Integration in Natural Language Processing*
Curriculum	*Computer Science and Engineering*
Supervisor(s)	*Prof. Paolo Torroni, Prof. Marco Lippi*
Evaluation	*5/5, Excellence*
Description	In this thesis, I analyze the methodology of integrating knowledge into deep learning models in the field of Natural Language Processing (NLP). I introduce Unstructured Knowledge Integration (UKI) as the process of integrating unstructured knowledge into machine learning models. I discuss UKI in the field of NLP, where knowledge is represented in a natural language format. UKI is a complex process comprised of multiple sub-processes, different knowledge types, and knowledge integration properties to guarantee. I propose a unified vision of structured knowledge extraction (KE) and UKI by identifying KE as a sub-process of UKI. KE is viewed as a form of symbolic representation. From this perspective, I remark on the need to define sophisticated UKI processes to verify the validity of knowledge integration. To this end, I advocate for frameworks capable of combining symbolic and sub-symbolic representations for learning as a solution.

2019

Type	**Summer School**, Certosa di Pontignano, Siena
Theme	*2nd Advanced Course on Data Science & Machine Learning*
Description	The course is a full-immersion five-day residential Course at the Certosa di Pontignano (Siena - Tuscany, Italy) on cutting-edge advances in Data Science and Machine Learning with lectures delivered by world-renowned experts.
Info	8 CFUs, 7 days.

Type	**Summer School**, University of Lille, Lille
Theme	*Reinforcement Learning Summer SCOOL (RLSS)*
Description	Two full weeks to discover the theory and practice of sequential decision making: reinforcement learning introduction, multi-armed bandits and deep RL.
Info	8 CFUs, 15 days.

Type	**Summer School**, University of Genova, Genova
Theme	*Machine Learning Crash Course (MLCC)*
Description	The course provides an introduction to the fundamental methods at the core of modern Machine Learning. It covers theoretical foundations as well as essential algorithms. Classes on theoretical and algorithmic aspects are complemented by practical lab sessions.
Info	2 CFUs, 5 days.

2018

Type	**Summer School**, University of Ferrara, Ferrara
Theme	*School is Statistical Relational Artificial Intelligence (StarAI)*
Description	Relational AI deals very effectively with complex domains involving many and even a varying number of entities connected by complex relationships, while statistical AI manages well the uncertainty that derives from incomplete and noisy descriptions of the domains.

Type	**Master Degree in Computer Engineering**
	110/110 *summa cum laude*, University of Bologna, Bologna
Title	*Argument structure prediction with stance classification features*
Supervisor(s)	*Prof. Paolo Torroni, Prof. Marco Lippi*

2015

Type	**Bachelor Degree in Computer Engineering**
	University of Bologna, Bologna
Title	*Monitoring, alerting and resource management of cloud applications based on Docker*
Supervisor(s)	*Prof. Paolo Bellavista*

International Projects

2025-Current

ID	JUST-2022-EJUSTICE
Title	*Principles Of Law In National and European VAT (POLINE)*
Reference	https://site.unibo.it/poline/en
Description	POLINE aims at developing an AI-powered pilot tool for the retrieval and analysis of judicial principles of law in the CJEU and national case-law in Value Added Tax (VAT). The tool relies on AI techniques for extracting, clustering and linking judicial principles of law and is embedded in a modular platform, consisting of a Legal Database, Link Visualization and Customised Detection Module. It covers the case-law of the CJEU and the Italian, Swedish and Bulgarian Supreme Courts and will be accessible to judges, other legal practitioners, tax policymakers and taxpayers. The development of the tool will be based on a multidisciplinary approach combining theory and practice of judicial decision-making for the study of the concept of "judicial principle of law" and the analysis of the case-law; legal informatics methods for the creation of an ontology of judicial concepts in VAT and training datasets of annotated judicial principles of law; AI, machine learning, and NLP techniques for the automatic extraction of principles, the detection of textual and semantic similarity, and network analysis. The tool will be trialled in 3 online national testing events and disseminated in 3 national demonstration events and 1 final international conference. The pilot tool provides a robust and trustworthy use case of AI technologies for justice. It will provide non-discriminatory and effective access to justice. Through its collection of principles of law and NLP-powered search engine, the tool will assist judges in accessing legal knowledge reducing their work overload. Moreover, through the Customised Detection Test Module, the tool will allow recipients of VAT measures to identify judicial principles of law applied in a specific case and assess whether VAT law is correctly applied. By developing open-access automated techniques of knowledge extraction, the methods developed can be easily reused and expanded to include other fields of law and other legal systems.
Role	Develop a transparent AI system that (i) processes legal documents (e.g., tax audit acts) to extract judicial principles of law; (ii) integrates domain-specific contextual knowledge, such as EU legal principles traceable back to the processed documents; (iii) produce a human-comprehensible explanation on the extraction process that is grounded on the integrated knowledge. These capabilities allow legal experts to explicitly assess the efficiency and inference process of the developed system.

