

Luca Gagliardelli

CURRICULUM VITAE

ULTIMO AGGIORNAMENTO: MAGGIO 2024

INFORMAZIONI PERSONALI

<i>Indirizzo</i>	Via Pietro Vivarelli 10, 41125 Modena, Italy
<i>email</i>	luca.gagliardelli@unimore.it
<i>Pagina personale</i>	www.lucagagliardelli.com
<i>Telefono</i>	(+39) 3403809971

ESPERIENZE DI RICERCA

Luca Gagliardelli è un ricercatore a tempo determinato (RTDA - PNRR) presso il Database Group (DBGroup) del Dipartimento di Ingegneria "Enzo Ferrari" dell'Università degli Studi di Modena e Reggio Emilia, situata a Modena, Italia. I suoi interessi di ricerca ricadono nell'area della Big Data integration e Big Data management. Con la sua ricerca ha portato avanzamenti nel campo dell'Entity Resolution, sviluppando tecniche di meta-blocking scalabili che possono essere sfruttate per integrare Big Data in contesti altamente eterogenei, che sono state rese disponibili in un framework open source chiamato [SparkER](#).

Dopo aver conseguito la laurea magistrale in Ingegneria Informatica ad aprile 2015, ha ottenuto un assegno di ricerca da maggio 2015 a ottobre 2016 presso l'Università degli Studi di Modena e Reggio Emilia. Durante questo periodo ha lavorato sul progetto "Open linked data dell'Osservatorio Giovani della Regione Emilia-Romagna" finanziato dal comune di Modena e dalla regione Emilia Romagna ([LINK](#)), integrando dati open e proprietari provenienti da sorgenti eterogenee utilizzando ed estendendo la piattaforma di integrazione dati MOMIS [[c1](#), [j1](#), [c3](#)]. Nel 2020 ha ottenuto un dottorato in Computer Science dalla scuola di dottorato in Information and Communication Technologies dell'Università di Modena e Reggio Emilia, difendendo la tesi intitolata "Techniques for Big Data Integration in Distributed Computing Environments". Il suo lavoro è stato focalizzato sul rendere scalabili tecniche di Entity Resolution sui Big Data in ambienti distribuiti [[j3](#), [j5](#), [c7](#), [c8](#), [c9](#), [c10](#)]. Durante questo periodo, ha anche lavorato su progetti secondari, come lo sviluppo del registro italiano per i malati di distrofia muscolare, sviluppato in collaborazione con DataRiver srl e il Dipartimento and the Dipartimento di Scienze Biomediche, Metaboliche e Neuroscienze dell'Università degli Studi di Modena e Reggio Emilia [[c2](#), [j8](#)].

Dopo il conseguimento del titolo di dottore di ricerca ha svolto un periodo di postdoc presso l'Università degli Studi di Modena e Reggio Emilia, ricoprendo diverse posizioni da assegnista di ricerca, lavorando su progetti multidisciplinari che necessitavano di integrazione dati:

- "LigAdvisor: a webservice for drug design based on big data" ([LINK](#)), finanziato dal FAR 2019 e sviluppato in collaborazione al Dipartimento di Scienze della Vita, prevedeva la progettazione e lo sviluppo di una piattaforma web per la scoperta di nuovi farmaci [[j6](#)].
- "Smart monitoring of Local Energy Community" finanziato dal MISE e sviluppato in collaborazione con ENEA, prevedeva la progettazione di una piattaforma per la gestione dei dati generati dalle comunità energetiche locali [[j9](#)].

- “Digital Experience Platform” finanziato dalla regione Emilia Romagna e sviluppato in collaborazione a Doxee spa, prevedeva la progettazione di una piattaforma per la gestione dei dati di fatturazione di bollette di luce e gas ([LINK](#)).
- “BD4M: Big Data 4 Manufacturing”, WP leader per il “WP3 - Design and Implementation of a Big Data analysis platform”. Progetto di ricerca industriale sviluppato in collaborazione con DataRiver srl e finanziato da Bi.Rex, prevedeva lo sviluppo di una piattaforma in ambito Industria 4.0 per raccogliere e analizzare grandi quantità di dati eterogenei raccolti da diversi macchinari industriali.

Durante il periodo di postdoc ha sviluppato collaborazioni internazionali con il professor Themis Palpanas dell’Università Paris Descartes di Parigi e con George Papadakis, PhD, dell’Università di Atene, pubblicando diversi lavori su Entity Resolution [[j10](#), [j4](#), [j7](#), [c13](#), [c16](#)]. Il lavoro principale [[c13](#), [j10](#)] che propone una nuova tecnica di meta-blocking è stato pubblicato e presentato presso VLDB 2022.

