

Matteo Brucato

SENIOR RESEARCHER · DATA SYSTEMS GROUP · MICROSOFT RESEARCH

Microsoft Research, Redmond WA, USA

✉ mbrucato@microsoft.com | 🌐 <https://matteo-brucato.github.io/> | 📷 [MatteoBrucato](#)

Summary

Senior Researcher at Microsoft Research, in the Data Systems Group.

Ph.D. from the University of Massachusetts Amherst. Thesis on systems for scalable and easy prescriptive analytics. Thesis title: “Package Queries: Enabling Declarative and Scalable Prescriptive Analytics in Relational Data”.

More than 8 years of experience in research, with several publications in top-tier conferences and journals in data and information management (SIGMOD, VLDB, CACM, VLDBJ). Multidisciplinary research, with publications in top NLP, IR and robotics/AI conferences. Work recognized by several ACM awards (CACM, SIGMOD Highlights, invited VLDBJ, best demos and best paper at VLDB, best young research nominee in RANLP), and several scholarships (Krithi Ramamritham Computer Science Scholarship by UMass; merit, research and exchange scholarships by the University of Bologna).

More than 8 years of mentorship and advising experience. Mentored more than 20 students, at all levels (undergrad, master, and doctoral students), and directly managed several projects.

Several industrial internships, both as researcher (MSR, IBM) and software developer (Cineca, Maggiore Hospital).

Research Interests

Research areas: Data Management, Prescriptive Analytics, Decision Making, Stochastic Constrained Optimization, Fair Machine Learning, Artificial Intelligence, Robotics, Natural Language Processing, Information Retrieval.

Research interests: Theoretical and systemic aspects of constrained optimization, decision making and prescriptive analytics at a very large scale; Structured, unstructured, and natural query languages; Scalable and approximate algorithms; Systems for interactive data exploration; Scalable systems for fair ML; Scalable systems for AI and robotics; Personalized text summarization.

Education

UMASS AMHERST

PH.D. IN COMPUTER SCIENCE

Amherst, MA, USA

2013 - 2021

- Co-advisors: Peter J. Haas, Alexandra Meliou, and Azza Abouzied (NYU)
- Thesis: *Package Queries: Enabling Declarative and Scalable Prescriptive Analytics in Relational Data*

UNIVERSITY OF BOLOGNA

Bologna, Italy

M.S. IN COMPUTER SCIENCE

2011 - 2013

- Summa cum laude
- Thesis Advisor: Danilo Montesi
- Thesis: *Temporal Information Retrieval*

UNIVERSITY OF BOLOGNA

Bologna, Italy

B.S. IN COMPUTER SCIENCE

2008 - 2011

- Summa cum laude
- Thesis Advisor: Danilo Montesi
- Thesis: *Mobile systems for patient-reported outcomes*

EXCHANGE PROGRAMS

ÅRHUS UNIVERSITY

Århus, Denmark

EXCHANGE PROGRAM DURING M.S., COMPUTER SCIENCE

Jan - Jun 2013

- Advisor: Christian S. Jensen

UC RIVERSIDE

Riverside, CA, USA

EXCHANGE PROGRAM DURING B.S., COMPUTER SCIENCE

Sep - Jun 2011

- Advisor: Vassilis J. Tsotras

Honors & Awards

2020	Best Demo Award , VLDB 2020	<i>Tokyo, Japan</i>
2020	Best Demo Runner-up Award , VLDB 2020	<i>Tokyo, Japan</i>
2019	Krithi Ramamritham Computer Science Scholarship	<i>Amherst, MA, USA</i>
2018	CACM Research Highlight Award (Invited journal paper in CACM Research Highlights)	<i>Amherst, MA, USA</i>
2017	ACM SIGMOD Research Highlight Award (invited journal paper in SIGMOD Records)	<i>Amherst, MA, USA</i>
2017	Invited journal paper in the VLDB Journal, Special Issue on Best Papers of VLDB 2016	<i>Amherst, MA, USA</i>

