

Donatella Firmani

E-mail	donatella.firmani@uniroma1.it
Web	https://sites.google.com/uniroma1.it/donatellafirmani
Scholar	2285 (citazioni), 21 (h-index)
Data	30 aprile 2026

Esperienza Lavorativa

2024–oggi	Professore Associato <i>Sapienza Università di Roma, Roma (IT)</i>
2021–2024	RTD-B (Ricercatore a tempo determinato di tipo B) <i>Sapienza Università di Roma, Roma (IT)</i>
2019–2021	RTD-A (Ricercatore a tempo determinato di tipo A) <i>Università degli Studi Roma Tre, Roma (IT)</i>
2016–2019	Assegnista di ricerca <i>Università degli Studi Roma Tre, Roma (IT)</i>
2013–2015	Assegnista di ricerca <i>Università degli Studi di Roma Tor Vergata, Roma (IT)</i>
2012	Visiting researcher <i>DIMACS – AT&T Labs, Florham Park (NJ) Tutor: Prof. Graham Cormode, Prof. Barna Saha, Dr. Divesh Srivastava</i>

Istruzione

Ott 2013	Dottorato di Ricerca in Ingegneria Informatica <i>Sapienza Università di Roma Titolo della tesi: "Large Scale Graph Algorithms and Applications" Relatore: Prof. Giorgio Ausiello</i>
Nov 2010	Esame di Stato per l'abilitazione all'esercizio della professione di Ingegnere
Mar 2009	Laurea Magistrale in Ingegneria Informatica <i>Sapienza Università di Roma, Roma (IT), con lode</i>
Feb 2007	Laurea Triennale in Ingegneria Informatica <i>Sapienza Università di Roma, Roma (IT), con lode</i>

Breve Biografia

Donatella Firmani è Professoressa Associata presso il Dipartimento di Scienze Statistiche della Sapienza Università di Roma. La sua ricerca si concentra sull'estrazione della conoscenza e sulla gestione affidabile dei dati, con applicazioni nell'elaborazione del linguaggio naturale e nell'analisi delle reti. Ha conseguito il Dottorato di Ricerca in Ingegneria Informatica nel 2013 presso la Sapienza, con una tesi sugli algoritmi su grafi di grandi dimensioni. Nel 2012 è stata ricercatrice visitatrice presso AT&T Labs, finanziata dalla Rutgers University, nell'ambito di attività di integrazione dei dati e data streaming.

I suoi contributi sono stati riconosciuti con due Best Paper Award, rispettivamente all'IEEE Services Congress 2019 e alla conferenza DL4KG 2021. È stata Program Chair per tre edizioni del workshop SIGMOD sull'uso di tecniche di intelligenza artificiale per la gestione dei dati (aiDM) ha recentemente ricoperto il ruolo di Program Chair per le conferenze SEBD 2025 (33° Simposio sui Sistemi Avanzati di Basi di Dati) e IRCDL 2025 (21^a Conferenza su Information and Research Science Connecting to Digital and Library Science). È coinvolta nei comitati di programma di conferenze di riferimento come SIGMOD, VLDB e ICDE, ed è attualmente Associate Editor per l'ACM Journal of Data and Information Quality e membro del Review Board del VLDB Journal.

Coordina diversi progetti di ricerca nazionali ed europei su operazioni sui dati basate sugli intent, costruzione di knowledge graph e machine learning spiegabile. Tra questi, coordina l'unità Sapienza del progetto europeo H2020 INTEND.

Progetti di Ricerca

2025–2027	Principal Investigator – Progetto Sapienza <i>Knowledge Graph technologies for the next-generation Computing Continuum</i>
2024–oggi	Unit Leader – Progetto Europeo H2020 <i>INTEND: Intent-based data operation in the computing continuum</i>
2022–2024	Principal Investigator – Progetto Sapienza <i>Trustworthy Technologies for Augmenting Knowledge Graphs</i>
2021–2024	Principal Investigator – Progetto SEED PNRR <i>FLOWER: Frontiers in Linking records: knOWledge graphs, Explainability and tempoRal data</i>
2020	Principal Investigator – Progetto Industriale <i>Document Labelling with OCR Technologies, MediatICA srl</i>
2018–2019	Collaboratrice – Progetto Regione Lazio <i>In Codice Ratio: estrazione della conoscenza da documenti antichi</i>
2014–2017	Collaboratrice – Progetto PRIN <i>AMANDA: Algorithms for MAssive and NEtworked DAta</i>
2010–2012	Collaboratrice – Progetto PRIN <i>AlgoDEEP: Algorithmic Challenges for Data-intensive Processing on Emerging Computing Platforms</i>

