Djellel Difallah

New York University Abu Dhabi, United Arab Emirates, P. O. Box 129188. Office: A1-185
Phone: +971 56 526 2829 | Email: djellel@nyu.edu | Web: djelleldifallah.com

Education

Ph.D. Computer Science 05/2011 – 06/2015
University of Fribourg, Switzerland
Thesis: *"Quality of Service in Crowd-Powered Systems"*
Advisor: Prof. Philippe Cudré-Mauroux

M.S. Computer Science (with Honors) 08/2009 – 05/2011
University of Louisiana at Lafayette, USA
Capstone Project: *"Frequent Association Action Rules Mining Using FP-Trees"*
Advisors: Prof. Vijay V. Raghavan and Prof. Ryan Benton
Scholarship: Fulbright Scholarship

Dipl.-Ing. Informatics 10/1999 – 12/2004
University of Science and Technology Houari Boumediene, Algeria
Capstone Project: *"Modeling the Information System of Air Algerie Health Centers"*

Professional Positions

Assistant Professor of Computer Science, NYU Abu Dhabi; Sept 2020–Present
Global Network Assistant Professor of Computer Science
Courant Institute of Mathematical Sciences, NYU

Affiliated Faculty, NYU-KAIST Global Innovation and Research Institute 2024–Present

Research Scientist Oct 2019 – August 2020
Wikimedia Foundation, Switzerland

Adjunct Assistant Professor Spring 2019
Leonard N. Stern School of Business, New York University, USA

Post-Doctoral Associate (Data Science Fellow) May 2017 – Oct 2019
Center for Data Science, New York University, USA

Senior Research Scientist Oct 2015 – May 2017
eXascale Infolab, University of Fribourg, Switzerland

Research Assistant May 2011 – July 2015
eXascale Infolab, University of Fribourg, Switzerland

Intern Summer 2013
Microsoft Research, Cloud and Information Services Lab, Mountain View, CA, USA

Intern (Google Summer of Code) Summer 2010
Mentoring Organization: DrizzleDB Project

Information Management Engineer 2006–2009
Schlumberger Information Solutions, Algeria

Engineer 2005–2006
EEPAD Internet Services Provider, Algeria

Teaching Experience

New York University Abu Dhabi, UAE

Instructor

- *Data Structures (CS-UH 1050); Enrollment: 23* — Spring 2025
- *Special Topics in Computer Science: Big Data Systems (CS-UH 3260); Enrollment: 24* — Fall 2024
- *Data Structures (CS-UH 1050); Enrollment: 24* — Fall 2023
- *Special Topics in Computer Science: Big Data Systems (CS-UH 3260); Enrollment: 24* — Spring 2023
- *Discrete Mathematics (CS-UH 1002); Enrollment: 23* — Fall 2022
- *Discrete Mathematics (CS-UH 1002-001); Two sections, 24 students each* — Fall 2021
- *Data Structures (CS-UH 1050); Enrollment: 24* — Spring 2021

- *Directed Study in Computer Science (CS-UH 3250): Graph Machine Learning; 6 students* — Spring 2025
- *Directed Study in Computer Science (CS-UH 3250): Probabilistic Machine Learning; 1 student* — Spring 2022

Leonard N. Stern School of Business, New York University, USA

Instructor

- *Data Mining for Business Analytics (Managerial) (INFO-GB.3336.20); Enrollment: 21* — Spring 2019

University of Fribourg, Switzerland

Co-Instructor

- *Big Data Infrastructures (Master); with Prof. Philippe Cudré-Mauroux* — Fall 2016
- *Data Science Seminar (Master); with Dr. Mourad Khayati* — Fall 2015

Teaching Assistant

- *Data Science Seminar (Master); Instructor: Prof. Philippe Cudré-Mauroux* — 2011-2014

Awards and Honors

NYU Abu Dhabi Palm Teaching Award — AY 2024/2025
Honorarium: 1,000 USD

Distinguished Reviewers Board, ACM Transactions on the Web — 2022

Outstanding Reviewer, The Web Conference — 2018

Best Computer Science PhD Thesis Award, University of Fribourg — February 2016
Awarded: 3,000 CHF