2024-Current

ID	FAIR
Title	*Future Artificial Intelligence Research*
Reference	https://fondazione-fair.it/
Description	The objective of the FAIR project is to contribute facing the research questions, methodologies, models, technologies, and ethical and legal rules to build AI systems capable of interacting and collaborating with humans, perceiving and acting in evolving contexts, to be conscious about their limits and capable to adapt to new situations, to be aware of the perimeters of safety and trust, and to be careful with the environmental and social impact that their creation and functioning may cause.
Role	Actively participating in the TP2 - Vision, Language and Multimodal Challenges. The focus is to develop challenging evaluating benchmarks for LLMs concerning reasoning capabilities.
2023-2024	

ID	H2020-ICT-2018-825619-AI4EU
Title	*A European AI On Demand Platform and Ecosystem*
Reference	https://cordis.europa.eu/project/id/825619
Description	The EU-funded AI4EU is working to change Europe's place in this race, by building the first European AI On-Demand Platform and Ecosystem that will share resources, tools, knowledge, algorithms and more between Member States. It will help to increase innovation and technology transfer, accelerate the growth of start-ups and SMEs, and fulfill the needs of the European AI community. The project will implement eight pilots led by industrial partners to demonstrate the platform's capabilities.
Role	Developing horizontal matchmaking solutions for AI resources. These solutions should be compliant with AI-on-demand online services.
2022 - 2023	

ID	H2020-ICT-2020-101017142-StairwAI
Title	*Stairway to AI: Ease the Engagement of Low-Tech users to the AI-on-Demand platform through AI*
Reference	https://cordis.europa.eu/project/id/101017142
Description	The StairwAI project targets low-tech users with the goal of facilitating their engagement on the AI4EU on-demand Platform. This will be achieved through a new service layer enriching the functionalities of the on-demand platform and containing: (1) a multi-lingual interaction layer enabling conversations with the Platform in the user's own language, (2) a horizontal matchmaking service for the automatic discovery of AI assets (tools, data sets, AI experts, consultants, papers, courses etc.) meeting the user business needs and, (3) a vertical matchmaking service that will dimension and provision hardware resources through a proper hardware provider (HPC, Cloud and Edge infrastructures).
Role	Developed several NLP solutions for AI assets, such as scientific papers and experts' profile, horizontal matchmaking (work package 5). The developed solutions are compliant with the project's ontology regarding AI terminology. The ontology was used to map input textual queries to AI assets to eventually perform the matchmaking. Moreover, I also participated in the work package 7 "Open Call Management" as a technical mentor. In particular, I collaborated with several companies in supervising their prototype AI projects (6-8 months development time).
2021-2023	

ID	JUST-JACC-EJU-AG-2020-1010074206-Adele
Title	*Analytics for DEcision of LEgal cases*
Reference	https://site.unibo.it/adele/en
Description	Project ADELE is premised on the ongoing paradigm shift towards cognitive computing and human-centered AI which is transforming many socio-economic activities, including justice. The project applies legal analytics (LA) – a blend of data science, machine learning, and natural language processing techniques – to judicial decisions. It aims to develop methods to extract knowledge and engage in outcome predictions and there build a pilot tool to support legal research and decision-making processes in the judiciary.

Role	I worked on two main problems: legal outcome prediction and argument mining. In particular, I collaborated in building new corpora and AI solutions, and I actively supervised interns and research fellows on the project.

2019-2022

ID	CLAUDETTE
Title	*automated CLAUse DETectEr*
Reference	http://claudette.eui.eu/
Description	CLAUDETTE is an interdisciplinary research project hosted at the Law Department of the European University Institute. The research objective is to test to what extent is it possible to automate reading and legal assessment of online consumer contracts and privacy policies, to evaluate their compliance with EU's unfair contractual terms law and personal data protection law (GDPR), using machine learning and grammar-based approaches.
Role	I'm involved in the development of NLP solutions for legal text classification. I focused on integrating legal expertise as unstructured textual knowledge into deep learning model to develop interpretable AI solutions.

National Projects

2024 - 2025

ID	BI-REX
Title	*Generative Models: Empowering Business Processes and Enhancing Workflows for Improved Performance (GeMEB)*
Reference	https://bi-rex.it/quarto-bando-bi-rex/
Description	The project developing ad-hoc LLM-based solutions to speed up existing user assistance systems while guaranteeing privacy.
Role	Principal technical investigator to develop LLMs-based solution for unstructured knowledge integration.
	AI-based Smart Collaborative Manufacturing System (SmartCasm)
Reference	https://bi-rex.it/quarto-bando-bi-rex/
Description	The project involves using LLMs to integrate unstructured knowledge into industrial pipelines to speed up production and foster technical advancement.
Role	Principal technical investigator to develop LLMs-based solution for unstructured knowledge integration.

2022

ID	PRIN2017NCPZ22 LAILA
Title	*Legal Analytics for Italian Law*
Reference	https://dsg.unibo.it/it/ricerca/progetti-di-ricerca/progetti-nazionali-e-di-ateneo/prin2017-laila-legal-analytics-for-italian-law
Description	The project regards the application of Legal Analytics methods to a vast and heterogeneous set of legal information: legislations, contracts, and judgments. The purpose is the application of Artificial Intelligence, Machine Learning, and Natural Language Processing to extract legal knowledge, infer relationships, and produce data-driven forecasts.
Role	I actively collaborated in developing NLP solutions for legal analytics, supervised research interns and coordinated with the team of legal experts.

2021

ID	FISR2020IP_01362 AMICA
Title	*Argument Mining In Covid-19 Articles*
Reference	http://amica.unimore.it/
Description	The objective of the AMICA project was to exploit the argumentative content present in the scientific literature regarding Covid-19 to improve the retrieval of relevant and reliable articles. The project involved both medical and artificial intelligence experts and aimed to develop an argument mining-based search engine, specifically designed for the analysis of scientific literature related to Covid-19.
Role	I contributed in developing NLP solutions and supervising a post-graduate student.

International Experience

Nov-Dec 2024	Visiting Researcher at Alan Turing Institute
Location	London, England
Supervisor(s)	*Prof. Maria Liakata*
Description	The visiting period focused on finalizing two research papers on rumor verification and large language model reasoning.

Feb-July 2021	Research internship at Ubiquitous Knowledge Processing (UKP) Lab
Location	Darmstadt, Germany
Supervisor(s)	*Prof. Mohsen Mesgar, Prof. Iryna Gurevych*
Description	My internship project concerned the development of AI dialogue argumentative chatbots on scientific literature. I developed a custom online platform for human dialogue collection. The collected data was used to investigate the impact of arguments in scientific papers to boost the capabilities of a dialogue agent.