Da gennaio 2022 fino a ottobre 2023 ha lavorato come ricercatore a tempo determinato (RTDA, DM 1062) sul progetto “Big Data e Intelligenza Artificiale per il Virtuosismo Energetico nelle Comunità Energetiche Locali” sviluppando tecniche di integrazione scalabili per la gestione di big data provenienti da comunità energetiche locali [[c12](#), [j9](#)]. Durante questo periodo, ha ricevuto una sovvenzione di 9200 € dall’Università di Modena e Reggio Emilia (FAR) per condurre le proprie attività di ricerca in ambito Big Data integration [[j10](#)], inoltre è stato coordinatore scientifico dell’attività di ricerca/consulenza “Progettazione e sviluppo dei microservizi per Record Linkage e Entity Matching per la piattaforma MOMIS” tra DataRiver srl e il Dipartimento di Ingegneria “Enzo Ferrari” dell’Università di Modena e Reggio Emilia, del valore di 5000 €.

Da Novembre 2023 è ricercatore a tempo determinato (RTDA, PNRR) sul WP5 “Digital Maktaba” del progetto PNRR “ITSERR - Italian Strengthening of the ESFRI RI RESILIENCE”, che ha lo scopo di costruire una libreria digitale per l’archiviazione e indicizzazione di documenti scritti in lingue non latine (es. Arabo). In questo contesto si vorrebbero sfruttare gli LLM per svolgere operazioni come la classificazione dei documenti, l’identificazione di duplicati (anche in lingue diverse), l’estrazione dei topic, etc. Siccome uno dei problemi dell’uso degli LLM è legato al costo di utilizzo, come piano di ricerca si propone lo sviluppo di tecniche di Entity Resolution che sfruttano gli LLM in modalità pay-as-you-go, per ottenere i risultati migliori avendo un budget predeterminato.

Sommaro			
Inizio	Fine	Istituto	Posizione
01/11/2023	Presente	Università degli Studi di Modena e Reggio Emilia	Ricercatore (RTDA - PNRR) DIEF presso DBGroup
01/01/2022	31/10/2023	Università degli Studi di Modena e Reggio Emilia	Ricercatore (RTDA - DM1062) DIEF presso DBGroup
16/07/2021	31/12/2021	Università degli Studi di Modena e Reggio Emilia	Postdoc (assegnista di ricerca senior) DIEF presso DBGroup
01/12/2020	30/06/2021	Università degli Studi di Modena e Reggio Emilia	Postdoc (assegnista di ricerca) DIEF presso DBGroup
01/12/2019	30/11/2020	Università degli Studi di Modena e Reggio Emilia	Postdoc (assegnista di ricerca) DSV
01/11/2016	31/10/2019	Università degli Studi di Modena e Reggio Emilia	Dottorando DIEF presso DBGroup
16/05/2015	31/10/2016	Università degli Studi di Modena e Reggio Emilia	Assegnista di ricerca DIEF presso DBGroup

11/2023-Presente
RTDA presso
UNIMORE

Ricercatore a tempo determinato di tipo A, PNRR, presso il DBGroup dell'Università degli Studi di Modena e Reggio Emilia.

Attività di ricerca svolta: l'attività di ricerca è condotta all'interno del Work Package 5 "Digital Maktaba" del progetto PNRR "ITSERR - Italian Strengthening of the ESFRI RI RESILIENCE" (Avviso MUR 3264/2022) funded by EU - NextGenerationEU - Grant No IR0000014. L'obiettivo è sviluppare tecniche per acquisire, indicizzare, categorizzare, e integrare documenti scritti in lingue non latine (principalmente in Arabo), con il fine di costruire una libreria digitale in grado di integrare dati provenienti da diverse librerie.

Supervisore: Prof.ssa Sonia BERGAMASCHI · Università degli Studi di Modena e Reggio Emilia

01/2022-10/2023
RTDA presso
UNIMORE

Ricercatore a tempo determinato di tipo A, DM 1062, presso il DBGroup dell'Università degli Studi di Modena e Reggio Emilia.

Attività di ricerca svolta: l'attività di ricerca è stata svolta all'interno del progetto "Big Data e Intelligenza Artificiale per il Virtuosismo Energetico nelle Comunità Energetiche Locali", finanziato dal programma "Fondo Sociale Europeo REACT-EU - PON "Ricerca e Innovazione" 2014 - 2020 - Azione IV.6 Contratti di ricerca su tematiche Green (D.M. 1062 del 10/08/2021) CUP: E95F21002330001". Il progetto prevedeva la realizzazione di una piattaforma per la gestione e l'integrazione di Big Data generati dalle comunità energetiche, come ad esempio informazioni sulla produzione di energia da sistemi fotovoltaici, consumi energetici, informazioni meteo, etc.

Supervisore: Prof.ssa Sonia BERGAMASCHI · Università degli Studi di Modena e Reggio Emilia

07/2021-12/2021
Assegnista presso
UNIMORE

Assegnista di ricerca presso il DBGroup dell'Università degli Studi di Modena e Reggio Emilia.

Attività di ricerca svolta: modellazione di nuovi algoritmi di data integration in contesto Big Data. In particolare, la ricerca è focalizzata sullo studio di tecniche scalabili di Entity Resolution (ossia identificare e fondere diverse rappresentazioni dello stesso oggetto del mondo reale) su architetture e framework paralleli, come ad esempio Apache Spark.

