PhD in Information Technology and Electrical Engineering
Università degli Studi di Napoli Federico II

PhD Student: Gian Marco Orlando

Cycle: XL

Training and Research Activities Report

Year: First

Tutor: Prof. Vincenzo Moscato

Co-Tutor: Dr. Valerio La Gatta

Date: October, 2025

Training and Research Activities Report

PhD in Information Technology and Electrical Engineering

Cycle: XL **Author: Gian Marco Orlando**

1. Information:

- **PhD student: Gian Marco Orlando**
- **DR number: DR999879**
- **Date of birth: 10/09/1995**
- **Master Science degree: Computer Engineering University: Università degli Studi di Napoli Federico II**
- **Doctoral Cycle: XL**
- **Scholarship type: DIETI - Dipartimento di Eccellenza**
- **Tutor: Prof. Vincenzo Moscato**
- **Co-tutor: Dr. Valerio La Gatta**

2. Study and training activities:

Activity	Type	Hours	Credits	Dates	Organizer	Certificate
PhD School - International School on Artificial Intelligence for Cognitive Technologies 2024	Doctoral School	30	6	10/12/2024 – 13/12/2024	Otto von Guericke University, Magdeburg, Germany, Co-organised and Hosted by: University of Naples Federico II	Y
Can we Rely on AI? Reliability Issues in Artificial Neural Networks and Potential Solutions for Autonomous Vehicles	Seminar	1	0.2	16/01/2025	Dr. Edoardo Giusto	Y
The Good, the Bad, and the Ugly in Quantum Computing: Computational Power, Intrinsic Noise, and Transient Faults	Seminar	1	0.2	17/01/2025	Dr. Edoardo Giusto	Y
FAIR Spoke Workshop 2025	Seminar	3	0.6	07/03/2025	Fondazione FAIR – Future Artificial Intelligence	Y

					Research	
On the security of Semantic Watermarking to detect AI-Generated Content	Seminar	1	0.2	29/04/2025	Prof. Luisa Verdoliva	Y
Methodologies and Tools for Conducting Systematic Literature Reviews and Systematic Mapping Studies	Course	14	3	28-29/04/2025 - 05-06-09-12-14/05/2025	Prof. Domenico Amalfitano	Y
Big Data Architecture and Analytics	Course	20	5	7-12-19-22-26-27-28/05/2025	Prof. Giancarlo Sperlì	Y
Sovranità Digitale: cos'è e quali sono le principali minacce al cyberspazio nazionale	Seminar	2	0.4	23/06/2025	Prof. Marcello Cinque	Y
Trusted Execution Environments for QPUs	Seminar	1	0.2	27/06/2025	Prof. Edoardo Giusto	Y
Advanced Models and Algorithms for Managing, Querying and Analyzing Big Multidimensional Data	Course	20	4	14-15-16-17-18/07/2025	Prof. Carlo Sansone	Y
Practical Network Intrusion Detection with Machine Learning and Generative AI	Course	14	4	1-2-7-8-9/10/2025	Dr. Giampaolo Bovenzi	Y
Healthcare Text Analytics in the AI Era	Course	23	6	29-30/09/2025 – 3-8-10-15-22/10/2025	Prof. Vincenzo Moscato	Y
Argumentation-Based Reasoning Frameworks for Public Interest Communication in Healthcare	Seminar	2	0.4	29/09/2025	Prof. Carlo Sansone, Prof. Elio Masciari	Y
Guardians or Threats? AI at the Frontlines of	Seminar	4	0.8	17/10/2025	Prof. Antonia Maria Tulino	Y

Cybersecurity						
Exact and Machine Learning-Guided Matheuristic Approaches for the Fliying Sidekick TSP	Seminar	1	0.2	17/10/2025	Prof. Maurizio Boccia	Y
Boosting Your Citations: Maximizing the Impact of Your Publications	Seminar	1	0.2	21/10/2025	Dr. Zachary M. Wilmot and Dr. Rong Han	Y