Research Visit Grant: European Science Foundation Research — Spring 2015
Awarded: 1,350 EUR

Travel Grant: HCOMP Doctoral Consortium — October 2014
Awarded: 700 USD

University of Louisiana at Lafayette Honors (for maintaining a GPA of 4.0) — 2011

Grants and Funded Projects

CityGraph: A Participatory Open Data Framework 2022–2024
Role: PI, with Prof. Etienne Wasmer (NYUAD Social Science)
Funding source: Center for Interacting Urban Networks, NYUAD
Grant Value: 2-year Postdoc, 1-month summer salary, 5,000 USD research funds

FashionBrain: Understanding Europe's Fashion Data Universe 2017
Role: Key personnel and grant writing (Left the project after six months to start a postdoc at NYU)
Funding Source: European Project Horizon 2020 (Big Data PPP)
Grant Value: 2.9M EUR total; University of Fribourg's share: 690,000 EUR

Query Expansion: Deep Learning and Crowdsourcing 2016
Role: Key personnel and lead grant writer
Funding Source: Swiss Commission of Technology and Innovation
Grant Value: 210,000 CHF

Publications

Google Scholar Metrics (as of June 2025): 3,573 citations; h-index:18; i10-index:26

Papers in Journals & Conference Proceedings

2025

1. **Djellel Difallah**. WikiRAG: Revisiting wikidata KGC datasets with community updates and retrieval-augmented generation. In *Proceedings of the 31st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '25)*, Toronto, ON, Canada, 2025. ACM. (To appear)

2024

2. Ryoji Kubo and **Djellel Difallah**. XGExplainer: Robust evaluation-based explanation for graph neural networks. In *Proceedings of the 2024 SIAM International Conference on Data Mining (SDM)*, pages 64–72. SIAM, 2024

2023

3. Farah Atif, Ola El Khatib, and **Djellel Difallah**. BeamQA: Multi-hop knowledge graph question answering with sequence-to-sequence prediction and beam search. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 781–790, 2023
4. Abdelouahab Khelifati, Mourad Khayati, Anton Dignös, **Djellel Difallah**, and Philippe Cudré-Mauroux. TSM-Bench: Benchmarking time series database systems for monitoring applications. *Proc. VLDB Endow.*, 16(11):3363–3376, 2023

2022

5. Nadya Abdel Majid, Ola Elkhatib, Shuan Gao, and **Djellel Difallah**. HyperKGQA: Question answering over knowledge graphs using hyperbolic representation learning. In *IEEE International Conference on Data Mining (ICDM)*, pages 309–318. IEEE, 2022
6. Natalia Ostapuk, **Djellel Difallah**, and Philippe Cudré-Mauroux. ParaGraph: Mapping wikidata tail entities to wikipedia paragraphs. In *IEEE Big Data*, pages 6008–6017. IEEE, 2022

7. **Djellel Difallah**, Diego Sáez-Trumper, Eriq Augustine, Robert West, and Leila Zia. Crosslingual section title alignment in wikipedia. In *IEEE Big Data*, pages 5892–5901. IEEE, 2022

2021

8. Michael Luggen, Julien Audiffren, **Djellel Difallah**, and Philippe Cudré-Mauroux. Wiki2Prop: A multimodal approach for predicting wikidata properties from wikipedia. In *Proceedings of the Web Conference 2021*, pages 2357–2366, 2021
9. Martin Gerlach, Marshall Miller, Rita Ho, Kosta Harlan, and **Djellel Difallah**. Multilingual entity linking system for wikipedia with a machine-in-the-loop approach. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management (CIKM)*, pages 3818–3827, 2021

2020

10. Lei Han, Alessandro Checco, **Djellel Difallah**, Gianluca Demartini, and Shazia Sadiq. Modelling user behavior dynamics with embeddings. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management (CIKM)*, pages 445–454, 2020