Supervisore: Prof.ssa Sonia BERGAMASCHI · Università degli Studi di Modena e Reggio Emilia

12/2020-06/2021
Assegnista presso
UNIMORE

Assegnista di ricerca presso il DBGroup dell'Università degli Studi di Modena e Reggio Emilia.
Attività di ricerca svolta: sviluppo di architetture software per l'integrazione e l'analisi di Big Data all'interno del progetto "SUPER: Supercomputing Unified Platform - EmiliaRomagna", finanziato dalla Regione Emilia Romagna.

Supervisore: Prof.ssa Sonia BERGAMASCHI · Università degli Studi di Modena e Reggio Emilia

12/2019-12/2020
Assegnista presso
UNIMORE

Assegnista di ricerca presso il DBGroup dell'Università degli Studi di Modena e Reggio Emilia.
Attività di ricerca svolta: sviluppo di una piattaforma per aiutare i ricercatori ad individuare nuove molecole in ambito farmaceutico all'interno del progetto di ricerca "LigAdvisor: a webserver for drug design based on Big Data" finanziato dal progetto interdisciplinare FAR 2019 tra il Dipartimento di Scienze Della Vita e il Dipartimento di Ingegneria "Enzo Ferrari".

Supervisore: Prof. Domenico BENEVENTANO · Università degli Studi di Modena e Reggio Emilia

11/2016-11/2019
Dottorando presso
UNIMORE

Dottorando presso il DBGroup dell'Università degli Studi di Modena e Reggio Emilia.
Attività di ricerca: sviluppo di tecniche di integrazione dati distribuite che consentono di scalare su grandi volumi di dati. In particolare, applicazione di tecniche di blocking/indexing e join di similarità su grandi quantità di dati eterogenei.

Supervisore: Prof.ssa Sonia BERGAMASCHI · Università degli Studi di Modena e Reggio Emilia
Co-Supervisore: PhD Giovanni SIMONINI · Università degli Studi di Modena e Reggio Emilia

05/2015-11/2016
Assegnista di
ricerca presso
UNIMORE

Assegnista di ricerca presso il DBGroup dell'Università degli Studi di Modena e Reggio Emilia.
Attività di ricerca: integrazione dati semantica per il progetto di ricerca "Open Linked Data dell'Osservatorio Giovani della Regione Emilia Romagna" ed estensione del sistema MOMIS al fine di realizzare dashboard interattive per la visualizzazione dei dati integrati, finanziato dal comune di Modena.

Supervisore: Prof.ssa Sonia BERGAMASCHI · Università degli Studi di Modena e Reggio Emilia

FORMAZIONE

03/2020
Dottorato presso
UNIMORE

UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA — ICT International Doctoral School

Tesi: *Techniques for Big Data Integration in Distributed Computing Environments.*

Tutor scientifico: Prof.ssa Sonia BERGAMASCHI · Università degli Studi di Modena e Reggio Emilia

Co-Tutor scientifico: PhD. Giovanni SIMONINI · Università degli Studi di Modena e Reggio Emilia

04/2015
Laurea magistrale
in Ingegneria
Informatica

UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

Voto: 110/110

10/2011
Laurea triennale in
Ingegneria
Informatica

UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

Voto: 93/110

QUALIFICHE

17/12/2023 –
17/12/2034

ABILITAZIONE SCIENTIFICA NAZIONALE ALLE FUNZIONI DI PROFESSORE UNIVERSITARIO DI SECONDA FASCIA
Settore Concorsuale: 09/H1 - Sistemi di Elaborazione delle Informazioni
Rilasciato da: MUR – Ministero dell'Università e della Ricerca

2015

ABILITAZIONE ALLA PROFESSIONE DI INGEGNERE - SEZIONE A - INGEGNERIA DELL'INFORMAZIONE
Rilasciato da: *Università degli Studi di Modena e Reggio Emilia*

ATTIVITÀ DI DOCENZA

Ha iniziato la sua attività di docenza durante il periodo di dottorato come cultore della materia per l'insegnamento "Basi di dati" del corso di Laurea Triennale in Ingegneria Informatica del Dipartimento di Ingegneria "Enzo Ferrari" dell'Università degli Studi di Modena e Reggio Emilia (UNIMORE), svolgendo esercitazioni in aula sulla progettazione ER, logica e linguaggio SQL. Successivamente, ha ricoperto lo stesso ruolo per l'insegnamento "Big Data Management and Governance" del corso di Laurea Magistrale in Ingegneria Informatica del Dipartimento di Ingegneria "Enzo Ferrari" di UNIMORE, svolgendo lezioni teoriche e pratiche su SQL avanzato (stored procedures, triggers, transazioni) e Apache Spark. Nell'anno accademico 2018/2019 è stato docente a contratto per l'insegnamento "Sistemi informativi e basi di dati L-Z", insegnando il linguaggio SQL attraverso lezioni pratiche agli studenti del corso di Laurea Triennale in Ingegneria Gestionale del Dipartimento di Scienze e Metodi dell'Ingegneria di UNIMORE presso la sede di Reggio Emilia. Nello stesso anno, ha tenuto un seminario su Apache Spark (teoria e pratica) all'interno del corso di alta formazione "Corso di perfezionamento Metodologie, Tecniche e Tool per Analisi di Big Data" [[LINK](#)]. Nel 2020, insieme al prof. Paolo Missier dell'Università di Newcastle, ha tenuto presso la scuola di dottorato in ICT di UNIMORE, il corso "Data Analytics" in cui sono state presentate le tecnologie abilitanti per elaborare Big Data, come il paradigma MapReduce, HDFS e l'uso pratico di Apache Spark.