Didattica Recente

2025–26 (1 anno)	Docente – Big Data Analytics <i>Sapienza Università di Roma, Laurea Magistrale in Metodi Statistici e Applicazioni</i>
2021–26 (5 anni)	Docente – Gestione dei Dati nella Statistica Ufficiale (poi: Pulizia e Integrazione dei Dati) <i>Sapienza Università di Roma, Laurea Magistrale in Metodi Statistici e Applicazioni</i>
2023–24 (1 anno)	Docente – Approcci Semantici per l'Entity Resolution <i>Sapienza Università di Roma, Dottorato Nazionale in Intelligenza Artificiale</i>
2021–26 (5 anni)	Docente – Informatica <i>Sapienza Università di Roma, Lauree Triennali in Statistica Impresa e Mercato, Statistica Economia Finanza e Assicurazioni, Statistica Economia e Società</i>
2019–20 (1 anno)	Docente – Approcci Moderni all'Entity Resolution <i>Università Roma Tre, Dottorato in Ingegneria Informatica</i>
2019–21 (2 anni)	Docente – Fondamenti di Informatica <i>Università Roma Tre, Laurea Triennale in Ingegneria Meccanica</i>

Attività di Servizio

Sono stata e sono attualmente membro del Program Committee (PC) per SEBD 2018–24, ICDE 2019–20, CIKM 2019 e 2026, EDBT 2020–26, SIGMOD 2023–26, VLDB 2024–27 e vari workshop internazionali.

2027	Senior PC member – ICDE <i>IEEE International Conference on Data Engineering</i>
2026	Diversity, Inclusion and Equity Chair – SIGMOD <i>2026 ACM SIGMOD/PODS International Conference on Management of Data</i>
2025	PC Chair – SEBD <i>33° Simposio sui Sistemi Avanzati di Basi di Dati</i>
2025	PC Chair – IRCDL <i>21ª Conferenza su Information and Research Science Connecting to Digital and Library Science</i>
2025	Research Area Chair – ICDE <i>IEEE International Conference on Data Engineering</i>
2025	Diversity, Inclusion and Equity Chair – SIGMOD <i>2025 ACM SIGMOD/PODS International Conference on Management of Data</i>
2022–2024	PC Chair – aiDM

	<i>International Workshop on Exploiting Artificial Intelligence Techniques for Data Management, co-located con SIGMOD22</i>
2020–21	Programming Contest Chair – SIGMOD <i>International Conference on Management of Data</i>
2019–21	PC Chair – PIE <i>International Workshop on Processing Information Ethically, co-located con CAiSE19, EDBT/ICDT20 e EDBT/ICDT21</i>
2020	Poster Chair – ICDE <i>36^a IEEE International Conference on Data Engineering</i>
2019–20	PC Chair – DI2KG <i>International Workshop on Challenges and Experiences from Data Integration to Knowledge Graphs, co-located con KDD19 e VLDB20</i>

Attività Editoriale

2020–oggi	Associate Editor <i>ACM Journal of Data and Information Quality</i>
2025–oggi	Review Board <i>VLDB Journal</i>
2021	Guest Editor <i>Frontiers in Big Data in Data Mining and Management</i>

Panel e Seminari su Invito

Giu 2025	Panel Moderator – SIGMOD 2025, Berlino (Germania) <i>"Towards a Broader Impact: Expanding the Reach of Data Management Research"</i> <i>Panelisti: Sihem Amer-Yahia (CNRS), Yael Amsterdamer (Bar-Ilan), Sourav S. Bhowmick (NTU), Carlo Curino (Microsoft), Sainyam Galhotra (Cornell), Gianmaria Silvello (Univ. di Padova)</i>
Mar 2024	Panelista – EDBT 2024, Paestum (Italia) <i>"How to write (and receive) reviews" Con Paolo Missier, Volker Markl, Anastassia Ailamaki, Benny Kimelfeld, Sudeepa Roy</i>
Mar 2021	Relatrice <i>"Modern Approaches for Entity Resolution", US Patent and Trademark Office, Alexandria (Virginia) Durante lo Symposium on Entity Resolution</i>
Mar 2021	Relatrice <i>"Beyond Crowdsourcing: Working at Scale Using AI and Machine Learning with Manuscripts" Harvard University, Cambridge (Massachusetts)</i>
Gen 2012	Relatrice <i>"MST and Connectivity of Large Scale Graphs in MapReduce", NII Shonan School, Tokyo (Giappone) Durante lo Shonan Meeting No. 2012-1</i>

