

Giovanni Pinna

Italiano (madrelingua), Inglese (fluente)

PROFILO

Ricercatore in IA e Ingegnere con un **Dottorato di Ricerca in Data Science Applicata & Intelligenza Artificiale** e 3 anni di esperienza tra industria e mondo accademico in elaborazione del linguaggio naturale (NLP), ottimizzazione di Large Language Models (LLM), Agenti IA e genetic improvement. Autore di **10+ pubblicazioni peer-reviewed** su sedi di prestigio (Scientific Reports, IEEE, Springer). Ho realizzato sistemi di IA proof-of-concept—inclusa una soluzione di Retrieval Augmented Generation (RAG) che ha **ridotto le chiamate al supporto clienti del 30%**. Ricercatore in Visita presso **UCL** (Londra) e **NOVA IMS** (Lisbona), con pubblicazioni a MSR, Scientific Reports e IEEE.

COMPETENZE

- Ho ottenuto la Laurea Magistrale in Ingegneria Elettronica e Informatica (LM-32) con la votazione di 108/110 presso l'Università di Trieste.
- Ho ottenuto a marzo il Dottorato di Ricerca in Applied Data Science and Artificial Intelligence *cum laude* presso l'Università di Trieste.
- Ho trascorso 8 mesi presso l'azienda Plus S.r.l. a Basovizza, dove ho acquisito conoscenze ed esperienza sui modelli LLM (Large Language Model), sullo sviluppo di tecniche RAG (Retrieval Augmented Generation), sulle metodologie basate su Agenti e sui metodi di addestramento e di fine tuning.
- Ho pubblicato 11 paper in conferenze e journal peer-reviewed nel campo dell'Intelligenza Artificiale, degli LLM e delle tecniche annesse, e nel campo dell'informatica. In molte di queste pubblicazioni ho affinato le tecniche riguardanti gli LLM, gli Agenti, le reti neurali artificiali e i metodi di addestramento e di fine tuning.

ESPERIENZA

University College London (UCL)

Ricercatore in Visita

Londra, Regno Unito

Set. 2025 — Dic. 2025

- Co-progettato **GA4GC**, il primo framework multi-obiettivo per l'ottimizzazione delle configurazioni di agenti IA per la generazione di codice basati su Large Language Model (LLM) tramite NSGA-II sul benchmark SWE-Perf, **ottenendo un miglioramento dell'hypervolume di 135×**, **una riduzione del runtime del 43,6%** mantenendo la correttezza e superando le baseline di random search (83,0% vs 53,1% di hypervolume cumulativo).
- Guidato uno studio empirico analizzando **7.156 pull request generate da IA** su 5 agenti IA di coding, evidenziando come la tipologia di task determini un **divario di 16 punti percentuali nei tassi di accettazione** (documentazione: 82,1% vs feature: 66,1%); articolo come primo autore accettato a MSR 2026.
- Co-progettato **HotCat**, un framework di feature selection green basato su NSGA-II per la tassonomia dei bug di hotfix, **ottenendo 0,59 di accuracy e 0,58 di NMI su 17 categorie in 129 secondi** senza overhead di risorse aggiuntivo rispetto alla configurazione di default.
- Co-autore di uno studio sull'inconsistenza messaggio-codice in **23.247 pull request agentiche**, annotando manualmente 974 PR e identificando 8 tipi di inconsistenza; **le PR ad alta inconsistenza hanno mostrato un'accettazione inferiore del 51,7% e tempi di merge 3,5×** più lunghi; pubblicato a MSR 2026.

NOVA Information Management School (NOVA IMS)

Ricercatore in Visita

Lisbona, Portogallo

Mag.—Ago. 2024 & Mag.—Ago. 2025

- Creato una nuova metrica di valutazione per Text-to-SQL (QAS) che integra la similarità semantica e strutturale delle query SQL, testata su **11 modelli** sul **dataset BIRD**; pubblicata su **Scientific Reports (Nature)**, superando i principali limiti degli approcci di valutazione binari.

- Sviluppato un Agente IA per task di Text-to-SQL utilizzando **solo modelli open-source di piccole dimensioni**, consentendo l'interrogazione di database in linguaggio naturale in modo economico e senza API proprietarie.