1) Courses, Seminar, Doctoral School, Research, Tutorship
2) Choose: Y or N

2.1. Study and training activities - credits earned

	Courses	Seminars	Research	Tutorship	Total
Bimonth 1	**6**	**0**	**5**	**0.2**	**11.2**
Bimonth 2	**0**	**0.4**	**5**	**0.3**	**5.7**
Bimonth 3	**0**	**0.8**	**5**	**0.2**	**6**
Bimonth 4	**8**	**0.6**	**5**	**0.3**	**13.9**
Bimonth 5	**4**	**0**	**5**	**0.3**	**9.3**
Bimonth 6	**10**	**1.6**	**10**	**0.3**	**21.9**
Total	**28**	**3.4**	**35**	**1.6**	**68**
Expected	**30 - 70**	**10 - 30**	**80 - 140**	**0 – 4.8**	

3. Research activity:

Generative Agent-Based Modeling

During my first year of PhD, my research activity focused on **developing, validating, and extending a Generative Agent-Based Modeling (GABM) framework**, an emerging simulation paradigm that integrates the reasoning and communication abilities of Large Language Models (LLMs) with traditional agent-based simulation (ABM) paradigms [1]. The goal of this framework is to reproduce social media dynamics by simulating realistic user behavior and emergent collective phenomena that arise from agent interactions.

Traditional ABM relies on manually crafted rules and predefined parameters to describe agents' actions and interactions [2]. While this approach has been instrumental in modeling collective dynamics such as opinion formation [3] or information diffusion [4], it suffers from **limited realism**, **parameter sensitivity**, and **researcher bias** due to oversimplified behavioral rules [5].

Recent advances in LLMs, however, have demonstrated remarkable capabilities in **reasoning, planning, and natural language interaction** [6], making them suitable to endow simulated agents with human-like decision-making, adaptability, and communication.

GABM leverages these capabilities to create **autonomous, context-aware agents** whose behaviors and

interactions emerge naturally from language-based reasoning rather than deterministic scripts, offering a more realistic and scalable way to simulate social phenomena.

The first phase of my research concentrated on the conceptualization, development and content-level validation of a GABM framework, culminating in the paper ***"Agent-Based Modelling Meets Generative AI in Social Network Simulations"*** (presented at the 16th International Conference on Advances in Social Network Analysis and Mining, **ASONAM 2024**). In this work, I introduced a modular architecture composed of three main components: *(i)* the **Characterization Module**, responsible for initializing each generative agent based on real users' personality traits and interests inferred from social media data through prompt engineering; *(ii)* the **Reasoning Module**, enabling agents to autonomously decide whether to post, re-share, or remain inactive, providing natural language justifications for their actions; and *(iii)* the **Interaction Module**, powered by a Retrieval-Augmented Generation (RAG) mechanism to expose agents to contextually relevant content and emulate recommendation strategies. Using a large-scale Twitter dataset from the 2020 U.S. election, we demonstrated that LLM-powered agents could accurately reproduce real users' linguistic styles, political orientations, and community-level behaviors. Moreover, we observed that preference-based recommendations led to stronger homophily [7] and echo chamber [8] formation, highlighting the framework's ability to replicate both individual- and group-level social media phenomena.

Building upon this foundation, the second phase of my research addressed the **structural validation** of the proposed framework by investigating whether well-known real-world global network phenomena naturally emerge within GABM simulations. This investigation resulted in the paper ***"Can Generative Agent-Based Modeling Replicate the Friendship Paradox in Social Media Simulations?"*** (presented at the 17th ACM Web Science Conference**, ACM WebSci 2025**). Specifically, we tested whether the **Friendship Paradox** [9]—the tendency for individuals to have fewer friends than their friends on average—and its generalizations (Activity Paradox [10], Virality Paradox [10], and Susceptibility Paradox [11]) arise spontaneously within our generative simulations. To this end, I extended the framework with a **dynamic social graph**, **agent memory modules** (short- and long-term memory), and an expanded set of social actions (follow, like, dislike, comment). Experiments conducted on datasets from the 2020 U.S. Election and the QAnon community revealed that the Friendship Paradox consistently emerged, driven primarily by infrequent but high-impact connections, mirroring patterns observed in real social networks. These findings provided strong empirical evidence that GABM can capture not only realistic content dynamics but also the emergent topological structures of online social systems.