April 2019	Research Short-Term Scientific Mission (STSM)
Location	Open University, Milton Keynes, England
Supervisor(s)	*Prof. Anna de Liddo*
Description	The aim of the scientific mission is to explore the possibility of integration of three technologies under development in the domain of computer-supported reflection and deliberation systems from three APPLY COST network institutions: the Open University (host institution), the University of Bologna (applicant's institution), and the Universidade Nova in Lisbon (Institution coordinating the APPLY COST action). The desired outcome of this Short-Term Scientific Missions (STSM) is a draft workplan leading to a focused pilot initiative in the context of the APPLY network, aiming to demonstrate the potential of innovative technologies in facilitating the policy debate through computer-aided reflection and deliberation systems.

Program Chair & Organization Committees

2026	
Title	13rd Workshop on Argument Mining (ArgMining)
Venue	Co-located with ACL 2026, Association for Computational Linguistics
Role	*Co-organizer*
Location	San Diego, USA
2025	
Title	CheckThat! Lab 2025, International Workshop on Fact-Checking
Venue	Co-located with CLEF 2025, Conference and Labs of the Evaluation Forum
Role	*Co-chair*
Location	Madrid, Spain

Title	MM-ArgFallacy-2025 Shared Task
Venue	ArgMining 2025, co-located with ACL 2025
Role	*Shared Task Organizer*
Source	https://nlp-unibo.github.io/mm-argfallacy/2025/
Location	Vienna, Austria
Sept 2024	
Title	CheckThat! Lab 2024, International Workshop on Fact-Checking
Venue	Co-located with CLEF 2024, Conference and Labs of the Evaluation Forum
Role	*Co-chair*
Location	Grenoble, France
Sept 2023	
Title	CheckThat! Lab 2023, International Workshop on Fact-Checking
Venue	Co-located with CLEF 2023, Conference and Labs of the Evaluation Forum
Role	*Co-chair*
Location	Thessaloniki, Greece
Sep 2022	
Title	Conference and Labs of the Evaluation Forum (CLEF)
Venue	CLEF 2022
Role	*Local Organizer*
Location	Bologna, Italy

Invited Speaker

2025	
Venue	CIRSFID, Bologna, Italy
Title	"Reasoning", Reasoning and Argumentation. Una breve panoramica critica su LLM
Description	The talk addresses three main aspects of LLMs and reasoning capabilities. First, we discuss what kind of reasoning type LLMs are tested for. The short answer is that in the majority of cases, it is unclear which reason type(s) is (are) considered. Second, we discuss to what extent do LLMs perform reasoning. Some view LLMs as stochastic parrots, while others believe they acquire true reasoning capabilities. Third, we show how reasoning and argumentation are tightly connected and discuss how argumentation is being progressively used as a way to assess reasoning capabilities in LLMs.
Role	Invited Speaker at *Incontro ALMA-AI | Workshop LLM: stato dell'arte fra limiti e prospettive terza edizione*
Venue	University of Milano Bocconi, Italy
Title	The importance of distilling and using textual knowledge in hate speech

Description | In NLP, to acquire new knowledge, we require the development of machine learning systems with the capability of understanding text to distill and organize novel (structured) knowledge, while guaranteeing other desiderata like model transparency, efficiency, robustness, adaptability to different contexts, and computational scalability. Concretely, we require a paradigm shift in how machine learning problems are formulated, going from the traditional approach of "Given input X, provide output Y" to "Given input X and knowledge K, provide output Y and update K". Here, the knowledge K refers to any information that contributes to the general human understanding of a given problem. This paradigm shift implies that when we design an NLP system, our purpose and measurement of its success should account for both model accuracy and the reason behind that accuracy. These capabilities constitute important desiderata for several real-world applications where extracted knowledge represents the basis for understanding underlying factors like human sociodynamics, dialogical interactions, and reasoning. A notable example is Hate Speech, where knowledge about the cultural domain of written text is necessary to discriminate hateful content.

Role | Invited Speaker at MilaNLP group

Venue | Alan Turing Institute, London

Title | From data annotation to modelling, the importance of distilling and using textual knowledge in hate speech.

Description | In NLP, to acquire new knowledge, we require the development of machine learning systems with the capability of understanding text to distill and organize novel (structured) knowledge, while guaranteeing other desiderata like model transparency, efficiency, robustness, adaptability to different contexts, and computational scalability. Concretely, we require a paradigm shift in how machine learning problems are formulated, going from the traditional approach of "Given input X, provide output Y" to "Given input X and knowledge K, provide output Y and update K". Here, the knowledge K refers to any information that contributes to the general human understanding of a given problem. This paradigm shift implies that when we design an NLP system, our purpose and measurement of its success should account for both model accuracy and the reason behind that accuracy. These capabilities constitute important desiderata for several real-world applications where extracted knowledge represents the basis for understanding underlying factors like human sociodynamics, dialogical interactions, and reasoning. A notable example is Hate Speech, where knowledge about the cultural domain of written text is necessary to discriminate hateful content. In this talk, I provide a brief overview of how textual knowledge can be integrated in hate speech as a tool for reasoning and as instructions, covering data annotation, problem formulation, and modelling.