2019

11. Michael Luggen, **Djellel Difallah**, Christina Sarasua, Gianluca Demartini, and Philippe Cudré-Mauroux. Non-parametric class completeness estimators for collaborative knowledge graphs - the case of wikidata. In *ISWC (1)*, volume 11778 of *Lecture Notes in Computer Science*, pages 453–469. Springer, 2019
12. **Djellel Difallah**, Alessandro Checco, Gianluca Demartini, and Philippe Cudré-Mauroux. Deadline-aware fair scheduling for multi-tenant crowd-powered systems. *ACM Trans. Soc. Comput.*, 2(1):3:1–3:29, 2019
13. Cristina Sarasua, Alessandro Checco, Gianluca Demartini, **Djellel Difallah**, Michael Feldman, and Lydia Pintscher. The evolution of power and standard wikidata editors: Comparing editing behavior over time to predict lifespan and volume of edits. *Comput. Support. Cooperative Work.*, 28(5):843–882, 2019

2018

14. **Djellel Difallah**, Elena Filatova, and Panos Ipeirotis. Demographics and dynamics of mechanical turk workers. In *WSDM*, pages 135–143. ACM, 2018
15. Julien Plu, Roman Prokofyev, Alberto Tonon, Philippe Cudré-Mauroux, **Djellel Difallah**, Raphaël Troncy, and Giuseppe Rizzo. Sanaphor++: Combining deep neural networks with semantics for coreference resolution. In *LREC*. European Language Resources Association (ELRA), 2018

2017

16. Roman Prokofyev, Michael Luggen, **Djellel Difallah**, and Philippe Cudré-Mauroux. Swisslink: High-precision, context-free entity linking exploiting unambiguous labels. In *Proceedings of the 13th International Conference on Semantic Systems*. ACM, 2017
17. Gianluca Demartini, **Djellel Difallah**, Ujwal Gadiraju, and Michele Catasta. An introduction to hybrid human-machine information systems. *Found. Trends Web Sci.*, 7(1):1–87, 2017

2016

18. **Djellel Difallah**, Gianluca Demartini, and Philippe Cudré-Mauroux. Scheduling human intelligence tasks in multi-tenant crowd-powered systems. In *WWW*, pages 855–865. ACM, 2016

19. Alberto Tonon, Victor Felder, **Djellel Difallah**, and Philippe Cudré-Mauroux. VoldemortKG: Mapping schema.org and web entities to linked open data. In *ISWC (2)*, volume 9982 of *Lecture Notes in Computer Science*, pages 220–228, 2016

2015

20. Roman Prokofyev, Alberto Tonon, Michael Luggen, Loic Vouilloz, **Djellel Difallah**, and Philippe Cudré-Mauroux. SANAPHOR: ontology-based coreference resolution. In *ISWC (1)*, volume 9366 of *Lecture Notes in Computer Science*, pages 458–473. Springer, 2015
21. **Djellel Difallah**, Michele Catasta, Gianluca Demartini, Panagiotis G. Ipeirotis, and Philippe Cudré-Mauroux. The dynamics of micro-task crowdsourcing: The case of amazon mturk. In *WWW*, pages 238–247. ACM, 2015

2014

22. Michele Catasta, Alberto Tonon, **Djellel Difallah**, Gianluca Demartini, Karl Aberer, and Philippe Cudré-Mauroux. TransactiveDB: Tapping into collective human memories. *Proc. VLDB Endow.*, 7 (14):1977–1980, 2014b
23. Carlo Curino, **Djellel Difallah**, Chris Douglas, Subru Krishnan, Raghu Ramakrishnan, and Sriram Rao. Reservation-based scheduling: If you're late don't blame us! In *SoCC*, pages 2:1–2:14. ACM, 2014
24. **Djellel Difallah**, Michele Catasta, Gianluca Demartini, and Philippe Cudré-Mauroux. Scaling-up the crowd: Micro-task pricing schemes for worker retention and latency improvement. In *HCOMP*, pages 50–58. AAAI, 2014

