Kashif Rabbani

✉ ch.kashirabbani@gmail.com ➤ https://kashif-rabbani.github.io/ 📍 Aalborg, Denmark

📇 Profile

I am a versatile and accomplished PhD researcher with expertise in databases, query optimization, and data quality. I will graduate in April 2024. I have a strong track record of published research and academic excellence, with almost two years of experience in the industry. I seek to apply advanced database skills and innovative problem-solving abilities to industry projects. My dedication lies in implementing data-driven solutions that can significantly impact a constantly evolving corporate environment.

💼 Professional Experience

PhD Researcher, *Dept. of Computer Science, Aalborg University* 🔗 — 09/2020 – present, Aalborg, Denmark
Research Areas: *Databases, Property Graphs, Knowledge Graphs, Query optimization, Schemas, and Data Quality*

Scientific Research Assistant, *Dept. of Computer Science, Aalborg University* 🔗 — 10/2019 – 08/2020, Aalborg, Denmark
Responsibilities: Managing research projects, maintaining and improving existing features, conducting surveys, and writing research reports.

Data Engineer Intern, *everis NTT Data Company* 🔗 — 06/2018 – 10/2018, Barcelona, Spain
Responsibilities: Data Modeling, Designing Ontologies, Designing Scalable Architecture, working on end-to-end solution for EU-CORDIS research projects.

Software Engineer, *eConceptions pvt Ltd* 🔗 — 01/2016 – 09/2017, Islamabad, Pakistan
Backend Development (Java), Front-end development (Javascript, MEAN Stack)

Web Developer, *Solution Providers(eSolPro) pvt Ltd* 🔗 — 08/2015 – 09/2015, Islamabad, Pakistan
APIs development (PHP), PSD design, HTML, CSS, Javascript

🎓 Education

PhD, *Dept. of Computer Science, Aalborg University* 🔗 — 09/2020 – 04/2024, Aalborg, Denmark
Thesis Title: Query optimization, Schema Extraction, and Data Quality for Graphs.
Supervisors: Prof. Katja Hose 🔗 , Assist. Prof. Matteo Lissandrini 🔗

Masters in Computer Science, *TU-Berlin* 🔗 — 09/2017 – 08/2019, Berlin, Germany
Master Degree in *"Big Data Management and Analytics"*. Final Grade: 1,4 (Sehr gut/very good)

Bachelors in Computer Science, *COMSATS University* 🔗 — 02/2013 – 02/2017, Islamabad, Pakistan
CGPA 3.84/4.00 (Campus Gold Medal & Institute Silver Medal)

🧠 Area of Expertise

Databases
MySQL, PostgreSQL, Oracle Database, Microsoft SQL Server, SQLite, MongoDB, Redis, Amazon DynamoDB

Big Data Processing
ETL, Spark, Apache Hadoop, Apache Flink

Programming Languages
Java, Python, C, C++, JavaScript

Misc.
Cloud Services (AWS), Tableau, Weka, RapidMiner, BPMN, Latex, Linux

Graphs & Semantic Web Technologies
Neo4j, GraphDB, Amazon Neptune, ArangoDB, OrientDB, Stardog, Virtuoso, Apache Jena, RDF4j, OWL, RDF, RDFS, SPARQL, SHACL

Web Development
Java Spring Framework, Vaadin Framework, MVC, HTML, CSS, Bootstrap, Ajax

Academic
Research, Reviewing, Paper Writing, Teaching, Supervision, PBL, Censoring, Dissemination, Networking, Collaboration

🎖 Achievements

PhD Study Abroad Travel Grants 2023, *Otto Mønsted & Danish Data Science Academy* 🔗   2023
I was awarded travel grants by Otto Mønsted and Danish Data Science Academy for study abroad during my PhD studies.

Big Data Talent Awards 2019 (Runner-up), *FIB, UPC Barcelona, Spain* 🔗   2019
I was selected as a runner-up for Big Data Talent Awards 2019 at UPC Barcelona, Spain, for my master thesis.

Erasmus Mundus Scholarship,   2017
Erasmus+ EU programme for education, training, youth and sport 🔗
I was awarded a fully funded scholarship for my master's degree in Big Data Management and Analytics.

Gold & Silver Medal, *COMSATS University, Islamabad Pakistan*   2017
I was awarded with a campus Gold Medal award and an institute Silver Medal award for the highest CGPA: 3.84/4.00 (Batch Spring 2013-17) of my Bachelor's degree.

🏠 Professional Activities

Supervisor, *Groups Supervisor for BSc Software Engineering students*   2021 – 2023
• 4th Semester Students: Supervision of groups working in medical domain for knowledge   Aalborg University, DK
  engineering. Specifically, with MIMIC healthcare dataset, aiming to design a common general purpose ontology using existing common data models like OMOP, SCDM, and PCORnet.
• 6th Semester Students: Supervision of groups working with large scale ships AIS data aiming to create a system to improve existing navigation and illegal fishing

Teaching Assistant, *Database Management System*   2021 – 2022
Responsibilities: Database design and querying exercise sessions   Aalborg University, DK

External Reviewer, *VLDB-2022, ISWC-2022, SIGMOD-2022*

🌐 Languages

English (C1), **Danish** (PD-3), **Urdu** (Mother Tongue)

📖 Publications

1. Extraction of Validating Shapes from very large Knowledge Graphs, *VLDB* 🔗   2023
Kashif Rabbani, *Matteo Lissandrini, and Katja Hose,* In Proceedings of the Very Large Databases 2023 (Volume 16 Issue 5, VLDB-2023), Vancouver Canada.

2. SHACTOR: Improving the Quality of Large-Scale Knowledge Graphs with Validating Shapes, *SIGMOD/ACM* 🔗   2023
Kashif Rabbani, *Matteo Lissandrini, and Katja Hose,* In Proceedings of the 2023 International Conference on Management of Data, (SIGMOD-Companion '23) Seattle, USA.

3. Incremental Schema Integration for Data Wrangling via Knowledge Graphs, *SWJ* 🔗   2023
Javier Flores, ***Kashif Rabbani,*** *Sergi Nadal, Cristina Gómez, Oscar Romero, Emmanuel Jamin, and Stamatia Dasiopoulou,* In Semantic Web Journal.

4. SHACL and ShEx in the Wild: A Community Survey on Validating Shapes Generation and Adoption, *TheWebConf/ ACM* 🔗   2022
Kashif Rabbani, *Matteo Lissandrini, and Katja Hose,* In Companion Proceedings of the Web Conference WWW, 2022

5. Optimizing SPARQL Queries using Shape Statistics, *EDBT* 🔗   2021
Kashif Rabbani, Matteo Lissandrini, and Katja Hose, In Proceedings of the 24th
International Conference on Extending Database Technology.

6. ODIN: A dataspace management system, *ISWC* 🔗   2019
*Nadal Francesch Sergi, **Kashif Rabbani**, Óscar Romero Moral, and Shumet Tadesse Nigatu,* In
Proceedings of the ISWC 2019 Satellite Tracks (Posters & Demonstrations, Industry, and
Outrageous Ideas) co-located with 18th International Semantic Web Conference (ISWC).

7. ARDI: Automatic Generation of RDFS Models from Heterogeneous Data Sources,   2019
EDOC 🔗
*Shumet Tadesse, Cristina Gómez, Oscar Romero, Katja Hose, **Kashif Rabbani**,* In IEEE 23rd
International Enterprise Distributed Object Computing Conference (EDOC).