Role | Invited Speaker at *Media in the digital age* special interest group

Reference | https://digitalmediasig.github.io/seminars2024/

2024

Venue | University of Bologna, Italy

Title | Reliability, Transparency and Reproducibility

Description | The meeting fosters discussions on the ingredients of research methodology, and the ongoing debate around reproducibility of results, with the aim of understanding if and how such dimensions can be translated into evaluation criteria to assess the quality of knowledge generated through research

Role | Selected speaker (6 selected out of 45 candidates from 22 departments)

2023

Venue | Alan Turing Institute, London

Title | Combining Transformers with Natural Language Explanations

Venue | Neurosymbolic Workshop (NeSY), Certosa di Pontignano, Siena

Title | Challenges for Neuro-Symbolic Approaches: Case Study on Legal Analytics and Argument Mining

Presentations

2025 |

Title	Interlocking-free Selective Rationalization Through Genetic-based Learning
Venue	ACL 2025
Reference	[S2]
Location	Vienna, Austria
Presentation	Poster
Title	Assessing the reasoning abilities of LLMs in the context of claim verification
Venue	ACL 2025
Reference	[S1]
Location	Vienna, Austria
Presentation	Poster
2024	
Title	A Corpus for Sentence-Level Subjectivity Detection on English News Articles
Venue	LREC-COLING 2024
Reference	[S6]
Location	Turin, Italy
Presentation	Poster
2023	
Title	A dataset of argumentative dialogues on scientific papers.
Venue	ACL 2023
Reference	[S10]
Location	Toronto, Canada.
Presentation	Poster
Title	On the definition of prescriptive annotation guidelines for language-agnostic subjectivity detection.
Venue	Text2Story 2023 Workshop, Co-located with ECIR
Reference	[O10]
Location	Dublin, Ireland (Online).
Presentation	Oral
2019	
Title	Deep learning for detecting and explaining unfairness in consumer contracts.
Venue	JURIX 2019
Reference	[O21]
Location	Madrid, Spain (Online).
Presentation	Poster

Funding

2024	
Title	"TRUSTWORTHY ARTIFICIAL INTELLIGENCE: Short Stay"
Description	The objective is to strengthen the scientific research of both Italian and UK researchers, promote the internationalization and the cooperation between Italy and UK and foster the technology transfers for the national industrial and societal challenges in both countries.
Funding	2.5k euros for a two-week visiting period.

Institution	Future Artificial Intelligence Research Foundation
Source	https://fondazione-fair.it/wp-content/uploads/2024/02/TAI-ShortStay-Exchange-program-2023.pdf

Awards

2024	
Award	Outstanding Reviewer at EMNLP 2024
Description	Been selected as one of the top reviewers based on several factors, such as review quality, availability, and active interactions with peers.
Source	https://aclanthology.org/2024.emnlp-main.0/

Peer Reviewing

We use † to denote class 1 conferences (according to GII-GRIN-SCIE rating) and Q1 journals (according to Scimago Journal Rank).

Associate Editor	
PLOS One	*Public Library of Science Journal* (2024, 2025)
Senior PC Member	
ECAI	*European Conference on Artificial Intelligence* (2024, 2025)
Area Chair	
ARR	*ACL Rolling Review* (2025)
PC Member	
Conferences	
†ACL	*Association for Computational Linguistics* (2022-2025)
†NAACL	*North American Association for Computational Linguistics* (2022-2025)
†EMNLP	*Empirical Methods in Natural Language Processing* (2022-2025)
†AAAI	*Association for the Advancement of Artificial Intelligence* (2022-2025)
†IJCAI	*International Joint Conference on Artificial Intelligence* (2024, 2025)
EACL	*European Association for Computational Linguistics* (2022-2025)
LREC	*International Conference on Language Resources and Evaluation* (2024)
COLING	*International Conference on Computational Linguistics* (2024, 2025)
HHAI	*Hybrid Human Artificial Intelligence* (2025)
ICTAI	*IEEE International Conference on Tools with Artificial Intelligence* (2022)
COMMA	*International Conference on Computational Models of Argument* (2022)
NLDB	*International Conference on Natural Language & Information Systems* (2024-2025)
Workshops	
ArgMining	*Workshop on Argument Mining* (2021-2025)
NLLP	*Natural Legal Language Processing* (2022)
LUHME	*Language Understanding in the Human-Machine Era* (2025)
Reviewer	
†AI & Law	*Artificial Intelligence and Law* (2021-2025)
†ACM	*ACM Computing Surveys* (2022)

†IEEE Access	*The Multidisciplinary Open Access Journal* (2022)
†AIR	*Artificial Intelligence Review* (2025)

Bibliometric Indexes

Last accessed 16 May 2025.

Source	**Scopus**
Citations	216
Citations by	158 documents
h-index	8
Source	**Google Scholar**
Citations	443
h-index	12

Selected Publications for Analytical Research Evaluation

Principal investigator or equal contribution entries are highlighted in **bold**.

We report the GII-GRIN-SCIE (GGS) and the International Computing Research and Education (ICORE) ratings for conferences, and Scimago quartiles for journals as standard practice for evaluating research. Workshops and pre-prints are not evaluated.

The GII-GRIN-SCIE (GGS) Conference Rating is sponsored by GII (Group of Italian Professors of Computer Engineering), GRIN (Group of Italian Professors of Computer Science), and SCIE (Spanish Computer-Science Society). The ICORE Conference Rating is sponsored by CORE (Computing Research and Education Association of Australasia), GII, GRIN, and SCIE. The SCImago Journal Rank (SJR) indicator is a measure of the prestige of scholarly journals that accounts for both the number of citations received by a journal and the prestige of the journals where the citations come from.

[S1] John Dougrez-Lewis, Mahmud Elahi Akhter, Federico Ruggeri, Sebastian Löbbers, Yulan He, and Maria Liakata. Assessing the reasoning capabilities of LLMs in the context of evidence-based claim verification. In Wanxiang Che, Joyce Nabende, Ekaterina Shutova, and Mohammad Taher Pilehvar, editors, *Findings of the Association for Computational Linguistics: ACL 2025*, pages 20604–20628, Vienna, Austria, July 2025. Association for Computational Linguistics. **GGS Rating: A++. ICORE Rating: A***.