2013

25. **Djellel Difallah**, Andrew Pavlo, Carlo Curino, and Philippe Cudré-Mauroux. OLTP-Bench: An extensible testbed for benchmarking relational databases. *Proc. VLDB Endow.*, 7(4):277–288, 2013c
26. **Djellel Difallah**, Gianluca Demartini, and Philippe Cudré-Mauroux. Pick-a-Crowd: tell me what you like, and i'll tell you what to do. In *WWW*, pages 367–374. ACM, 2013b
27. Gianluca Demartini, **Djellel Difallah**, and Philippe Cudré-Mauroux. Large-scale linked data integration using probabilistic reasoning and crowdsourcing. *VLDB J.*, 22(5):665–687, 2013
28. **Djellel Difallah**, Philippe Cudré-Mauroux, and Sean A. McKenna. Scalable anomaly detection for smart city infrastructure networks. *IEEE Internet Comput.*, 17(6):39–47, 2013a

2012

29. Gianluca Demartini, **Djellel Difallah**, and Philippe Cudré-Mauroux. Zencrowd: leveraging probabilistic reasoning and crowdsourcing techniques for large-scale entity linking. In *WWW*, pages 469–478. ACM, 2012

Workshop Papers, Demonstrations, and Resource Papers

30. Luca Althaus, Mourad Khayati, Abdelouahab Khelifati, Anton Dignös, **Djellel Difallah**, and Philippe Cudré-Mauroux. SEER: an end-to-end toolkit for benchmarking time series database systems in monitoring applications. *Proc. VLDB Endow.*, 17(12):4361–4364, 2024
31. Lifan Yu, Nadya Abdel Madjid, and **Djellel Difallah**. CrunchQA: A synthetic dataset for question answering over crunchbase knowledge graph. In *IEEE Big Data*, pages 4635–4641. IEEE, 2022

32. Natalia Ostapuk, **Djellel Difallah**, and Philippe Cudré-Mauroux. Sectionlinks: Mapping orphan wikidata entities onto wikipedia sections. In *Wikidata@ISWC*, volume 2773 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2020

33. Dana Van Aken, **Djellel Difallah**, Andrew Pavlo, Carlo Curino, and Philippe Cudré-Mauroux. Benchpress: Dynamic workload control in the oltp-bench testbed. In *SIGMOD Conference*, pages 1069–1073. ACM, 2015

34. Michele Catasta, Alberto Tonon, **Djellel Difallah**, Gianluca Demartini, Karl Aberer, and Philippe Cudré-Mauroux. Hippocampus: answering memory queries using transactive search. In *WWW (Companion Volume)*, pages 535–540. ACM, 2014a

35. Michele Catasta, Alberto Tonon, **Djellel Difallah**, Gianluca Demartini, Karl Aberer, and Philippe Cudré-Mauroux. TransactiveDB: Tapping into collective human memories. *Proc. VLDB Endow.*, 7 (14):1977–1980, 2014b

36. Carlo Curino, **Djellel Difallah**, Andrew Pavlo, and Philippe Cudré-Mauroux. Benchmarking oltp/web databases in the cloud: the oltp-bench framework. In *CloudDB@CIKM*, pages 17–20. ACM, 2012

37. **Djellel Difallah**, G Demartini, and P Cudré-Mauroux. Mechanical cheat: Spamming schemes and adversarial techniques on crowdsourcing platforms. In *CrowdSearch*, volume 842 of *CEUR Workshop Proceedings*, pages 26–30. CEUR-WS.org, 2012

38. **Djellel Difallah**, Ryan G. Benton, Vijay V. Raghavan, and Tom Johnsten. FAARM: frequent association action rules mining using fp-tree. In *ICDM Workshops*, pages 398–404. IEEE Computer Society, 2011