Premi e Riconoscimenti

2021	Best Paper Award – DL4KG <i>Per il lavoro "Knowledge Graph Embeddings or Bias Graph Embeddings? A Study of Bias in Link Prediction Models" (con Andrea Rossi e Paolo Merialdo)</i>
2019	Best Paper Award – IEEE ICWS <i>Per il lavoro "On Computing Throttling Rate Limits in Web APIs through Statistical Inference" (con Francesco Leotta e Massimo Mecella)</i>
2018	Reproducibility Award – SIGMOD <i>Per il lavoro "Robust Entity Resolution using Random Graphs" (con Sainyam Galhotra, Barna Saha, Divesh Srivastava)</i>
2012	Visiting Grant – Rutgers University, Piscataway (NJ) <i>Finanziato da NSF/CCF, award no. 1144502</i>

Pubblicazioni Selezionate

L'elenco completo è disponibile online tramite i seguenti servizi:

DBLP → dblp.org/pid/05/10127 | Google Scholar → scholar.google.com/citations?user=aoB1hvoAAAAJ

Conferenze

[Conf1] Jerin George Mathew e Donatella Firmani, Divesh Srivastava. Evaluating methods for efficient entity count estimation. PVLDB, VLDB Endowment, 2025.

[Conf2] Sihem Amer-Yahia, Lorenzo Balzotti, Jerin George Mathew, Donatella Firmani, e Riccardo Torlone. R-fairness: Assessing fairness of ranking in subjective data. ACL 2025.

[Conf3] Donatella Firmani, Giorgio Grani, e Flavia Tagliafierro. Breaking down accuracy with subspace optimization. EDBT 2024, pp. 790–793.

[Conf4] Andrea Rossi, Donatella Firmani, Paolo Merialdo, e Tommaso Teofili. Explaining link prediction systems based on knowledge graph embeddings. SIGMOD 2022, pp. 2062–2075.

[Conf5] Sainyam Galhotra, Donatella Firmani, Barna Saha, e Divesh Srivastava. Hierarchical entity resolution using an oracle. SIGMOD 2022, pp. 414–428.

[Conf6] Sainyam Galhotra, Donatella Firmani, Barna Saha, e Divesh Srivastava. Robust entity resolution using random graphs. SIGMOD 2018, pp. 3–18.

[Conf7] Alessio Conte, Donatella Firmani, Maurizio Patrignani, Caterina Mordente, e Riccardo Torlone. Fast enumeration of large k-plexes. KDD 2017, pp. 115–124.

Riviste

[Riv1] Andrea De Angelis, Maurizio Mazzei, Federico Piai, Paolo Merialdo, Giovanni Simonini, Luca Zecchini, Sonia Bergamaschi, Donatella Firmani, Xu Chu, Peng Li, e Renzhi Wu. Experiences and lessons learned from the SIGMOD entity resolution programming contests. SIGMOD Rec., 52(2):43–47, 2023.

[Riv2] Andrea Rossi, Denilson Barbosa, Donatella Firmani, Antonio Matinata, e Paolo Merialdo. Knowledge graph embedding for link prediction: A comparative analysis. ACM Trans. Knowl. Discov. Data, 15(2):14:1–14:49, 2021.

[Riv3] Elena Nieddu, Donatella Firmani, Paolo Merialdo, e Marco Maiorino. In codice ratio: A crowd-enabled solution for low resource machine transcription of the vatican registers. Inf. Process. Manag., 58(5):102606, 2021.

[Riv4] Sainyam Galhotra, Donatella Firmani, Barna Saha, e Divesh Srivastava. Efficient and effective ER with progressive blocking. VLDB J., 30(4):537–557, 2021.

[Riv5] Alessio Conte, Donatella Firmani, Maurizio Patrignani, e Riccardo Torlone. A meta-algorithm for finding large k-plexes. Knowl. Inf. Syst., 63(7):1745–1769, 2021.

[Riv6] Donatella Firmani, Barna Saha, e Divesh Srivastava. Online entity resolution using an oracle. PVLDB, 9(5):384–395, 2016.

[Riv7] Donatella Firmani, Massimo Mecella, Monica Scannapieco, e Carlo Batini. On the meaningfulness of "big data quality". Data Science and Engineering, 1(1):6–20, 2015.