PLUS S.r.l., Area Science Park

Scienziato di IA Applicata

Trieste, Italia

Gennaio 2023 — Settembre 2023

- Progettato e distribuito un sistema proof-of-concept di Retrieval Augmented Generation (RAG), fine tuning, utilizzando Python, LangChain, LlamaIndex e HuggingFace Transformers, abilitando l'interrogazione in linguaggio naturale della documentazione aziendale e **riducendo il volume delle chiamate al call center clienti del 30%**.
- Supervisionato 1 tirocinante per 16 settimane nella progettazione e distribuzione di soluzioni di Retrieval Augmented Generation (RAG) e LLM basate su IA per clienti enterprise, guidando l'intero ciclo di sviluppo dalla progettazione dell'architettura al rilascio del proof-of-concept.
- **Progettato e presentato un sistema proof-of-concept di Agente IA** per l'interrogazione automatica di portali di servizi comunali, permettendo a utenti non tecnici di navigare procedure regolamentari complesse tramite il linguaggio naturale.

Università di Trieste

Assistente alla Didattica e Tutor di Dipartimento

Trieste, Italia

Gen. 2024 — Dic. 2024

- Progettato esercitazioni di laboratorio e valutato i progetti per **Basi di Dati** (Laurea Triennale in Ingegneria Informatica), guidando **100+ studenti** attraverso i cicli di progettazione di database e implementazione SQL.
- Ristrutturato i siti web dei corsi di Laurea Triennale, Magistrale e di Dottorato del Dipartimento di Data Science & IA, migliorando l'accessibilità delle informazioni per **200+ studenti**.

BDF S.p.A. / Fincantieri S.p.A.

Specialista di Supporto IT

Trieste, Italia

Apr. 2019 — Ago. 2019

U-BLOX Italia S.p.A.

Tirocinante Collaudo Qualità

Sgonico, Italia

Lug. 2015 — Set. 2015

FORMAZIONE

Dottorato di Ricerca in Data Science Applicata e Intelligenza Artificiale

Supervisionato dal Prof. Luca Manzoni e dal Prof. Andrea De Lorenzo

Trieste, Italia

Università di Trieste

Gen. 2023 — Dic. 2025

- Tesi: "[Application of Large Language Models: Addressing Real-World Challenges](#)" — Voto: **con Lode**
- Realizzato una pipeline di Genetic Improvement basata su Grammatical Evolution per correggere automaticamente il codice generato da LLM, **migliorando l'accuratezza delle soluzioni nel 50–75% dei casi** su 4 LLM e 25 problemi di benchmark; pubblicato a EuroGP 2024 (Springer) e su SN Computer Science.
- Supervisionato **3 tesisti magistrali e 1 triennale** su ricerche in Natural Language Processing (NLP) e generazione di codice; 2 tesi hanno contribuito direttamente a pubblicazioni peer-reviewed a EuroGP 2024 e IEEE WI-IAT 2024.

Laurea Magistrale in Ingegneria Elettronica e Informatica (LM-32)

Università di Trieste

Trieste, Italia

Set. 2019 — Ott. 2022

- Tesi in **Computer Vision**: "An Automatic Tool for the Recognition of Punches in Late-Medieval Panel Paintings". ([Articolo](#)) — Voto: 108/110

Scambio Erasmus+

Montanuniversität Leoben

Leoben, Austria

Set. 2021 — Feb. 2022

- Corsi principali: Production Data Analysis & Modeling, Fundamentals of Logistics Systems Engineering, Computational Methods and Tools, Decision-Making and Risk Analysis.