A partire dall'anno accademico 2022/2023 è docente di riferimento dell'insegnamento "Basi di dati" del corso triennale in Ingegneria Informatica del Dipartimento di Ingegneria "Enzo Ferrari" di UNIMORE - distaccamento di Mantova, e dell'insegnamento "Business Intelligence" del corso di Laurea Magistrale in International Management presso il Dipartimento di Economia "Marco Biagi" di UNIMORE con sede a Modena. Il primo insegnamento è erogato in lingua italiana e fornisce le conoscenze per progettare e implementare una base di dati tramite lezioni teoriche e pratiche. Mentre il secondo è erogato in lingua inglese, in quanto molti studenti sono stranieri, e fornisce un approccio pratico sulle metodologie e tecnologie utilizzate in una tipica pipeline di Business Intelligence. Gli studenti imparano a gestire, preparare, e integrare dati al fine di costruire dashboard interattive e eseguire alcune operazioni base di data mining.

Durante le sue attività di ricerca, ha supervisionato diversi studenti dei corsi triennali e magistrali in Ingegneria Informatica come riportato successivamente.

Anno accademico
2023/2024
UNIMORE

Docente · 72 ore di lezioni frontali per il corso Basi di Dati, presso Dipartimento di Ingegneria Enzo Ferrari, corso di laurea in Ingegneria Informatica - Sede di Mantova.
Lingua erogazione: Italiano. **Studenti iscritti:** 43.

Docente · 38 ore di lezioni frontali per il corso Business Intelligence, presso il Dipartimento di Economia Marco Biagi, corso di laurea magistrale in Management Internazionale.

Lingua erogazione: Inglese. **Studenti iscritti:** 80.

Anno accademico
2022/2023
UNIMORE

Docente · 60 ore di lezioni frontali per il corso Basi di Dati, presso Dipartimento di Ingegneria Enzo Ferrari, corso di laurea in Ingegneria Informatica - Sede di Mantova.

Lingua erogazione: Italiano. **Studenti iscritti:** 40.

Docente · 38 ore di lezioni frontali per il corso Business Intelligence, presso il Dipartimento di Economia Marco Biagi, corso di laurea magistrale in Management Internazionale.

Lingua erogazione: Inglese. **Studenti iscritti:** 80.

Anno accademico
2021/2022
UNIMORE

Docente · 36 ore di lezioni frontali per il corso Basi di Dati, presso Dipartimento di Ingegneria Enzo Ferrari, corso di laurea in Ingegneria Informatica - Sede di Mantova.

Lingua erogazione: Italiano. **Studenti iscritti:** 38.

Anno accademico
2020/2021
UNIMORE

Docente a contratto · 32 ore di lezioni frontali per il corso Basi di Dati, presso Dipartimento di Ingegneria Enzo Ferrari, corso di laurea in Ingegneria Informatica - Sede di Mantova.

Lingua erogazione: Italiano. **Studenti iscritti:** 50.

Docente di corso di dottorato · Docente per un modulo per la scuola di dottorato in ICT dell'Università degli Studi di Modena e Reggio Emilia, nome del corso: "Data Analytics". – [[LINK](#)].

Lingua erogazione: Inglese. **Studenti iscritti:** 10.

Anno accademico
2019/2020
UNIMORE

Docente a contratto · 32 ore di lezioni frontali per il corso Basi di Dati, presso Dipartimento di Ingegneria Enzo Ferrari, corso di laurea in Ingegneria Informatica - Sede di Mantova.

Anno accademico
2018/2019
UNIMORE

Docente a contratto · 27 ore di didattica frontale per il corso Sistemi informativi e basi di dati L-Z, presso Dipartimento di Scienze e Metodi dell'Ingegneria, corso di laurea in Ingegneria Gestionale.

Lingua erogazione: Italiano. **Studenti iscritti:** 80.

Docente a contratto · 4 ore di lezioni frontali per il "Corso di perfezionamento Metodologie, Tecniche e Tool per Analisi di Big Data".

Lingua erogazione: Italiano. **Studenti iscritti:** 10.

SUPERVISIONE DI STUDENTI

PhD Students

(Co-tutor) Giovanni Sullutrone. "Exploring the Potential of Large Language Models for Multilingual Historical Document Analysis and Semantic Cataloguing in Digital Libraries"
XXXIX Ciclo [[LINK](#)].