Overall, my research this year contributed both **conceptual and empirical validation** of GABM as a viable approach for studying social phenomena. The first study confirmed the framework's fidelity at the micro and meso levels (agent realism, echo chambers), while the second validated its ability to naturally reproduce macro-level phenomena (network asymmetries and structural hierarchies). This dual validation lays the groundwork for future developments, including the integration of multi-agent coordination strategies, more advanced recommendation mechanisms, and analyses of emergent phenomena such as polarization and misinformation diffusion.

Applications of Generative Agents to Real-World Problems

Building upon the conceptual advances achieved through the development of the Generative Agent-Based Modeling (GABM) framework, I extended my research toward exploring **practical applications of generative agents**—autonomous, LLM-powered entities capable of reasoning, collaborating, and interacting with minimal human supervision. While the GABM framework focused on simulating large-scale social simulations, these studies investigated how generative agents can be employed to tackle domain-specific analytical and reasoning problems.

The first application, presented in the paper ***"Assessing the Potential of Generative Agents in Crowdsourced Fact-Checking"*** (published in ***Online Social Networks and Media***, Elsevier, 2025), focuses on the domain of **crowdsourced fact-checking**, a collaborative approach where large groups of human contributors assess the veracity of online claims [12]. Although this strategy has proven effective in scaling verification efforts and leveraging collective intelligence, it remains limited by human constraints such as low scalability, inconsistent judgment criteria, and cognitive biases. In my work, I designed a framework that employs crowds of generative agents—each endowed with realistic demographic and ideological profiles—to replicate the diversity and subjectivity of human annotators. Agents retrieved external evidence, evaluated claims across multiple quality dimensions (e.g., accuracy, completeness, speaker trustworthiness), and issued veracity judgments through structured questionnaires. Experimental results demonstrated that artificial crowds **outperformed human evaluators** in accuracy and agreement with expert fact-checkers, while showing reduced cognitive bias. This work represents the first attempts to operationalize the ***"wisdom of synthetic crowds"***, offering a scalable and less biased alternative for crowd-based misinformation detection.

The second application, detailed in ***"Generative Agent-Based Modeling with Large Language Models for Insider Threat Detection"*** (published in ***Engineering Applications of Artificial Intelligence***), extends the GABM paradigm to the cybersecurity domain. Here, I proposed a **hierarchical multi-agent framework** composed of *Specialized Agents*—responsible for analyzing distinct log categories—and a *Supervisor Agent* that synthesizes their reports through **Chain-of-Thought reasoning** to classify activities as benign or malicious. Validated on two benchmark datasets (PicoDomain and CERT r5.2), the system achieved **state-of-the-art recall** while introducing interpretability, proving effective in detecting subtle behavioral patterns associated with insider threats.

Together, these two applications extend the impact of GABM beyond social simulations, illustrating how generative agents can function also as **cognitive collaborators** capable of performing analytical, evaluative, and decision-making tasks in real-world settings.

Other Research Activities

In parallel to my main line of research on Generative Agent-Based Modeling and the design of generative agents, I engaged in several complementary projects aimed at advancing the understanding, safety, and transparency of AI systems, as well as exploring LLM-based methodologies for large-scale information extraction and multimodal reasoning. These works were carried out in collaboration with other members of the PRAISELab research group supervised by Prof. Vincenzo Moscato and, in some cases, through joint initiatives with external organizations such as the AI Incident Database (AIID).

A first line of work, presented in ***"Automating AI Failure Tracking: Semantic Association of Reports in AI Incident Database"*** (presented at the 28th European Conference on Artificial Intelligence, **ECAI**

2025), focused on developing a semantic retrieval framework to automate the linking of new AI failure reports to existing incidents within the AIID. This work addressed the scalability limitations of current manual curation, framing the problem as a semantic ranking task where transformer-based models compute similarity between new and historical reports. The proposed approach demonstrated that combining textual embeddings from titles and descriptions significantly improves retrieval performance, enabling a more consistent and scalable method for monitoring real-world AI failures. This collaboration with AIID contributes to the broader goal of establishing evidence-based infrastructures for AI safety and accountability.