Laurea Triennale in Ingegneria Elettronica e Informatica

Università di Trieste

Trieste, Italia

Set. 2015 — Mar. 2019

PUBBLICAZIONI SELEZIONATE

1. **Pinna, G.**, Gong, J., Williams, D., Sarro, F. (2026). “Comparing AI Coding Agents: A Task-Stratified Analysis of Pull Request Acceptance.” *MSR 2026*, in concomitanza con ICSE 2026.
2. Gong, J., **Pinna, G.**, Bian, Y., & Zhang, J. M. (2026). “Analyzing Message-Code Inconsistency in AI Coding Agent-Authored Pull Requests.” *MSR 2026 (MSR 2026 Distinguished Mining Challenge Paper Award)*, in concomitanza con ICSE 2026.
3. **Pinna, G.**, Perezhohin, Y., Manzoni, L., Castelli, M., De Lorenzo, A. (2025). “Redefining Text-to-SQL Metrics by Incorporating Semantic and Structural Similarity.” *Scientific Reports* 15.1.
4. **Pinna, G.**, Ravalico, D., Rovito, L., Manzoni, L., De Lorenzo, A. (2025). “Exploring the Effect of Genetic Improvement for LLM-Generated Code.” *SN Computer Science* 6.7.
5. **Pinna, G.**, Ravalico, D., Rovito, L., Manzoni, L., De Lorenzo, A. (2024). “Enhancing LLMs-Based Code Generation by Leveraging Genetic Improvement.” *EuroGP 2024*, LNCS vol. 14631, Springer.
6. **Pinna, G.**, Tugnoli, D., Manzoni, L., De Lorenzo, A. (2024). “From Courts to Comprehension: Can LLMs Make Judgments More Accessible?” *IEEE/WIC WI-IAT 2024*.
7. Gong, J., et al. incl. **Pinna, G.** (2025). “GA4GC: Greener Agent for Greener Code via Multi-Objective.” *SSBSE 2025*.
8. de la Cal, L., et al. incl. **Pinna, G.** (2025). “HotCat: Green and Effective Feature Selection for HotFix Bug Taxonomy.” *SSBSE 2025*.
9. Zulich, M., Macovaz, V., **Pinna, G.**, Pellegrino, F.A. (2023). “An AI System for Automatic Recognition of Punches in 14th-Century Panel Painting.” *IEEE Access*.
10. **Pinna, G.**, Ravalico, D., Rovito, L., Manzoni, L., & De Lorenzo, A. (2025). Improving LLM-Generated Code via Genetic Improvement: A Summary of Recent Advances. *CEUR Workshop Proceedings*.
11. Fenu, C., **Pinna, G.** “L’Edicola della Storia di Trieste: NLP e distant reading per il confronto tra periodici storici italiani e sloveni nella Trieste del 1902.” *AIUCD 2026*.

PROGETTI

Pipeline NLP per l’Analisi di Giornali Storici Multilingue. Sviluppato l’intera pipeline software (Python, BERTopic, spaCy, modelli di sentiment HuggingFace) per un progetto di Digital Humanities che confronta giornali italiani e sloveni della Trieste del 1902, includendo correzione OCR assistita da LLM, topic modeling e sentiment analysis cross-lingua su **50 numeri**. Articolo accettato ad **AIUCD 2026**.

Contamination Lab (CLab), Università di Trieste, 2020–2022. Raggiunta la **top 7** nella competizione universitaria di imprenditorialità, sviluppando un’idea di business con mentori accademici ed esperti del settore. ([Presentazione Video](#))

DIDATTICA E SUPERVISIONE

CORSI

Basi di Dati, Università di Trieste, Laurea Triennale in Ingegneria Informatica e Laurea Triennale in Intelligenza Artificiale e Data Analytics, 2024.

STUDENTI MAGISTRALI

Damiano Ravalico, Tesi in Ingegneria Informatica, 2023.

Davide Tugnoli, Tesi in Ingegneria Informatica, 2024.

Lorenzo Medeot, Tesi in Ingegneria Informatica, 2025.

STUDENTI TRIENNALI

Andrea Giacomazzi, Tesi in Intelligenza Artificiale e Data Analytics, 2025.

PREMI E FINANZIAMENTI

Finanziamento di ricerca per crediti di calcolo (5k \$), Google Cloud, 2024.

CERTIFICAZIONI

McKinsey Forward Program — percorso di sviluppo professionale di 10 settimane a cura di McKinsey.org.

CONFERENZE E SUMMER SCHOOL

Scuole di Specializzazione: AthNLP (Atene, Set. 2024; selezionato tra 200+ candidati), LxMLS (Lisbona, Lug. 2024), ALPS (Alpi Francesi, Apr. 2024; ~50 partecipanti selezionati), DeepLearn (Gran Canaria, Lug. 2023), **OxML percorso finanza (Oxford, Lug. 2023)**.
Volontariato: EACL 2024 (St Julian's, Malta), IEEE CCTA 2022 (Trieste, Italia).
Affiliazioni: AILC (Associazione Italiana di Linguistica Computazionale), AI2S, Mentors4U (mentee).

LINGUE

Italiano (madrelingua), Inglese (fluente).