[S2] **Ruggeri, Federico** and Gaetano Signorelli. Interlocking-free selective rationalization through genetic-based learning. In Wanxiang Che, Joyce Nabende, Ekaterina Shutova, and Mohammad Taher Pilehvar, editors, *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1175–1191, Vienna, Austria, July 2025. Association for Computational Linguistics. **Principal Investigator. GGS Rating: A++. ICORE Rating: A*.**

[S3] Katerina Korre, Arianna Muti, Federico Ruggeri, and Alberto Barrón-Cedeño. Untangling hate speech definitions: A semantic componential analysis across cultures and domains. *Findings of the Association for Computational Linguistics: NAACL 2025*, pages 3184–3198, April 2025. **GGS Rating: A+. ICORE Rating: A.**

[S4] Eleonora Mancini, Ana Tanevska, Andrea Galassi, Alessio Galatolo, Federico Ruggeri, and Paolo Torroni. Promoting the responsible development of speech datasets for mental health and neurological disorders research. *J. Artif. Intell. Res.*, 82:937–972, 2025. **Scimago Quartile: Q1.**

[S5] Arianna Muti, Federico Ruggeri, Khalid Al Khatib, Alberto Barrón-Cedeño, and Tommaso Caselli. Language is scary when over-analyzed: Unpacking implied misogynistic reasoning with argumentation theory-driven prompts. In Yaser Al-Onaizan, Mohit Bansal, and Yun-Nung Chen, editors, *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 21091–21107, Miami, Florida, USA, November 2024. Association for Computational Linguistics. **GGS Rating: A+. ICORE Rating: A*.**

[S6] Francesco Antici, Federico Ruggeri, Andrea Galassi, Katerina Korre, Arianna Muti, Alessandra Bardi, Alice Fedotova, and Alberto Barrón-Cedeño. A corpus for sentence-level subjectivity detection on English news articles. In Nicoletta Calzolari, Min-Yen Kan, Veronique Hoste, Alessandro Lenci, Sakriani Sakti, and Nianwen Xue, editors, *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 273–285, Torino, Italy, May 2024. ELRA and ICCL. **GGS Rating: A. ICORE Rating: B. Corresponding Author**.

[S7] Arianna Muti, Federico Ruggeri, Cagri Toraman, Alberto Barrón-Cedeño, Samuel Algherini, Lorenzo Musetti, Silvia Ronchi, Gianmarco Saretto, and Caterina Zapparoli. PejorativITy: Disambiguating pejorative epithets to improve misogyny detection in Italian tweets. In Nicoletta Calzolari, Min-Yen Kan, Veronique Hoste, Alessandro Lenci, Sakriani Sakti, and Nianwen Xue, editors, *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 12700–12711, Torino, Italy, May 2024. ELRA and ICCL. **GGS Rating: A. ICORE Rating: B.**

[S8] Eleonora Mancini, Federico Ruggeri, and Paolo Torroni. Multimodal fallacy classification in political debates. In Yvette Graham and Matthew Purver, editors, *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 170–178, St. Julian's, Malta, March 2024. Association for Computational Linguistics. **GGS Rating: A+. ICORE Rating: A.**

[S9] Eleonora Mancini, Andrea Galassi, Federico Ruggeri, and Paolo Torroni. Disruptive situation detection on public transport through speech emotion recognition. *Intelligent Systems with Applications*, 21:200305, 2024. **Scimago Quartile: Q1.**

[S10] **Federico Ruggeri**, Mohsen Mesgar, and Iryna Gurevych. A dataset of argumentative dialogues on scientific papers. In Anna Rogers, Jordan Boyd-Graber, and Naoaki Okazaki, editors, *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 7684–7699, Toronto, Canada, July 2023. Association for Computational Linguistics. **Principal Investigator. GGS Rating: A++. ICORE Rating: A*.**

[S11] Gianfranco Brambilla, Antonella Rosi, Francesco Antici, Andrea Galassi, Daniele Giansanti, Fabio Magurano, Federico Ruggeri, Paolo Torroni, Evaristo Cisbani, and Marco Lippi. Argument mining as rapid screening tool of covid-19 literature quality: Preliminary evidence. *Frontiers in Public Health*, 10, 2022. **Scimago Quartile: Q1.**

[S12] **Federico Ruggeri**, Francesca Lagioia, Marco Lippi, and Paolo Torroni. Detecting and explaining unfairness in consumer contracts through memory networks. *Artif. Intell. Law*, 30(1):59–92, 2022. **Principal Investigator. Scimago Quartile: Q1.**

Other Publications

First author or equivalent contribution entries are highlighted in **bold**.

[O1] Katerina Korre, Federico Ruggeri, and Alberto Barrón-Cedeño. A Grice-ful examination of offensive language: Using NLP methods to assess the co-operative principle. In Rui Sousa-Silva, Henrique Lopes Cardoso, Maarit Koponen, Antonio Pareja Lora, and Márta Seresi, editors, *Proceedings of the First LUHME Workshop*, pages 12–19, Santiago de Compostela, Spain, October 2024. CLUP, Centro de Linguística da Universidade do Porto FLUP - Faculdade de Letras da Universidade do Porto.

[O2] Eleonora Mancini, Federico Ruggeri, Stefano Colamonaco, Andrea Zecca, Samuele Marro, and Paolo Torroni. MAMKit: A comprehensive multimodal argument mining toolkit. In Yamen Ajjour, Roy Bar-Haim, Roxanne El Baff, Zhexiong Liu, and Gabriella Skitalinskaya, editors, *Proceedings of the 11th Workshop on Argument Mining (ArgMining 2024)*, pages 69–82, Bangkok, Thailand, August 2024. Association for Computational Linguistics.