A second direction explored the use of LLMs for knowledge graph generation and dynamic information organization, presented in ***"Scaling LLM-Based Knowledge Graph Generation: A Case Study of Italian Geopolitical News"*** (presented at the **IEEE International Conference on Big Data 2024**). In this work, I contributed to designing a modular and horizontally scalable architecture that continuously extracts and structures entity–relation triples from Italian geopolitical news articles. The system integrates an LLM-based NER/RE engine with a reliable queue architecture to ensure continuous, fault-tolerant data ingestion and mitigate hallucinations in triple extraction. This study showcased how LLMs can support real-time, multilingual knowledge integration, offering a robust tool for dynamic domains such as political or media analysis.

Complementing this work, ***"EuropeanLawAdvisor: An Open-Source Search Engine for European Laws"*** (presented at the **IEEE International Conference on Big Data 2024**) extended the idea of semantic retrieval to the legal domain. The system leverages transformer-based embeddings and vector similarity search to support advanced semantic navigation across multilingual European legislative texts. By improving the accessibility and interpretability of complex legal corpora, this project exemplifies how LLM-powered retrieval pipelines can enhance transparency and civic engagement in public governance contexts. The project also aligns with ongoing European efforts toward AI-assisted legal information systems, highlighting the societal value of explainable retrieval technologies.

A further research line addressed multimodal and social computing applications. In ***"A Multilingual Multimodal Medical Examination Dataset for Visual Question Answering in Healthcare"*** (38th IEEE International Symposium on Computer-Based Medical Systems, **CBMS 2025**, Special Track on Multimodal Artificial Intelligence in Healthcare), I contributed to the creation of a multilingual benchmark for evaluating Vision–Language Models (VLMs) in medical reasoning tasks across Spanish, English, and Italian. In parallel, in ***"From Explanation to Exploration: Promoting Diversity in Recommendation Systems"*** (**First International Workshop on Recommender Systems for Sustainability and Social Good, RecSoGood 2024, co-located with RecSys 2024**), I co-developed FEEDERS, a framework that leverages explanations within recommender systems to quantify and mitigate the "rabbit hole" effect in repeated user interactions. By promoting explanation-driven diversity, FEEDERS contributes to fairer and more transparent content recommendation strategies, addressing concerns around polarization and content homogeneity.

[1] Joon Sung Park, Joseph O'Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. 2023. Generative Agents: Interactive Simulacra of Human Behavior. In Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (UIST '23). Association for

Computing Machinery, New York, NY, USA, Article 2, 1–22, https://doi.org/10.1145/3586183.3606763

[2] E. Elliott, L.D.K.: Agent-based modeling in the social and behavioral sciences. Nonlinear Dynamics, Psychology, and Life Sciences, Vol. 8, No. 2, April, 2004 (2004)

[3] Mastroeni, L., Vellucci, P., Naldi, M.: Agent-based models for opinion formation: A bibliographic survey. IEEE Access 7, 58836–58848 (2019). https://doi.org/ 10.1109/ACCESS.2019.2913787

[4] Tseng, S., Nguyen, T.S.: Agent-based modeling of rumor propagation using expected integrated mean squared error optimal design. Applied System Innovation (2020)

[5] Bonabeau, E.: Agent-based modeling: Methods and techniques for simulating human systems. Proceedings of the national academy of sciences (2002)

[6] Kojima, T., Gu, S.S., Reid, M., Matsuo, Y., Iwasawa, Y.: Large language models are zero-shot reasoners. In: Annual Conference on Neural Information Processing Systems 2022, NeurIPS 2022, New Orleans, LA, USA, November 28 - December 9, 2022 (2022)

[7] McPherson, M., Smith-Lovin, L., Cook, J.M.: Birds of a feather: Homophily in social networks. Annual Review of Sociology (2001)

[8] Bruns, A.: Echo chamber? what echo chamber? reviewing the evidence. School of Communication. Digital Media Research Centre, Cardiff (2017)

[9] Scott L Feld. 1991. Why your friends have more friends than you do. American journal of sociology 96, 6 (1991), 1464–1477.