Papers Under Review

39. Mourad Khayati, **Djellel Difallah**, Zakhar Tymchenko, Quentin Nater, and Philippe Cudré-Mauroux. From missing to meaningful: Assessing imputation's impact on common time series downstream tasks, 2025 (submitted to ICDE 2025, currently under review)

40. Prince Larbi Ampofo, Ryoji Kubo, and **Djellel Difallah**. DBRAG: Multi-table retrieval-augmented generation for complex database queries, 2025 (submitted to CIKM 2025, currently under review)

41. Ola Elkhatib and **Djellel Difallah**. QALLM: Path generation and knowledge graph embeddings for multi-hop KGQA, 2025 (submitted to CIKM Short Papers 2025, currently under review)

42. Michael Luggen, Benedikt Hitz, Julien Audiffren, **Djellel Difallah**, Jean-Luc Cochard, and Philippe Cudré-Mauroux. Open government data as multi-dimensional 5 star data: cube.link, 2025 (submitted to ISWC In-Use Track 2025, currently under review)

Papers in Preparation (Preprint Available on ArXiv)

43. Dongzhe Fan, Yi Fang, Jiajin Liu, **Djellel Difallah**, and Qiaoyu Tan. MLaGA: Multimodal large language and graph assistant, 2025 (manuscript in preparation for submission to KDD 2026)

44. Ryoji Kubo and **Djellel Difallah**. RAW-Explainer: Post-hoc explanations of graph neural networks on knowledge graphs, 2025 (manuscript in preparation for submission to LoG (Learning on Graphs) 2026)

Thesis

45. **Djellel Difallah**. *Quality of Service in Crowd-powered Systems*. PhD thesis, University of Fribourg, Switzerland, 2015

Talks and Tutorials

Talks

- *Knowledge Discovery and Exploration with LLMs and Graph ML*, AIE 2025 (NYU Abu Dhabi), Invited Talk.
- *Machine-in-the-loop: A New Paradigm of Crowdsourcing for Wikipedia Editing* , WSDM 2023 Crowd Science Workshop (Singapore), Invited Talk.
- *Deadline-Aware Fair Scheduling for Multi-Tenant Crowd-Powered Systems*, CSCW 2019 (Texas, USA), invited presentation.
- *Crowd Size Estimation*, Moore-Sloan Data Science Summit 2018 (Utah, USA).
- *Big Data for Executives*, IIMT guest lecture (Fribourg, Switzerland, 2016).
- *Quality and Performance Optimizations in Microtask Crowdsourcing*, MIT CSAIL (Boston, MA, 2014).
- *Quality and Performance Optimizations in Microtask Crowdsourcing*, CMU (Pittsburgh, PA, 2014).
- *Extending the OLTP-Bench Framework for Big Data Systems*, Workshop on Big Data Benchmarking 2014 (Potsdam, Germany).
- *Mechanical Cheat: Spamming Schemes and Adversarial Techniques on Crowdsourcing Platforms*, Crowd-Search 2012 workshop at WWW (Lyon, France).

Conference Tutorials

- *Aggregation Techniques in Crowdsourcing: Multiple Choice Questions and Beyond*
  CIKM 2021 (online) – Organizers: Djellel Difallah, Alessandro Checco.
- *Using Crowdsourcing Effectively for Social Media Research*
  ICWSM 2016 (Cologne, Germany) – Organizers: Ujwal Gadiraju, Gianluca Demartini, Djellel Difallah, Michele Catasta.
- *It's Getting Crowded! How to Use Crowdsourcing Effectively for Web Science Research*
  Web Science 2016 (Hannover, Germany) – Organizers: Ujwal Gadiraju, Gianluca Demartini, Djellel Difallah, Michele Catasta.