[O3] Alberto Barrón-Cedeño, Firoj Alam, Tanmoy Chakraborty, Tamer Elsayed, Preslav Nakov, Piotr Przyby a, Julia Maria Struß, Fatima Haouari, Maram Hasanain, Federico Ruggeri, Xingyi Song, and Reem Suwaileh. The clef-2024 checkthat! lab: Check-worthiness, subjectivity, persuasion, roles, authorities, and adversarial robustness. In *Advances in Information Retrieval: 46th European Conference on Information Retrieval, ECIR 2024, Glasgow, UK, March 24–28, 2024, Proceedings, Part V*, page 449–458, Berlin, Heidelberg, 2024. Springer-Verlag.

[O4] Julia Maria Struß, Federico Ruggeri, Alberto Barrón-Cedeño, Firoj Alam, Dimitar Dimitrov, Andrea Galassi, Georgi Pachov, Ivan Koychev, Preslav Nakov, Melanie Siegel, Michael Wiegand, Maram Hasanain, Reem Suwaileh, and Wajdi Zaghouani. Overview of the CLEF-2024 checkthat! lab task 2 on subjectivity in news articles. In Guglielmo Faggioli, Nicola Ferro, Petra Galuscáková, and Alba García Seco de Herrera, editors, *Working Notes of the Conference and Labs of the Evaluation Forum (CLEF 2024), Grenoble, France, 9-12 September, 2024*, volume 3740 of *CEUR Workshop Proceedings*, pages 287–298. CEUR-WS.org, 2024.

[O5] Alberto Barrón-Cedeño, Firoj Alam, Tommaso Caselli, Giovanni Da San Martino, Tamer Elsayed, Andrea Galassi, Fatima Haouari, Federico Ruggeri, Julia Maria Struß, Rabindra Nath Nandi, Gullal S. Cheema, Dilshod Azizov, and Preslav Nakov. The clef-2023 checkthat! lab: Checkworthiness, subjectivity, political bias, factuality, and authority. In Jaap Kamps, Lorraine Goeuriot, Fabio Crestani, Maria Maistro, Hideo Joho, Brian Davis, Cathal Gurrin, Udo Kruschwitz, and Annalina Caputo, editors, *Advances in Information Retrieval*, pages 506–517, Cham, 2023. Springer Nature Switzerland. **GGS Rating: A-. ICORE Rating: A.**

[O6] Alberto Barrón-Cedeño, Firoj Alam, Andrea Galassi, Giovanni Da San Martino, Preslav Nakov, Tamer Elsayed, Dilshod Azizov, Tommaso Caselli, Gullal S. Cheema, Fatima Haouari, Maram Hasanain, Mücahid Kutlu, Chengkai Li, Federico Ruggeri, Julia Maria Struß, and Wajdi Zaghouani. Overview of the CLEF-2023 checkthat! lab on checkworthiness, subjectivity, political bias, factuality, and authority of news articles and their source. In Avi Arampatzis, Evangelos Kanoulas, Theodora Tsikrika, Stefanos Vrochidis, Anastasia Giachanou, Dan Li, Mohammad Aliannejadi, Michalis Vlachos, Guglielmo Faggioli, and Nicola Ferro, editors, *Experimental IR Meets Multilinguality, Multimodality, and Interaction - 14th International Conference of the CLEF Association, CLEF 2023, Thessaloniki, Greece, September 18-21, 2023, Proceedings*, volume 14163 of *Lecture Notes in Computer Science*, pages 251–275. Springer, 2023.

[O7] Andrea Galassi, Federico Ruggeri, Alberto Barrón-Cedeño, Firoj Alam, Tommaso Caselli, Mücahid Kutlu, Julia Maria Struß, Francesco Antici, Maram Hasanain, Juliane Köhler, Katerina Korre, Folkert Leistra, Arianna Muti, Melanie Siegel, Mehmet Deniz Türkmen, Michael Wiegand, and Wajdi Zaghouani. Overview of the CLEF-2023 checkthat! lab: Task 2 on subjectivity detection. In Mohammad Aliannejadi, Guglielmo Faggioli, Nicola Ferro, and Michalis Vlachos, editors, *Working Notes of the Conference and Labs of the Evaluation Forum (CLEF 2023), Thessaloniki, Greece, September 18th to 21st, 2023*, volume 3497 of *CEUR Workshop Proceedings*, pages 236–249. CEUR-WS.org, 2023.

[O8] Mirko Del Moro, Serban Cristian Tudosie, Francesco Vannoni, Andrea Galassi, and Federico Ruggeri. Inception models for fashion image captioning: An extensive study on multiple datasets. In Avi Arampatzis, Evangelos Kanoulas, Theodora Tsikrika, Stefanos Vrochidis, Anastasia Giachanou, Dan Li, Mohammad Aliannejadi, Michalis Vlachos,

Guglielmo Faggioli, and Nicola Ferro, editors, *Experimental IR Meets Multilinguality, Multimodality, and Interaction - 14th International Conference of the CLEF Association, CLEF 2023, Thessaloniki, Greece, September 18-21, 2023, Proceedings*, volume 14163 of *Lecture Notes in Computer Science*, pages 3–14. Springer, 2023.