(Co-tutor) Luca Sala. "Data Management, analytics and intelligent AI-based knowledge extraction for multilingual and multi-alphabetic heritages"
XXXVIII Ciclo [[LINK](#)].

(Co-tutor) Sania Aftar. "Data Management, analytics and intelligent AI-based knowledge extraction for multilingual and multi-alphabetic heritages"
XXXVIII Ciclo [[LINK](#)].

*Laurea magistrale
in Ingegneria
Informatica*

(Co-tutor) Marco Esposito. "Gestione e analisi di Big Data conformi al GDPR: il progetto del database Digital eXperience Platform"
2022 UNIMORE.

*Laurea triennale in
Ingegneria
Informatica*

(Tutor) Manuel Olmi. "Full-stack project "Gestione Esami""
2023 UNIMORE.

(Co-tutor) Michele Rinaldi. "Entity Resolution progressiva con graph embedding"
2022 UNIMORE.

(Co-tutor) Arturo Bianchi. "Documentazione e aggiunte alla libreria per l'Entity Resolution SparkER"
2022 UNIMORE.

(Co-tutor) Davide Bucciarelli. "SIARD Format: analisi di un nuovo modo di preservare database"
2021 UNIMORE.

(Co-tutor) Davide Abba. "Full-stack project "Gastronomia La Cilba" "
2021 UNIMORE.

(Co-tutor) Mario Cristiano. "Relational Database Management System: Analisi dell'ambiente MariaDB"
2021 UNIMORE.

(Co-tutor) Davide Corradi. "Full-Text Search e applicazioni: Confronto tra Microsoft SQL Server ed Elasticsearch"
2021 UNIMORE.

(Co-tutor) Francesca Prontera. "Scalabilità di Tecniche di Analisi di Dati con la libreria Modin"
2021 UNIMORE.

(Co-tutor) Luisa Bottiglieri. "TimescaleDB: analisi di time-series di consumi energetici"
2021 UNIMORE.

(Co-tutor) Giovanni Bonisoli. "Sviluppo di un'applicazione web per l'utilizzo del framework SparkER"
2017 UNIMORE.

PARTECIPAZIONE E RESPONSABILITÀ IN PROGETTI DI RICERCA

4/2023–11/2023

Progettazione e sviluppo dei microservizi per Record Linkage e Entity Matching per la piattaforma MOMIS

Responsabile scientifico del contratto di ricerca/consulenza tra DataRiver srl e il dipartimento di Ingegneria "Enzo Ferrari" dell'Università degli Studi di Modena e Reggio Emilia per l'attività "Progettazione e sviluppo dei microservizi per Record Linkage e Entity Matching per la piattaforma MOMIS".

10/2020–01/2022

Digital Experience Platform

Progetto finanziato dalla Regione Emilia Romagna sviluppato in collaborazione con l'azienda Doxee. Lo scopo del progetto è sviluppare una piattaforma in grado di gestire grandi quantità di dati provenienti da bollette di gas e luce al fine di eseguire analisi dati per migliorare l'esperienza degli utenti. In particolare, si vogliono utilizzare tecniche di machine learning per ridurre i costi (es. suggerire agli utenti un'offerta migliore) ed evitare l'abbandono degli utenti.

10/2020–04/2021

Smart monitoring of Local Energy Community

Il progetto è stato finanziato e si è svolto in collaborazione con ENEA. L'obiettivo era progettare una piattaforma in grado di raccogliere e analizzare grandi quantità di dati provenienti da sensori per

monitorare la rete energetica al fine di consentire un utilizzo coscienzioso dell'energia da parte degli utilizzatori finali. Per raggiungere questo obiettivo, si è eseguita un'analisi approfondita di diverse tecnologie per la raccolta e la memorizzazione di grandi quantità di dati. Infine, è stata progettata una piattaforma per la gestione di Big Data combinando diversi strumenti (Apache Spark, Delta Lake, TimescaleDB, Airflow) per poter gestire tutto il ciclo di vita dei dati.

7/2020–01/2022

BD4M: Big Data 4 Manufacturing

Work Package leader del "WP3 - Design e Implementazione di una piattaforma per Big Data analysis". Il progetto di ricerca industriale è stato svolto in collaborazione con DataRiver srl e finanziato da Bi.Rex "Area 2 – Big Data per il Manufacturing" di cui in attuazione dell'art. 1, comma 115, Legge 11 dicembre 2016 n. 232 e in conformità e nel rispetto dell'art. 4, D.Lgs. 19 agosto 2016 n. 175, del D.M. Sviluppo Economico 12 settembre 2017 n. 214, del Decreto Direttoriale del Ministero dello Sviluppo Economico 29 gennaio 2018. Lo scopo del progetto era lo sviluppo di una piattaforma Big Data in ambito Industria 4.0 per la raccolta e analisi di dati eterogenei raccolti da diversi macchinari industriali.

VISITE & COLLABORAZIONI

01/03/2021–
01/08/2021
Paris Descartes
University

Tirocinio svolto sotto la supervisione del Prof. Themis Palpanas presso Descartes University, Paris. Durante questo periodo ho lavorato su di un progetto di **Big Data Cleaning** che ha avuto come risultato un articolo che è stato pubblicato al convegno internazionale VLDB 2022 [c13].