Service

Senior Program Committee

- ICDM (Area Chair) : 2023, 2024
- CIKM (Research Track SPC): 2025, 2024, 2023
- CIKM (Applied Research Track SPC): 2022, 2021, 2020
- HCOMP (SPC): 2024
- AAAI: 2023, 2024

Program Committee

- IEEE International Conference on Data Mining (ICDM):
  2025, 2022, 2021, 2020, 2019, 2018
- ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD):
  2025, 2024, 2023, 2022, 2021, 2020, 2019, 2018, 2017
- ACM International Conference on Information and Knowledge Management (CIKM):
  2025, 2024, 2023, 2022, 2021, 2020, 2018, 2017
- ACM Conference on Research and Development in Information Retrieval (SIGIR):
  2025, 2024, 2023, 2022, 2021, 2020
- ACM International Conference on Web Search and Data Mining (WSDM): 2025, 2024, 2023, 2022, 2021
- The Web Conference (WWW): 2023, 2022, 2021, 2020, 2019, 2018, 2017
- AAAI Conference on Artificial Intelligence (AAAI): 2022, 2021, 2020

- IEEE International Conference on Data Engineering (ICDE): 2025, 2026
- ACM International Conference on Management of Data (SIGMOD): 2025
- Proceedings of the VLDB Endowmen (PVLDB): 2026
- AAAI Conference on Human Computation and Crowdsourcing (HCOMP): 2024, 2022, 2021, 2020, 2019
- International Semantic Web Conference (ISWC): 2022, 2021, 2020, 2019
- SIAM International Conference on Data Mining (SDM): 2022
- Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD): 2025
- The International AAAI Conference on Web and Social Media (ICWSM): 2018
- The International Conference on Web Engineering (ICWE): 2016
- International Conference on Social Informatics (SocInfo): 2022, 2020, 2019
- Wiki Workshop: 2023, 2022, 2021
- Wiki NLP Workshop: 2025

Conference Session Chair

- Session Chair: ICDM 2024, ISWC 2022, SIGIR 2021, KDD 2021

Editorial Board, Associate Editor

- Knowledge and Information Systems (KAIS) (Since Dec 2024)

Journal Reviewer

- Distinguished Reviewers Board, ACM Transactions on the Web (Since Oct 2022)
- ACM Transactions on the Web (3 papers)
- ACM Transactions on Computer-Human Interaction (1 paper)
- ACM Transactions on Social Computing (1 paper)
- SIGMOD Record (1 paper)
- ACM Transactions on Computer-Human Interaction (2 paper)

External Reviewer

- ACM Collective Intelligence (CI): 2024
- IEEE International Conference on Data Engineering (ICDE): 2016
- ACM Special Interest Group on Management of Data (SIGMOD): 2016
- International ACM Conference on Web Science (WebSci): 2016
- ACM Symposium on Cloud Computing (SoCC): 2016
- Proceedings of the VLDB Endowmen (PVLDB): 2014, 2016
- International Conference on Extending Database Technology (EDBT): 2013
- AAAI Conference on Artificial Intelligence (AAAI): 2014, 2015
- IEEE International Conference on Big Data (IEEE BigData): 2013

University Service – at NYU Abu Dhabi

- Science and Engineering Curriculum Committee (CS representative) (09/2022–present)
- Undergraduate Curriculum Committee (At-large member) (09/2021–12/2023)
- Master of Interdisciplinary Data Science and AI – Proposal working group member (2021–2022)
- Search Committee Member (CS Professor, Open Rank) (AY21/22)
- Search Committee Member (CS Instructor) (AY21/22)
- Contract Faculty Review Committees (Chair of one committee)
- Undergraduate Research Grants Reviewer (CS proposals) (2021–present)
- Piano Prize Committee Member (2022–2023)

Students Advising

PhD students

- Ola Elkhatib, NYU Tandon. Expected Graduation May 2026.