2016 **Best Papers of**, VLDB 2016
 2013 **M.S. with Honors**, University of Bologna
 2013 **Best Young Researcher Award Nominee**, RANLP 2013
 2013 **Merit Scholarship**, awarded to 0.2% of all students of the University of Bologna
 2012 **Research Scholarship**, University of Bologna
 2012 **Merit Certificate**, University of Bologna
 2012 **Complimentary AAAI Membership**, Stanford Online AI Course
 2011 **B.S. with Honors**, University of Bologna
 2011 **Merit Certificate**, University of Bologna
 2010 **Merit Certificate**, University of Bologna
 2010 **Exchange Scholarship**, University of Bologna

New Delhi, India
 Bologna, Italy
 Hissar, Bulgaria
 Bologna, Italy
 Bologna, Italy
 Bologna, Italy
 Stanford, CA, USA
 Bologna, Italy
 Bologna, Italy
 Bologna, Italy
 Bologna, Italy

Publications

1. N. Yadav, **M. Brucato**, A. Fariha, O. Youngquist, J. Killingback, A. Meliou, P. J. Haas
 SUBSUME: A Dataset for Subjective Summary Extraction from Wikipedia Documents
NewSum @ EMNLP 2021 (workshop paper)
2. S. B. Nashed, J. Svegliato, **M. Brucato**, C. Basich, R. Grupen, S. Zilberstein
 Solving Markov Decision Processes with Partial State Abstractions
ICRA 2021
3. **M. Brucato**, M. Mannino, A. Abouzied, P. J. Haas, A. Meliou
 sPaQLTools: A Stochastic Package Query Interface for Scalable Constrained Optimization
VLDB 2020 (demonstration paper) pp.2881-2884
 **Best Demo Award. Video: <https://www.youtube.com/watch?v=1v5KevGKbfE>**
4. A. Fariha, **M. Brucato**, P. J. Haas, A. Meliou
 SuDocu: Summarizing Documents by Example
VLDB 2020 (demonstration paper) pp.2861-2864
 **Best Demo Runner-Up Award. Video: <https://www.youtube.com/watch?v=ZGzF6cBSYjo>**
5. **M. Brucato**, N. Yadav, A. Abouzied, P. J. Haas, A. Meliou
 Stochastic Package Queries in Probabilistic Databases
SIGMOD 2020 pp.269-283
6. **M. Brucato**, A. Abouzied, A. Meliou
 Scalable Computation of High-Order Optimization Queries
Communications of the ACM (CACM) 2019 pp.108-116
7. **M. Brucato**, A. Abouzied, A. Meliou
 Package queries: Efficient and scalable computation of high-order constraints
The VLDB Journal 2018 (Special Issue on Best Papers of VLDB 2016) pp.693-718
8. **M. Brucato**, A. Abouzied, A. Meliou
 A Scalable Execution Engine for Package Queries
SIGMOD Record 2017 (ACM SIGMOD Research Highlight Award 2017) pp.24-31
9. **M. Brucato**, A. Abouzied, C. Blauvelt
 Redistributing Funds across Charitable Crowdfunding Campaigns
arXiv:1706.00070 2017
10. **M. Brucato**, J. F. Beltran, A. Abouzied, A. Meliou
 Scalable Package Queries in Relational Database Systems
VLDB 2016 pp.576-587
 **Best papers of VLDB 2016**
 **Invited journal paper in the VLDB Journal (Special Issue on Best Papers of VLDB 2016)**
 **ACM SIGMOD Research Highlight Award (Invited journal paper in SIGMOD Records 2017)**
 **CACM Research Highlight Award (Invited journal paper in CACM Research Highlight 2018)**
11. **M. Brucato**, A. Abouzied, and A. Meliou
 Improving Package Recommendations through Query Relaxation
Data4U @ VLDB 2014 (workshop paper) pp.13-18
12. **M. Brucato**, R. Ramakrishna, A. Abouzied, A. Meliou
 PackageBuilder: From Tuples to Packages
VLDB 2014 (demonstration paper) pp.1593-1596
13. K. Fernandes, **M. Brucato**, R. Ramakrishna, A. Abouzied, A. Meliou
 PackageBuilder: Querying for packages of Tuples
SIGMOD 2014 (poster paper) pp.1613-1614
14. **M. Brucato**, D. Montesi
 Metric Spaces for Temporal Information Retrieval
ECIR 2014 pp.385-397