[10] Nathan Hodas, Farshad Kooti, and Kristina Lerman. 2013. Friendship paradox redux: Your friends are more interesting than you. In Proceedings of the International AAAI Conference on Web and Social Media, Vol. 7. 225–233

[11] Luca Luceri, Jinyi Ye, Julie Jiang, and Emilio Ferrara. 2024. The Susceptibility Paradox in Online Social Influence. arXiv preprint arXiv:2406.11553 (2024).

[12] Allen, J., Arechar, A. A., Pennycook, G., & Rand, D. G. (2021). Scaling up fact-checking using the wisdom of crowds. *Science advances*, *7*(36), eabf4393.

4. Research products:

1. Antonino Ferraro, Antonio Galli, Valerio La Gatta, Marco Postiglione, **Gian Marco Orlando**, Diego Russo, Giuseppe Riccio, Antonio Romano, Vincenzo Moscato; “Agent-Based Modelling Meets Generative AI in Social Network Simulations”; International Conference on Advances in Social Networks Analysis and Mining; ASONAM; **Published**; 2024

2. Antonino Ferraro, Antonio Galli, Valerio La Gatta, Marco Postiglione, Diego Russo, **Gian Marco Orlando**, Giuseppe Riccio, Antonio Romano, Vincenzo Moscato; "From Explanation to Exploration: Promoting DivErsity in Recommendation Systems"; International Workshop on Recommender Systems for Sustainability and Social Good; RecSoGood; **Published**; 2024

3. Diego Russo, **Gian Marco Orlando**, Antonio Romano, Giuseppe Riccio, Valerio La Gatta, Marco Postiglione, Vincenzo Moscato; "Scaling LLM-Based Knowledge Graph Generation: A Case Study of Italian Geopolitical News"; 2024 IEEE International Conference on Big Data (BigData); **Published**; 2024

4. Raffaele Russo, Diego Russo, **Gian Marco Orlando**, Antonio Romano, Giuseppe Riccio, Valerio La Gatta, Marco Postiglione, Vincenzo Moscato; "EuropeanLawAdvisor: an open source search engine for European laws"; 2024 IEEE International Conference on Big Data (BigData); **Published**; 2024

5. **Gian Marco Orlando**, Valerio La Gatta, Diego Russo, Vincenzo Moscato; "Can Generative Agent-Based Modeling Replicate the Friendship Paradox in Social Media Simulations?"; 17th ACM Web Science Conference 2025; WebSci 2025; **Published**; 2025

6. Giuseppe Riccio, Antonio Romano, Mariano Barone, **Gian Marco Orlando**, Diego Russo, Marco Postiglione, Valerio La Gatta, Vincenzo Moscato; "A Multilingual Multimodal Medical Examination Dataset for Visual Question Answering in Healthcare"; IEEE 38th International Symposium on Computer-Based Medical Systems; CBMS; **Published**; 2025

7. Luigia Costabile, **Gian Marco Orlando**, Valerio La Gatta, Vincenzo Moscato; "Assessing the Potential of Generative Agents in Crowdsourced Fact-Checking"; Online Social Networks and Media; OSNEM; **Published**; 2025

8. Diego Russo, **Gian Marco Orlando**, Valerio La Gatta, Vincenzo Moscato; "Automating AI Failure Tracking: Semantic Association of Reports in AI Incident Database"; 28th European Conference on Artificial Intelligence, ECAI 2025; **Published**; 2025

9. Antonino Ferraro, **Gian Marco Orlando**, Diego Russo; "Generative Agent-Based Modeling with Large Language Models for insider threat detection"; Engineering Applications of Artificial Intelligence; EAAI; **Published**; 2025

10. Mariano Barone, **Gian Marco Orlando**, Marco Perillo, Giuseppe Riccio, Antonio Romano, Diego Russo, Ferdinando Tammaro, Valerio La Gatta, Marco Postiglione, Vincenzo Moscato; "Improving Preliminary Diagnosis in Italian Through a Retrieval-Augmented Medical Chatbot"; 4th Italian Conference on Big Data and Data Science; ITADATA 2025; **Accepted**; 2025

