

# Giovanni Simonini

Professore Associato – Sistemi di Elaborazione delle Informazioni (IINF-05/A)  
Università degli Studi di Modena e Reggio Emilia, Dip. di Ingegneria “Enzo Ferrari”

[simonini@unimore.it](mailto:simonini@unimore.it) · [www.giovanisimonini.com](http://www.giovanisimonini.com)

Aggiornato: settembre 2026

## Profilo di ricerca

Giovanni Simonini è Professore Associato presso l'Università degli Studi di Modena e Reggio Emilia e svolge attività di ricerca su scalable data management, data integration e AI-driven methods for industrial data analysis. La sua attività di ricerca attuale comprende **LLM context management** e l'applicazione di **agentic LLM-based techniques** in contesti industriali, con particolare attenzione a sistemi di data-centric AI affidabili, process support, automation e human-in-the-loop data management.

## Posizioni attuali e precedenti

- |                                                                                                                                                                                                                                                                                                   |                        |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------|
| <b>Professore Associato</b><br><i>Università degli Studi di Modena e Reggio Emilia, Dip. di Ingegneria “Enzo Ferrari”</i><br>Area di ricerca: Scalable Data Management; LLM context management. Abilitato alla prima fascia, Abilitazione Scientifica Nazionale, IINF-05/A, dal 19 novembre 2024. | Nov 2022 – Presente    |
| <b>Ricercatore a tempo determinato, RTD-B</b><br><i>Università degli Studi di Modena e Reggio Emilia</i><br>Attività di ricerca nell'ambito dello scalable data management.                                                                                                                       | Nov. 2019 – Ott 2022   |
| <b>Postdoctoral Associate</b><br><i>Massachusetts Institute of Technology – Computer Science and Artificial Intelligence Laboratory</i><br>Attività di ricerca con il Prof. Michael Stonebraker su automated data cleaning e preparation in machine learning pipelines.                           | Sett. 2018 – Nov. 2019 |
| <b>Assegnista di ricerca</b><br><i>Università degli Studi di Modena e Reggio Emilia</i><br>Attività di ricerca su scalable data integration ed entity resolution.                                                                                                                                 | Gen 2016 – Set 2018    |
| <b>Dottorando di ricerca</b><br><i>Università degli Studi di Modena e Reggio Emilia</i><br>Attività di ricerca di dottorato su data integration ed entity resolution. Co-advisor: Prof. H. V. Jagadish, University of Michigan.                                                                   | Gen 2013 – Gen 2016    |

## Progetti di ricerca scientifica e industriale selezionati

- |                                                                                                                                                                                                                                                                                                                                                 |                 |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------|
| <b>Collaborazione industriale con SPAL Automotive</b><br>Ruolo: Coordinatore Scientifico. Collaborazione attiva sull'applicazione di <b>AI and LLM-based agentic techniques</b> in contesti industriali, inclusa la supervisione di due posizioni di dottorato focalizzate su AI-driven industrial data analysis, process support e automation. | 2026 – Presente |
| <b>AI for Industry – AIMAG Spa</b><br>Progetto industriale sull'applicazione di metodi di AI a dati e processi industriali.                                                                                                                                                                                                                     | 2026            |
| <b>Emilia-Romagna PR-FESR 2022 DATHA</b><br>Ruolo: Principal Investigator. Progetto regionale di ricerca e innovazione su data-driven methods and technologies.                                                                                                                                                                                 | 2023 – 2026     |
| <b>PRIN 2022 National Project – “Discount Quality for Responsible Data Science”</b><br>Ruolo: Local Principal Investigator. Progetto finanziato dal MIUR su data quality e responsible data science.                                                                                                                                            | 2023 – 2025     |
| <b>ML Techniques for Wearable Devices</b><br>Ruolo: Principal Investigator. Progetto industriale su machine learning techniques for wearable devices.                                                                                                                                                                                           | 2021            |
| <b>ECOSISTER, PNRR</b><br>Ruolo: Partecipante. Progetto nell'ambito del Piano Nazionale di Ripresa e Resilienza.                                                                                                                                                                                                                                | 2025            |

## Competenze principali

- **Agentic LLMs**: LLM-based agents for industrial data analysis, automation, and decision support.
- **LLM context management**: data preparation, retrieval, and integration for reliable LLM applications.
- **Scalable data management**: data integration, data preparation, data quality, and data cleaning at scale.
- **Vector and hybrid search**: adaptive hybrid search in vector databases and scalable indexing.
- **Industrial AI**: company collaborations on analytics, automation, AI workflows, and applied machine learning.

## Didattica e supervisione di dottorato

- Big Data Management, MSc in Computer Engineering and AI Engineering, UNIMORE, 2023 – Presente.

- Quantitative Methods and Computer Science, Data Analysis for Economics, UNIMORE, 2023 – Presente.