Nota: Il tirocinio è stato svolto da remoto a causa delle restrizioni imposte dalla pandemia di Covid-19.

RICONOSCIMENTI E SOVVENZIONI

Riconoscimenti

2017 · Young Professional event @ RTSI2017 · (secondo posto).

Sovvenzioni

2022 · FAR dipartimentale - Starter Kit per ricercatori del valore di 9000€

2019 · Christine Collet EDBT/ICDT Student Participation Award · (Registrazione e pranzi gratuiti alla conferenza EDBT)

2016 · Borsa di dottorato della durata di tre anni sovvenzionata dalla Regione Emilia Romagna.

PRESENTAZIONI

Giugno 2023 · SEBD, Galzignano Terme, Italia: "How To Generalize Supervised Meta-blocking"

Settembre 2022 · VLDB, Sidney, Australia (da remoto a causa dalla pandemia di Covid-19): "Generalized Supervised Meta-blocking"

Marzo 2022 · DataPlat, Edinburgh, UK (da remoto a causa dalla pandemia di Covid-19): "ECDP: A Big Data Platform for the Smart Monitoring of Local Energy Communities"

Giugno 2020 · SEBD, Villasimius, Italia (da remoto a causa dalla pandemia di Covid-19): "Scaling Up Record-level Matching Rules"

Marzo 2020 · EDBT, Copenhagen, Danimarca (da remoto a causa dalla pandemia di Covid-19): "RulER: Scaling Up Record-level Matching Rules"

Marzo 2019 · EDBT, Lisbona, Portogallo: "Sparker: Scaling entity resolution in spark"

Luglio 2018 · TE2018, Modena, Italia: "BigDedup: a big data integration toolkit for duplicate detection in industrial scenarios"

SERVIZIO PER LA COMUNITÀ SCIENTIFICA

*Membro di
comitato di
programma*

- 33rd ACM International Conference on Information and Knowledge Management (CIKM 2024) - Resource Track.
- 25th International Conference on Information Integration and Web Intelligence (iiWAS 2024).
- 25th International Conference on Big Data Analytics and Knowledge Discovery (DaWaK 2024).
- 25th International Conference on Information Integration and Web Intelligence (iiWAS 2023).
- 25th International Conference on Big Data Analytics and Knowledge Discovery (DaWaK 2023).
- 24th International Conference on Information Integration and Web Intelligence (iiWAS 2022).
- 24th International Conference on Big Data Analytics and Knowledge Discovery (DaWaK 2022).
- 23rd International Conference on Big Data Analytics and Knowledge Discovery (DaWaK 2021).

Revisore

- Elsevier Information Systems.
- Data Mining & Knowledge Discovery.
- Data & Knowledge Engineering.
- IEEE Transactions on Knowledge and Data Engineering (TKDE).
- Knowledge-Based Systems (KNOSIS).
- ACM SIGMOD Reproducibility 2023.

Revisore esterno

- ICDE 2022.
- SIGMOD 2022.
- SIGMOD 2021.
- VLDB 2020.
- EDBT 2020.
- ICDE 2018.

ESPERIENZA LAVORATIVA

*2015–Presente
Collaboratore
presso DataRiver*

Consulente per progetti di integrazione dati.
<http://datariver.it/chi-siamo/team/>

*2/2019– 3/2019
Sviluppatore full
stack*

Sviluppo di una mappa interattiva per il comune di Modena, titolo del progetto "Mappa multimediale di Modena, le architetture del Novecento". [Sito web]

COMPETENZE INFORMATICHE

*Linguaggi di
programmazione*

SQL; PYTHON; SCALA; JAVA; JAVASCRIPT; SHELL

*Big Data
Framework*

SPARK (SPARK, SPARKSQL, MLlib); DELTALAKE

Altri Framework

DOCKER; IPYTHON; ANGULAR; SCIKIT-LEARN; SCRAPY; MONGODB; SQL SERVER; MYSQL; POSTGRESQL; L^AT_EX;

ALTRE INFORMAZIONI

Lingue

ITALIANO · Madrelingua
INGLESE · Conoscenza professionale

PUBBLICAZIONI

Metriche (a Luglio 2024)		
	Google Scholar	Scopus
h-index	10	8
Numero citazioni	318	217

Elenco di pubblicazioni suddiviso per categoria e ordinate per data di pubblicazione

[c] = conferenza [j] = rivista [bc] = capitolo di libro [b] = libro [t] = tesi
[sw] = software