[O9] Piera Santin, Giulia Grundler, Andrea Galassi, Federico Galli, Francesca Lagioia, Elena Palmieri, Federico Ruggeri, Giovanni Sartor, and Paolo Torroni. Argumentation structure prediction in cjeu decisions on fiscal state aid. In *Proceedings of the Nineteenth International Conference on Artificial Intelligence and Law*, ICAIL '23, page 247–256, New York, NY, USA, 2023. Association for Computing Machinery. **GGS Rating: B-. ICORE Rating: C.**

[O10] **Federico Ruggeri**, Francesco Antici, Andrea Galassi, Katerina Korre, Arianna Muti, and Alberto Barrón-Cedeño. On the definition of prescriptive annotation guidelines for language-agnostic subjectivity detection. In Ricardo Campos, Alípio Mário Jorge, Adam Jatowt, Sumit Bhatia, and Marina Litvak, editors, *Proceedings of Text2Story - Sixth Workshop on Narrative Extraction From Texts held in conjunction with the 45th European Conference on Information Retrieval (ECIR 2023), Dublin, Ireland, April 2, 2023*, volume 3370 of *CEUR Workshop Proceedings*, pages 103–111. CEUR-WS.org, 2023.

[O11] Sezen Perçin, Andrea Galassi, Francesca Lagioia, Federico Ruggeri, Piera Santin, Giovanni Sartor, and Paolo Torroni. Combining WordNet and word embeddings in data augmentation for legal texts. In *Proceedings of the Natural Legal Language Processing Workshop 2022*, pages 47–52, Abu Dhabi, United Arab Emirates (Hybrid), December 2022. Association for Computational Linguistics.

[O12] Pavlo Seroyizhko, Zhanel Zhexenova, Muhammad Zohaib Shafiq, Fabio Merizzi, Andrea Galassi, and Federico Ruggeri. A sentiment and emotion annotated dataset for bitcoin price forecasting based on Reddit posts. In *Proceedings of the Fourth Workshop on Financial Technology and Natural Language Processing (FinNLP)*, pages 203–210, Abu Dhabi, United Arab Emirates (Hybrid), December 2022. Association for Computational Linguistics.

[O13] Federico Galli, Giulia Grundler, Alessia Fidelangeli, Andrea Galassi, Francesca Lagioia, Elena Palmieri, Federico Ruggeri, Giovanni Sartor, and Paolo Torroni. Predicting outcomes of italian VAT decisions. In Enrico Francesconi, Georg Borges, and Christoph Sorge, editors, *Legal Knowledge and Information Systems - JURIX 2022: The Thirty-fifth Annual Conference, Saarbrücken, Germany, 14-16 December 2022*, volume 362 of *Frontiers in Artificial Intelligence and Applications*, pages 188–193. IOS Press, 2022. **GGS Rating: C. ICORE Rating: C.**

[O14] Giulia Grundler, Piera Santin, Andrea Galassi, Federico Galli, Francesco Godano, Francesca Lagioia, Elena Palmieri, Federico Ruggeri, Giovanni Sartor, and Paolo Torroni. Detecting arguments in CJEU decisions on fiscal state aid. In Gabriella Lapesa, Jodi Schneider, Yohan Jo, and Sougata Saha, editors, *Proceedings of the 9th Workshop on Argument Mining, ArgMining@COLING 2022, Online and in Gyeongju, Republic of Korea, October 12 - 17, 2022*, pages 143–157. International Conference on Computational Linguistics, 2022.

[O15] Marco Lippi, Francesco Antici, Gianfranco Brambilla, Evaristo Cisbani, Andrea Galassi, Daniele Giansanti, Fabio Magurano, Antonella Rosi, Federico Ruggeri, and Paolo Torroni. AMICA: an argumentative search engine for COVID-19 literature. In Luc De Raedt, editor, *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI 2022, Vienna, Austria, 23-29 July 2022*, pages 5932–5935. ijcai.org, 2022. **GGS Rating: A++. ICORE Rating: A*.**

[O16] Eleonora Mancini, **Federico Ruggeri**, Andrea Galassi, and Paolo Torroni. Multimodal argument mining: A case study in political debates. In Gabriella Lapesa, Jodi Schneider, Yohan Jo, and Sougata Saha, editors, *Proceedings of the 9th Workshop on Argument Mining,*

ArgMining@COLING 2022, Online and in Gyeongju, Republic of Korea, October 12 - 17, 2022, pages 158–170. International Conference on Computational Linguistics, 2022.

[O17] Mattia Silvestri, Allegra De Filippo, Federico Ruggeri, and Michele Lombardi. Hybrid offline/online optimization for energy management via reinforcement learning. In Pierre Schaus, editor, *Integration of Constraint Programming, Artificial Intelligence, and Operations Research - 19th International Conference, CPAIOR 2022, Los Angeles, CA, USA, June 20-23, 2022, Proceedings*, volume 13292 of *Lecture Notes in Computer Science*, pages 358–373. Springer, 2022. **GGS Rating: B-. ICORE Rating: B.**

[O18] **Federico Ruggeri**. *Towards Unstructured Knowledge Integration in Natural Language Processing*. PhD thesis, alma, Giugno 2022.

[O19] Francesco Antici, Luca Bolognini, Matteo Antonio Inajetovic, Bogdan Ivasiuk, Andrea Galassi, and Federico Ruggeri. Subjectivita: An italian corpus for subjectivity detection in newspapers. In K. Selçuk Candan, Bogdan Ionescu, Lorraine Goeuriot, Birger Larsen, Henning Müller, Alexis Joly, Maria Maistro, Florina Piroi, Guglielmo Faggioli, and Nicola Ferro, editors, *Experimental IR Meets Multilinguality, Multimodality, and Interaction - 12th International Conference of the CLEF Association, CLEF 2021, Virtual Event, September 21-24, 2021, Proceedings*, volume 12880 of *Lecture Notes in Computer Science*, pages 40–52. Springer, 2021.