11. Antonino Ferraro, Antonio Galli, Valerio La Gatta, Marco Postiglione, Giuseppe Riccio, Antonio Romano, **Gian Marco Orlando**, Diego Russo, Vincenzo Moscato; "MediCARE:

Medical Collaborative Agents REasoning over Interpretable Heterogeneous Graphs"; Artificial Intelligence in Medicine; **Submitted**

12. **Gian Marco Orlando**, Valerio La Gatta, Diego Russo, Vincenzo Moscato; "Validating Generative Agent-Based Modeling in Social Media Simulations through the Lens of the Friendship Paradox"; Information Processing and Management; IPM; **Submitted**

13. **Gian Marco Orlando**, Diego Russo, Valerio La Gatta, Vincenzo Moscato; "Coordinated Deception: Generative Agents for Multi-Agent Fake News Generation"; ACM Transactions on Intelligent Systems and Technology; TIST; **Submitted**

14. **Gian Marco Orlando**, Marco Perillo, Diego Russo, Vincenzo Moscato; "An Agent-Driven Architecture for Harmful Meme Detection through Multimodal Decomposition"; 27th International Symposium on Multimedia; ISM 2025; **Submitted**

15. Valerio La Gatta, **Gian Marco Orlando**, Marco Perillo, Ferdinando Tammaro, Vincenzo Moscato; "From Who They Are to How They Act: Behavioral Traits in Generative Agent-Based Models of Social Media"; 2026 ACM Web Conference; WWW 2026; **Submitted**

16. **Gian Marco Orlando**, Jinyi Ye, Valerio La Gatta, Mahdi Saeedi, Vincenzo Moscato, Emilio Ferrara, Luca Luceri; "Emergent Coordinated Behaviors in Networked LLM Agents: Modeling the Strategic Dynamics of Information Operations"; 2026 ACM Web Conference; WWW 2026; **Submitted**

17. Diego Russo, Valerio La Gatta, **Gian Marco Orlando**, Vincenzo Moscato; "Promoting Temporal Diversity in Recommender Systems via Explanations"; ACM Transactions on Information Systems; TOIS; **Submitted**

5. Conferences and seminars attended

1. 16th International Conference on Advances in Social Networks Analysis and Mining; ASONAM 2024; University of Calabria, Rende (CS), Calabria, Italy; 02/09/2024 - 05/09/2024; Presented the paper "Agent-Based Modelling Meets Generative AI in Social Network Simulations"

2. International School on Artificial Intelligence for Cognitive Technologies 2024; ISACT 2024; University of Naples Federico II, Centro Congressi Federico II, Via Partenope, 36, 80121 Naples NA, Italy; 10/12/2024 - 13/12/2024

3. 17th ACM Web Science Conference 2025; WebSci 2025; Rutgers University, New Brunswick, New Jersey, USA; 20/05/2025 - 23/05/2025; Presented the paper "Can Generative Agent-Based Modeling Replicate the Friendship Paradox in Social Media Simulations?"

4. 4th Italian Conference on Big Data and Data Science; ITADATA 2025; Università di Torino, Torino, Italy; 09/09/2025 – 11/09/2025; Presented the paper "Improving Preliminary Diagnosis in Italian Through a Retrieval-Augmented Medical Chatbot"

5. 28th European Conference on Artificial Intelligence; ECAI 2025; Bologna Congress Center, Bologna, Italy; 25/10/2025 – 30/10/2025; Presented the paper "Automating AI Failure Tracking: Semantic Association of Reports in AI Incident Database"

6. Activity abroad:

7. Activity in partner companies:

8. Tutorship

- Co-supervisor of nine master theses in Computer Engineering
- Weekly two hours of teaching activities regarding practical lectures/seminars during the course "Basi di Dati", Bachelor Degree in Computer Engineering
- Weekly two hours of teaching activities regarding practical lectures/seminars during the course "Sistemi Informativi", Master Degree in Management Engineering
- Weekly two hours of teaching activities regarding practical lectures/seminars during the course "Big Data Engineering", Master Degree in Computer Engineering
- 48 hours for the project "POLO REGIONALE PER L'INNOVAZIONE DIGITALE EVOLUTA - P.R.I.D.E." at the CERICT – Centro Regionale Information Communication Technology.