Riviste

- [j11] **Stream-aware indexing for distributed inequality join processing** - Information Systems, 2024. - A. ASLAM, G. SIMONINI, L. GAGLIARDELLI, L. ZECCHINI, S. BERGAMASCHI [PDF]
- [j10] **GSM: A generalized approach to Supervised Meta-blocking for scalable entity resolution** - Information Systems, 2024. - L. GAGLIARDELLI, G. PAPADAKIS, G. SIMONINI, S. BERGAMASCHI, T. PALPANAS. [PDF]
- [j9] **A big data platform exploiting auditable tokenization to promote good practices inside local energy communities** - Future Generation Computer Systems, 2023 - L. GAGLIARDELLI, L. ZECCHINI, L. FERRETTI, D. BENEVENTANO, G. SIMONINI, S. BERGAMASCHI, M. ORSINI, L. MAGNOTTA, E. MESCOLI, A. LIVALDI, N. GESSA, P. DE SABBATA, G. D'AGOSTA, F. PAOLUCCI, F. MORETTI. [PDF]
- [j8] **The Italian National Registry for FSHD: an enhanced data integration and an analytics framework towards Smart Health Care and Precision Medicine for a rare disease** - Orphanet Journal of Rare Diseases, 2021 - C. BETTIO, V. SALSÌ, M. ORSINI, E. CALANCI, L. MAGNOTTA, L. GAGLIARDELLI, J. KINOSHITA, S. BERGAMASCHI, R. TUPLER. [PDF]
- [j7] **Reproducible experiments on Three-Dimensional Entity Resolution with JedAI** - Information Systems, 2021 - G. PAPADAKIS, G. MANDILARAS, L. GAGLIARDELLI, G. SIMONINI, E. THANOS, G. GIANNAKOPOULOS, S. BERGAMASCHI, T. PALPANAS, M. KOUBARAKIS. [PDF]
- [j6] **LigAdvisor: a versatile and user-friendly web-platform for drug design** - Nucleic Acids Research (NAR), 2021 - L. PINZI, A. TINIVELLA, L. GAGLIARDELLI, D. BENEVENTANO, G. RASTELLI. [PDF]
- [j5] **BLAST2 An Efficient Technique for Loose Schema Information Extraction from Heterogeneous Big Data Sources** - Journal of Data and Information Quality (JDIQ), 2020 - D. BENEVENTANO, S. BERGAMASCHI, L. GAGLIARDELLI, G. SIMONINI. [PDF]
- [j4] **Three-dimensional Entity Resolution with JedAI** - Information Systems, 2020 - G. PAPADAKIS, G. MANDILARAS, L. GAGLIARDELLI, G. SIMONINI, E. THANOS, G. GIANNAKOPOULOS, S. BERGAMASCHI, T. PALPANAS, M. KOUBARAKIS. [PDF]
- [j3] **Scaling entity resolution: A loosely schema-aware approach** - Information Systems, 2019 - G. SIMONINI, L. GAGLIARDELLI, S. BERGAMASCHI, HV Jagadish. [PDF] [CODE]
- [j2] **BigBench workload executed by using Apache Flink** - Procedia Manufacturing 11 (2017): 695-702. - S. BERGAMASCHI, L. GAGLIARDELLI, G. SIMONINI, S. ZHU. [PDF]
- [j1] **Driving Innovation in Youth Policies With OpenData** - IC3K 2015. Communications in Computer and Information Science, 2016 - D. BENEVENTANO, S. BERGAMASCHI, L. GAGLIARDELLI, L. Po. [PDF]