[O20] Ruta Liepina, **Federico Ruggeri**, Francesca Lagioia, Marco Lippi, Kasper Drazewski, and Paolo Torroni. Explaining potentially unfair clauses to the consumer with the CLAUDETTE tool. In Nikolaos Aletras, Ion Androutsopoulos, Leslie Barrett, Adam Meyers, and Daniel Preotiuc-Pietro, editors, *Proceedings of the Natural Legal Language Processing Workshop 2020 co-located with the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD 2020), Virtual Workshop, August 24, 2020*, volume 2645 of *CEUR Workshop Proceedings*, pages 61–64. CEUR-WS.org, 2020.

[O21] Francesca Lagioia, **Federico Ruggeri**, Kasper Drazewski, Marco Lippi, Hans-Wolfgang Micklitz, Paolo Torroni, and Giovanni Sartor. Deep learning for detecting and explaining unfairness in consumer contracts. In Michal Araszkiewicz and Víctor Rodríguez-Doncel, editors, *Legal Knowledge and Information Systems - JURIX 2019: The Thirty-second Annual Conference, Madrid, Spain, December 11-13, 2019*, volume 322 of *Frontiers in Artificial Intelligence and Applications*, pages 43–52. IOS Press, 2019.

Under Review

First author or equivalent contribution entries are highlighted in **bold**.

[R1] Mahmud Elahi Akhter, Federico Ruggeri, Iman Munire Bilal, Robert Procter, and Maria Liakata. On the merits of complex claim decomposition as a step to verification. *EACL*, 2025. **GGS Rating A. ICORE Rating: A**.

[R2] Luca Bolognini, Elena Palmieri, Nicolò Donati, Giulia Grundler, Gianmarco Pappacoda, Federico Ruggeri, Andrea Galassi, and Paolo Torroni. Detecting SDG contributions in scientific articles. *Nature Scientific Data*, 2025. **Scimago Quartile: Q1**.

[R3] Andrea Gerardo Russo, Federico Ruggeri, Davide Bombini Ivan Tomarchio, Paolo Torroni, and Giuseppe-Emiliano La Cara. Retrieval-augmented generation to to support railways engineering tasks: A case study. *Engineering Applications of Artificial Intelligence*, 2025. **Scimago Quartile: Q1**.

[R4] **Federico Ruggeri**, Eleonora Misino, Arianna Muti, Katerina Korre, Paolo Torrini, and Alberto Barrón-Cedeño. Let guidelines guide you: A prescriptive guideline-centered data annotation methodology. *TACL*, 2025. **Principal Investigator. Scimago Quartile: Q1**.

[R5] Sezen Perçin, Piera Santin, Andrea Galassi, Francesca Lagioia, Ruta Liepina, Federico Ruggeri, Giovanni Sartor, and Paolo Torroni. Automatic extraction of legal principles of the court of justice of the european union. *AI & Law*, 2024. **Scimago Quartile: Q1**.

Pre-Print

First author or equivalent contribution entries are highlighted in **bold**.

[P1] **Federico Ruggeri**, Marco Lippi, and Paolo Torroni. Combining transformers with natural language explanations. *ArXiv*, 2024.

[P2] **Federico Ruggeri**, Marco Lippi, and Paolo Torroni. Tree-constrained graph neural networks for argument mining. *ArXiv*, 2021.

Software Products

2024	
Python Library	*cinnamon*
Description	Cinnamon is a simple framework for general-purpose configuration and code logic decoupling. It was developed to offer two main functionalities: De-coupling a code logic from its regulating parameters; Re-use of code logic without effort
Role	Creator, Principal Maintainer.
Github	https://github.com/nlp-unibo/cinnamon
Course Repository	
Description	This is the official repository of the PhD course "Robust and Reproducible Research".
Role	Creator, Principal Maintainer.
Github	https://github.com/federicoruggeri/phdlectures-r3
Research Repository	
Description	This repository contains the code and resources for the project "Interlocking-free Selective Rationalization Through Genetic-based Learning"
Role	Co-Creator, Co-Maintainer
Github	https://github.com/federicoruggeri/gen-spp
Reference	[S2]
Research Repository	
Description	This repository contains the code and resources for the project "Let Guidelines Guide You: A Prescriptive Guideline-Centered Data Annotation Methodology"
Role	Creator, Maintainer
Github	https://github.com/federicoruggeri/gcam
Reference	[R4]
Research Repository	
Description	This repository contains the code and resources for the project "MAMKit: A Comprehensive Multimodal Argument Mining Toolkit".
Role	Co-Creator, Co-Maintainer.
Github	https://github.com/nlp-unibo/mamkit
Reference	[O2]
Research Repository	
Description	This repository contains the code and resources for the project "Combining Transformers with Natural Language Explanations".
Role	Creator, Principal Maintainer.
Github	https://github.com/nlp-unibo/bert-natural-explanations
Reference	[P1]

Research Repository
Description | This repository contains the code and resources for the project "A Corpus for Sentence-Level Subjectivity Detection on English News Articles".
Role | Creator, Co-Maintainer.
Github | https://github.com/nlp-unibo/newssd-eng
Reference | [S6]

Research Repository
Description | This repository contains the code and resources for the project "Multimodal Fallacy Classification in Political Debates".
Role | Co-Maintainer.
Github | https://github.com/nlp-unibo/multimodal-am-fallacy
Reference | [S8]

2023
Research Repository
Description | This repository contains the code and resources for the project "A Dataset of Argumentative Conversational Discussions on Scientific Papers".
Role | Co-Maintainer.
Github | https://github.com/UKPLab/acl2023-argscichat
Reference | [S10]

2022
Research Repository
Description | This repository contains the code and resources for the project "Multimodal Argument Mining: A Case Study in Political Debates".
Role | Co-Maintainer.
Github | https://github.com/nlp-unibo/multimodal-am
Reference | [O16]

Languages

Italian: Mothertongue
English: Advanced, C1
IELTS, British Council, 30 June 2018
Score: 7.5/9