15. **M. Brucato**, L. Derczynski, H. Llorens, K. Bontcheva, C. S. Jensen
Recognising and Interpreting Named Temporal Expressions
RANLP 2013 pp.113-121
🌟 **Best Young Researcher Award Nominee**

Professional Experience

INDUSTRY

Microsoft Research

SENIOR RESEARCHER

Redmond, WA, USA

Oct 2021 - present

Microsoft Research

RESEARCH INTERN

Redmond, WA, USA

May - Aug 2017

Worked under the supervision of **Kaushik Chakrabarti** on natural language to SQL using neural networks.

IBM Watson

T. J. Watson Research Center, Yorktown Heights, NY, USA

RESEARCH INTERN, WATSON HEALTH

Jun - Aug 2016

Worked under the supervision of **Shilpa Mahatma** and **Yajuan Wang** on “conceptual queries” for healthcare applications.

CINECA

Casalecchio di Reno, Bologna, Italy

SOFTWARE ENGINEERING INTERN

Jun - Oct 2012

- Responsible for developing and managing a data warehouse and analytics system for a nationwide videoconferencing system.
- Duties included generating aggregated data and analytics, deployed through usable interfaces to let decision makers understand usage and performance of the videoconferencing system.

Maggiore Hospital

Bologna, Italy

SOFTWARE DEVELOPER (VOLUNTEER)

2011 - 2013

- Co-developed software to help the Emergency Room visit their patients using portable ultrasound technology.
- Software deployed on medical tablets to help doctors collect patient's information on the field.
- The collected data was then used to make better informed decisions in the Emergency Room.



ACADEMIA

DREAM Lab, University of Massachusetts

Amherst, MA, USA

RESEARCH ASSISTANT

Sep 2013 - May 2021

- Working with **Prof Peter J. Haas**, **Prof Alexandra Meliou**, and **Prof Azza Abouzied** (NYU Abu Dhabi) to create  **Package Builder**, a system that extends query engines to support the generation of *packages*, i.e., collections of tuples with global properties that are true of the entire collection.
- Technical paper at SIGMOD 2020
- Demo papers at VLDB 2020
- Invited journal paper at CACM Research Highlights 2019.
- Invited journal paper at the VLDB Journal 2018.
- Invited journal paper at SIGMOD Records 2017.
- Technical paper at VLDB 2016; workshop paper and demo paper at VLDB 2014; undergraduate research poster at SIGMOD 2014.
- Project website: packagebuilder.cs.umass.edu.
- DREAM Lab website: dbgroup.cs.umass.edu.
- Project code: Scalable-PaQL-Queries on  GitHub

Design Technology Lab, NYU Abu Dhabi

Abu Dhabi, UAE

VISITING PH.D. CANDIDATE

Mar - May 2016

- Paper: Redistributing Funds across Charitable Crowdfunding Campaigns (arXiv:1706.00070).

AMPLab, UC Berkeley

Berkeley, CA, USA

VISITING PH.D. STUDENT

May - Aug 2015

- Worked on extensions to Package Builder to identify ideal packages by example, under NYU Summer internship.

Data-Intensive Systems Lab, Århus University

Århus, Denmark

VISITING M.S. STUDENT

Jan - Jun 2013

- Worked on M.S. thesis on *Temporal Information Retrieval* under the supervision of **Prof Christian S. Jensen**.
- Actively collaborated with researchers within the group.
- Published two papers in main international conferences in the fields of NLP and IR.

Database Lab, UC Riverside

Riverside, CA, USA

UNDERGRADUATE RESEARCH INTERN (VOLUNTEER)

Jan - Jun 2013

- Worked closely with a Ph.D. candidate and **Prof Vassilis J. Tsotras** on query optimization techniques for query result diversification.
- Developed a query parser for a SQL-like diversification query language.
- Devised and developed a query optimizer to select optimal diversification query plans.
- Developed client/server technology to issue and evaluate diversification queries.