Conferenze

- [c19] **A Novel Methodology for Topic Identification in Hadith**, IRCDL 2024. - S. AFTAR, L. GAGLIARDELLI, A. EL GANADI, F. RUOZZI, S. BERGAMASCHI.
- [c18] **A Big Data Platform for the Management of Local Energy Communities Data**, AIPREF @ IEEE BigData 2023. - S. BERGAMASCHI, L. GAGLIARDELLI.
- [c17] **HKS: Efficient Data Partitioning for Stateful Streaming**, DAWAK 2023. - A. ASLAM, G. SIMONINI, L. GAGLIARDELLI, A. MOZZILLO, S. BERGAMASCHI.
- [c16] **A general approach to Supervised Meta-Blocking**, SEBD 2023. - L. GAGLIARDELLI, G. PAPADAKIS, G. SIMONINI, S. BERGAMASCHI, T. PALPANAS.
- [c15] **Big Data Integration for Data-Centric AI**, ItaData 2022. - S. BERGAMASCHI, D. BENEVENTANO, G. SIMONINI, L. GAGLIARDELLI, A. ASLAM, G. DE SABBATA, L. ZECCHINI.
- [c14] **Progressive Entity Resolution with Node Embeddings**, SEBD 2022. - G. SIMONINI, L. GAGLIARDELLI, M. RINALDI, L. ZECCHINI, G. DE SABBATA, A. ASLAM, D. BENEVENTANO, S. BERGAMASCHI. [PDF]
- [c13] **Generalized Supervised Meta-blocking** - 48th International Conference on Very Large Databases, VLDB 2022. - L. GAGLIARDELLI, G. PAPADAKIS, G. SIMONINI, S. BERGAMASCHI, T. PALPANAS. [PDF]
- [c12] **ECDP: A Big Data Platform for the Smart Monitoring of Local Energy Communities** - 1st International Workshop on Data Platform Design, Management, and Optimization, DataPlat 2022. - L. GAGLIARDELLI, L. ZECCHINI, D. BENEVENTANO, G. SIMONINI, S. BERGAMASCHI, M. ORSINI, L. MAGNOTTA, E. MESCOI, A. LIVALDI, N. GESSA, P. DE SABBATA, G. D'AGOSTA, F. PAOLUCCI, F. MORETTI. [PDF]
- [c11] **The Case for Multi-task Active Learning Entity Resolution** - 29th Italian Symposium on Advanced Database Systems, SEBD 2021. - G. SIMONINI, H. SACCANI, L. GAGLIARDELLI, L. ZECCHINI, D. BENEVENTANO, S. BERGAMASCHI. [PDF]
- [c10] **Scaling Up Record-level Matching Rules** - 28th Italian Symposium on Advanced Database Systems, SEBD 2020. - L. GAGLIARDELLI, G. SIMONINI, S. BERGAMASCHI. [PDF]
- [c9] **RuLER: Scaling Up Record-level Matching Rules** - 23rd International Conference on Extending Database Technology, EDBT 2020. - L. GAGLIARDELLI, G. SIMONINI, S. BERGAMASCHI. [PDF]
- [c8] **Entity resolution and data fusion: An integrated approach** - 27th Italian Symposium on Advanced Database Systems, SEBD 2019. - D. BENEVENTANO, S. BERGAMASCHI, L. GAGLIARDELLI, G. SIMONINI. [PDF]
- [c7] **Sparker: Scaling entity resolution in Spark** - 22nd International Conference on Extending Database Technology, EDBT 2019. - L. GAGLIARDELLI, G. SIMONINI, D. BENEVENTANO, S. BERGAMASCHI. [PDF] [CODE]
- [c6] **How improve Set Similarity Join based on prefix approach in distributed environment.** - IEEE International Conf. on High Performance Computing & Simulation, HPCS 2018. - S. ZHU, L. GAGLIARDELLI, D. BENEVENTANO, G. SIMONINI. [PDF]
- [c5] **Enhancing Loosely Schema-aware Entity Resolution with User Interaction.** - IEEE International Conf. on High Performance Computing & Simulation, HPCS 2018. - G. SIMONINI, L. GAGLIARDELLI, S. ZHU, S. BERGAMASCHI. [PDF]
- [c4] **Bigdedup: a big data integration toolkit for duplicate detection in industrial scenarios** - IEEE 25th International Conference on Transdisciplinary Engineering, TE 2018. - L. GAGLIARDELLI, S. ZHU, G. SIMONINI, S. BERGAMASCHI. [PDF]
- [c3] **Momis dashboard: A powerful data analytics tool for industry 4.0** - IEEE 25th International Conference on Transdisciplinary Engineering, TE 2018. - L. MAGNOTTA, L. GAGLIARDELLI, G. SIMONINI, M. ORSINI, S. BERGAMASCHI. [PDF]

- [c2] **The Italian FSHD Registry: an enhanced data integration and analytics framework for smart health care** - IEEE 3rd International Forum on Research and Technologies for Society and Industry, RTSI 2017. - M. ORSINI, E. CALANCI, L. MAGNOTTA, L. GAGLIARDELLI, M. GOVI, F. MELE, R. TUPLER. [PDF]
- [c1] **Open Data for Improving Youth Policies** - International Joint Conference on Knowledge Discovery, Knowledge Engineering, IC3K 2015 - D. BENEVENTANO, S. BERGAMASCHI, L. GAGLIARDELLI, L. PO. [PDF]

Capitoli di libri

- [bc1] **From Data Integration to Big Data Integration.** - "A Comprehensive Guide Through the Italian Database Research 2018", Springer. - S. BERGAMASCHI, D. BENEVENTANO, F. MANDREOLI, R. MARTOGLIA, F. GUERRA, M. ORSINI, L. PO, M. VINCINI, G. SIMONINI, S. ZHU, L. GAGLIARDELLI, L. MAGNOTTA. [PDF]

Libri

- [b1] **Progetto di Basi di Dati Relazionali**, Bonomo Editore. D. BENEVENTANO, S. BERGAMASCHI, L. GAGLIARDELLI, F. GUERRA, M. VINCINI. ISBN 9788869723032 [LINK]

Tesi

- [t] **Techniques for Big Data Integration in Distributed Computing Environments** - PhD Thesis (2020) - L. Gagliardelli. [PDF]

Prodotti software

- [sw] **JedAI** - è un framework completo che contiene diversi tool per eseguire integrazione dati. Personalmente mi sono occupato di sviluppare la parte distribuita basata su Apache Spark. [LINK]
- [sw] **RuLER** - è un tool che utilizza una nuova tecnica per trovare record simili applicando regole complesse di join di similarità su uno o più attributi realizzato utilizzando il framework Apache Spark. [LINK]
- [sw] **SparkER** - è un tool che implementa tecniche avanzate di blocking per eseguire Entity Resolution realizzato utilizzando il framework Apache Spark. [LINK]