Research Associates and Assistants

- Ryoji Kubo. Sept 2023 - Present.
- Prince Ampofo. Sept 2023 - Present.
- Shahid Rabbani. Oct 2023 - March 2025. (now: NYU Abu dhabi)
- Wenqing Li. Sept 2023 - Aug 2024. (now: Hunan University)
- Rashid Alyassi. Jan 2023 - Aug 2023. (now: PhD Student, ETH Zurich)
- Farah Atif. Apr 2022 - Apr 2023. (now: PhD Student, MBZUAI)
- Nadya Abdel Madjid. Feb 2021 - Aug 2022. (now: PhD Student, Khalifa University)
- Bilal Saleem. Sept 2020 - Aug 2021. (now: PhD Student, Purdue University)

Undergraduate Capstone Students at NYUAD

AY 2024-2025

- Ali Abdullah Jafri: *"Predicting Formula 1 Race Outcomes: A Machine Learning Approach"*
- Kristina Aslanian: *"Mapping the UAE's Cultural Heritage Using Computational Text Analysis"*
- Sasha Malik: *"Leveraging Tree of Thoughts Strategy with Large Language Models for Question Answering on Structured Data"*
- Janindu Nanayakkara: *"Kosmograph: Knowledge Graph Visualization Tool"*

AY 2023-2024

- Dariga Shokayeva: *"Named-Entity Linking: Mapping Column Names to Wikidata Properties"*
- Sagynbek Talgatuly: *"VocabBooster: A Personalized English Vocabulary Learning App Powered by Machine Learning and Taxonomies"*

AY 2022-2023

- Ryoji Kubo: *"Learning Explanations for Graph Neural Networks based Knowledge Graph Tasks"*
- Kartikey Singhal: *"Multi-Modal Content Retrieval in Wikipedia"*
- Yernar Mukayev: *"University orientated knowledge platform using crowdsourcing for summarization"*
- Hanzalla Usman: *"Kariger: A Crowd Sourcing Web-App for Physical Tasks in the Gig Economy"*

AY 2021-2022

- Fuseini Sunnuma: *"WikiContribute: A Light-weight Context-aware Mobile Application for Editing and Browsing Wikidata"*
- Wajahat Mirza: *"Wikidata2Vec Embeddings & Systematization of research on Entity Enrichment of Transformer based Language Models"*
- Prabhav Arora: *"An evaluation of state of the art Neural Architectures utilized for Time Series Forecasting"*
- Seif Kassem: *"Reinforcement Learning Framework for Data Cleaning"*
- Omar Hussein: *"Building a Tool to Assess the Quality of Writing in Wikipedia Articles"*

AY 2020-2021

- Branden Kang: *"Link Predictions Using PyTorch-BigGraph and Gensim"*
- Estelle Ocran: *"Learning Entity Embeddings from Wikipedia and Wikidata for Entity or Article Prediction"*

Master Students at NYU New York

- Shuang Gao: "Hyperbolic Question Answering over Knowledge Graphs" – Fall 2020

Dissertation Committees

- Abdelouahab Khelifati Ph.D. student, University of Fribourg, Switzerland (2025)
  Thesis: "A Holistic Approach for Time Series Management: Unifying Data Storage and Data Processing"
- Natalia Ostapuk, Ph.D. student, University of Fribourg, Switzerland (2024)
  Thesis: "Enhancing Structured Knowledge Representation in Large-Scale Data Environments"
- Michael Luggen, Ph.D. student, University of Fribourg, Switzerland (2020)
  Thesis: "Exploiting and Enriching Highly Dynamic Linked Data"

Qualifying Exam Committees

- Anh Mai, NYU Courant. 2024

Master and Bachelor Student Projects at the University of Fribourg

- Alexander Striffeler: "Benchmarking Big Data Infrastructures" – 2015
- Marwa Bouzeyane: "Grammatical Errors Detection Based On N-gram Analysis" – 2015
- Stefan Nüesch: "Real-Time Anomaly Detection in Water Distribution Networks using Spark Streaming" – 2014
- Phokham Nonava: "HDFS Blocks Placement Strategy" – 2014
- Simpal Kumar: "Real Time Data Analysis for Water Distribution Network using Storm" – 2014
- Dani Rotzetter: "Crowd-Flow Designer: An Open-Source Toolkit to Design and Run Complex Crowd-Sourced Tasks" – 2014
- Victor Felder: "Openturk: An Implementation of the Pick-A-Crowd Architecture" – 2013