### Ulteriori attività didattiche

- Visiting Professor, Aalborg University, corso di dottorato su Big Data Integration, 2024.
- Tools and Techniques for Big Data Analysis, CINECA, quattro edizioni, 2014 – 2017.

### Supervisione di dottorato

- 2 dottorandi attualmente supervisionati come advisor: Bilel Arfaoui, finanziato da Skybackbone Engenio; Giovanni Malaguti, finanziato da Leonardo S.p.A.
- 2 dottorandi che hanno completato il percorso sotto la mia supervisione come advisor: Adeel Aslam, attualmente Postdoc presso University of Southampton; Giulio de Sabbata, attualmente Data Scientist.
- 1 dottorando che ha completato il percorso sotto la mia supervisione come co-advisor: Luca Zecchini, attualmente Researcher presso TU Berlin.
- 2 posizioni di dottorato nell'ambito della collaborazione attiva con SPAL Automotive su AI and LLM-based agentic techniques for industrial applications.

### Collaborazioni e network selezionati

---

Felix Naumann, Hasso Plattner Institute, Germania, su entity resolution, data integration e table overlap; Michael Stonebraker, MIT CSAIL, USA, su automated data cleaning and preparation in ML pipelines; H. V. Jagadish, University of Michigan, USA, su data integration ed entity resolution; Themis Palpanas, Université Paris Cité, Francia, su scalable entity resolution e meta-blocking; George Papadakis, University of Athens / Athena RC, Grecia, su entity resolution e meta-blocking.

### Riconoscimenti e attività di servizio alla comunità scientifica

---

- Abilitazione Scientifica Nazionale alla prima fascia, IINF-05/A, 2024.
- Honorable Mention for AI and Big Data, Gruppo 2003, 2021.
- Best PhD Thesis Award, IEEE Computer Society – Italy Section, 2017.
- Coordinatore del curriculum Computer Science and Engineering, Dottorato in ICT, UNIMORE, 2024 – Presente.
- Membro del Collegio dei Docenti, Dottorato in ICT, UNIMORE, 2019 – Presente.
- Proceedings Chair, VLDB 2025 e 2026.
- Co-Chair, ACM SIGMOD Programming Contest, 2021 e 2022.
- Program Committees (più edizioni): ICDE, SIGMOD, VLDB, EDBT.
- Reviewer per IEEE TKDE, Data & Knowledge Engineering, ACM JDIQ.

### Pubblicazioni selezionate

---

1. A. Aslam, L. Gagliardelli, E. Rezig, G. Konstantinidis, G. Simonini. Versatile Sketch-Based Attribute Filtering for Hybrid Vector Search. *IEEE Transactions on Knowledge and Data Engineering*, pp. 6270-6284, vol. 38, 2026.
2. A. Aslam, R. Khan, G. Simonini, and G. Konstantinidis. Adaptive query-aware hybrid search in vector databases. *EDBT*, 2026.
3. L. Zecchini, V. Efthymiou, F. Naumann, and G. Simonini. Deduplicated sampling on-demand. *Proceedings of the VLDB Endowment*, 2025.
4. L. Zecchini, Z. Abedjan, V. Efthymiou, and G. Simonini. RadLER: Deduplicated sampling on-demand. *Proceedings of the VLDB Endowment*, 2025.
5. B. Pernici et al. Sustainable quality in data preparation. *ACM Journal of Data and Information Quality*, 2025.
6. L. Zecchini, T. Bleifuß, G. Simonini, S. Bergamaschi, and F. Naumann. Determining the largest overlap between tables. *Proceedings of the ACM on Management of Data*, 2024.
7. G. Simonini, L. Zecchini, S. Bergamaschi, and F. Naumann. Entity resolution on-demand. *Proceedings of the VLDB Endowment*, 2022.
8. L. Gagliardelli, G. Papadakis, G. Simonini, S. Bergamaschi, and T. Palpanas. Generalized supervised meta-blocking. *Proceedings of the VLDB Endowment*, 2022.
9. E. K. Rezig, L. Cao, G. Simonini, M. Schoemans, S. Madden, N. Tang, M. Ouzzani, and M. Stonebraker. Dagger: A data, not code, debugger. *CIDR*, 2020.
10. E. K. Rezig et al. Data Civilizer 2.0: A holistic framework for data preparation and analytics. *Proceedings of the VLDB Endowment*, 2019.
11. G. Simonini, G. Papadakis, T. Palpanas, and S. Bergamaschi. Schema-agnostic progressive entity resolution. *IEEE TKDE*, 2019.

12. G. Simonini, L. Gagliardelli, S. Bergamaschi, and H. V. Jagadish. Scaling entity resolution: A loosely schema-aware approach. *Information Systems*, 2019.
13. G. Simonini, S. Bergamaschi, and H. V. Jagadish. BLAST: A loosely schema-aware meta-blocking approach for entity resolution. *Proceedings of the VLDB Endowment*, 2016.