Advising & Mentoring

Research Advisor for Ph.D. students Riddho Ridwanul and Anh Mai

UMASS AMHERST

Riddho (UMass Amherst) and Anh (NYU Abu Dhabi) are first-year Ph.D. students and working on package queries.

Amherst, MA

2021

Research Advisor for M.S. students Oscar Youngquist and Julian killingback

UMASS AMHERST

Oscar is working on enhancing SuDocu, our system for extractive document summarization by example. Julian helped us collecting user-annotated data from Mechanical Turk.

Amherst, MA

2021

Research Advisor for undergraduate student Agam Sandhu

UMASS AMHERST

Agam Sandhu worked on scalable stochastic constrained optimization under costly sampling methods.

Amherst, MA

2020

Research Advisor for PhD student Ahmad Chatha (NYU)

UMASS AMHERST

Ahmad Chatha worked on combining scalable constrained optimization methods with scenario-based methods for stochastic optimization.

Amherst, MA

2020

Project Mentor (in collaboration with Microsoft Research)

UMASS AMHERST

Mentored MS students (Abdul Hannan Kanji, Parul Gupta, Ge Gao, Xinlu Liu) for their Independent Class project (CS 696DS), in collaboration with Microsoft Research, to improve the Fairlearn library (<https://fairlearn.github.io>).

Amherst, MA

2020

Research Advisor for MS student Kanchi Masalia

UMASS AMHERST

Kanchi Masalia worked on parts of SuDocu, our system for personalized efficient document summarization by example.

Amherst, MA

2020

Research Advisor for undergrad Genglin Liu

UMASS AMHERST

Genglin Liu worked on the frontend of SuDocu, our system for personalized efficient document summarization by example.

Amherst, MA

2020

Research Mentor for undergraduate student Kyaw Htet

UMASS AMHERST

Kyaw worked on analyzing scenario reductions methods for stochastic optimization.

Amherst, MA

Fall 2019

Project Mentor (in collaboration with Microsoft)

UMASS AMHERST

Mentored a group of four MS students (Katie House, Shanu Vashishtha, Deeksha Razdan, Cheng-Yin Eng) for their Independent Class project (CS 696DS), in collaboration with Microsoft to produce notebooks that showcase their Machine-as-a-service platform Azure ML.

Amherst, MA

Spring 2019

Research Mentor for undergraduate student Nishad Ranade

UMASS AMHERST

Nishad worked on Monte Carlo scenario reductions methods.

Amherst, MA

2018–2019

Research Mentor for 1st-year PhD student Nishant Yadav

UMASS AMHERST

Nishant and I worked together on stochastic package queries.

Amherst, MA

2018

Research Mentor for graduate student Sandhya Sankaranarayanan

UMASS AMHERST

Sandhya and I worked together on extending the package query language PaQL.

Amherst, MA

Fall 2015

Research Mentor for undergraduate student Kevin Fernandez

UMASS AMHERST

Kevin and I worked together on evaluation techniques for package queries. This work was invited to present at the ACM SIGMOD 2014 Undergraduate Research Poster Competition.

Amherst, MA

Spring 2014

Research Mentor for graduate student Rahul Ramakrishna

UMASS AMHERST

Rahul and I worked together on interface abstractions for package queries. Our work was demonstrated at VLDB 2014.

Amherst, MA

2013 - 2014

Teaching

Teaching Assistant for Graduate Database Systems (CS 645)

UMASS AMHERST

Office hours, question answering, grading, projects.

Amherst, MA

Spring 2021

Teaching Assistant for Undergraduate Database Systems (CS 345)

UMASS AMHERST

Office hours, question answering, grading and labs.

Amherst, MA

Fall 2020

Teaching Assistant for Undergraduate Class “A Networked World” (INFO 203)

UMASS AMHERST

Homework grading.

Amherst, MA

Spring 2018

Teaching Assistant for Undergraduate Introduction To Computation (CS 250)

UMASS AMHERST

Responsible for discussion sessions, occasional lectures, office hours, homework and discussion grading.

Amherst, MA

Fall 2017

Teaching Assistant for Undergraduate Database Systems (CS 345)

UMASS AMHERST

Responsible for occasional lectures, project advising, office hours, homework and project grading.

Amherst, MA

Fall 2015

Professional Service & Leadership

Reviewer

2020

- *IEEE Transactions of Knowledge and Data Engineering (TKDE)*.

Guest Editorial Board

2014 - 2015

- *Information Processing and Management*, special issue on *Time and Information Retrieval*, Elsevier, 2014/2015.

External Review Committee

2013 - 2015

- VLDB 2017
- SIGMOD 2015
- VLDB 2015
- RANLP 2013

Lab Seminar Organizer

Amherst, MA, USA

UMASS AMHERST DATABASE LAB

2014 - 2015

- Coordinated weekly discussion seminar, including invited talks and paper readings, for the UMass Amherst Database Research group (~10 Ph.D. students and faculty), and webmaster duties.

Student Seminar Leader

Amherst, MA, USA

UMASS AMHERST DATABASE LAB

Fall 2015

- Organized and coordinated a weekly discussion seminar among Ph.D. and M.S. students in the UMass Amherst Database Research group to foster research collaborations, exchange of ideas, group discussions, and group cohesion.

Talks

INDUSTRY

Package Queries: Enabling Declarative and Scalable Prescriptive Analytics in Relational Data

Virtually

MEGAGON LABS

Jan 2021

Stochastic Package Queries for In-Database Prescriptive Analytics

Virtually

SISU DATA

Sep 2021

Scalable Package Queries in Relational Database Systems

Redmond, WA, USA

MICROSOFT RESEARCH

Jun 2017

Conceptual Queries for Multiple Data Sources

Austin, TX, USA

IBM WATSON

Aug 2016

Package Queries: From Tuples to Packages

Berkeley, CA, USA

BERKELEY LABS

Jul 2015

ACADEMIA

Package Queries: Scalable Prescriptive Analytics Close to the Data

Virtually

GEORGIA TECH

Mar 2021

Package Queries: Scalable Prescriptive Analytics Close to the Data

Virtually

CARNEGIE MELLON UNIVERSITY

Mar 2021

Package Queries: Enabling Declarative and Scalable Prescriptive Analytics in Relational Data

Virtually

HARVARD UNIVERSITY, DASLAB

Jan 2021

Stochastic Package Queries For In-database Optimization Under Uncertainty

Virtually at INFORMS 2020

INFORMS 2020

Nov 2020

sPaQLToolS: A Stochastic Package Query Interface for Scalable Constrained Optimization

Tokyo, Japan (virtually)

VLDB 2020

Sep 2020

- Video: <https://www.youtube.com/watch?v=1v5KevGKbfE>

Stochastic Package Queries in Probabilistic Databases

Portland, OR, USA (virtually)

SIGMOD 2020

Jun 2020

- Video: <https://www.youtube.com/watch?v=UrEOEplK06I>

Package Queries: Enabling Declarative and Scalable Prescriptive Analytics in Relational Data

MIT CSAIL, DATA SYSTEMS GROUP

Cambridge, MA, USA

Feb 2020

Stochastic Package Queries for In-Database Constrained Optimization Under Uncertainty

MIT, NORTH EAST DATABASE DAY, SAMBERG CONFERENCE CENTER

Cambridge, MA, USA

Jan 2020

Scalable Package Queries in Relational Database Systems

VLDB 2016

New Delhi, India

Sep 2016

- Slides: packagebuilder.cs.umass.edu/papers/scalable-paql-presentation-share.pdf

Package Queries: From Tuples to Packages

AMPLAB, UC BERKELEY

Berkeley, CA, USA

Jun 2015

Improving Package Recommendations through Query Relaxation

VLDB 2014

Hangzhou, China

Sep 2014

- Slides: packagebuilder.cs.umass.edu/papers/package-recomm-slides-web.pdf

Metric Spaces for Temporal Information Retrieval

ECIR 2014

Amsterdam, the Netherlands

Apr 2014

Metric Spaces for Temporal Information Retrieval

UNIVERSITY OF BOLOGNA

Bologna, Italy

Dec 2013

Metric Spaces for Temporal Information Retrieval

ÅRHUS UNIVERSITY

Århus, Denmark

Jan 2013