International Workshop on Quantum Data Science and Management (QDSM)

Valter Uotila[1], Sven Groppe[2], Le Gruenwald[3], Jiaheng Lu[1] and Wolfgang Mauerer[4]

[1]University of Helsinki, Finland
[2]Institute of Information Systems (IFIS), University of Lübeck, Germany
[3]The University of Oklahoma, USA
[4]Technical University of Applied Science Regensburg, Germany

Abstract

The first international workshop on Quantum Data Science and Management (QDSM), co-located with VLDB 2023, is centered around addressing the possibilities of quantum computing for data science and data management. Quantum computing is a relatively new and emerging field that is believed to have huge computational potential in the future. In the QDSM workshop, we want to provide a venue for discussing and publishing novel results of applying quantum computing to hard data science and data management problems. These problems include join order optimization, designing efficient quantum feature maps, studying possibilities of solving linear programs with quantum algorithms, and divergent index tuning with quantum machine learning. Besides, we include a short and visionary survey on quantum computing for databases. The workshop provides a platform for active discussion on these and related topics.

1. Introduction

The field of quantum computing has experienced remarkable progress after decades of research and development. Prototypes of quantum computers already exist and have been made available to users through cloud services (e.g., IBM Q experience, Google Quantum AI, or Xanadu Cloud). Although large-scale fault-tolerant quantum computers are not available yet, the potential of this new technology is undeniable. Quantum algorithms have the proven ability to either outperform classical approaches for various tasks or are impossible to be efficiently simulated by classical means under reasonable complexity-theoretic assumptions. Even imperfect current-day technology is speculated to exhibit computational advantages over classical systems.

For most database researchers, quantum computing and quantum machine learning are still new research fields. The goal of this workshop is to bring together academic researchers and industry practitioners from multiple disciplines (e.g., database, AI, software, physics, etc.) to discuss the challenges, solutions, and applications of quantum computing and quantum machine learning that have the potential to advance the state of the art of data science and data management technologies. Our purpose is to foster the interaction between database researchers and more traditional quantum disciplines, as well as industrial users. The workshop serves as a forum for the growing quantum computing community to connect with database researchers to discuss the wider questions and applications of how quantum resources can benefit data science and data management tasks, and how quantum software can support this endeavor.

We believe that many unsolved and interesting issues can be found at boundaries and intersections between different fields and that there are insufficient venues to publish such cross-disciplinary results. We also believe that an important aspect of future quantum computing will concern issues of handling data in one way or another. The workshop Quantum Data Science and Management will serve as a venue not only to discuss early, experimental results in research but also to feature a demonstration part with the intention of providing attendees with first-hand experience in using novel quantum computing techniques that go beyond the simple examples offered by various web services. This will give researchers a realistic intuition about quantum computing for data science and data management tasks.

2. Types of papers

The workshop solicits papers of the following categories.

Research Papers propose new approaches, theories or techniques related to quantum data science and management, including new data structures, protocols, and algorithms. They should make substantial theoretical and empirical contributions to the research field.

System Papers describe new systems and whole

Joint Workshops at 49th International Conference on Very Large Data Bases (VLDBW'23) — International Workshop on Quantum Data Science and Management (QDSM'23), August 28 - September 1, 2023, Vancouver, Canada

valter.uotila@helsinki.fi (V. Uotila); sven.groppe@uni-luebeck.de (S. Groppe); ggruenwald@ou.edu (L. Gruenwald)
0000-0001-8201-397X (V. Uotila); 0000-0001-5196-1117 (S. Groppe)

frameworks for enabling quantum data science and management.

Experiments and Analysis Papers focus on the experimental evaluation of existing approaches, including data structures and algorithms for quantum data science and management, and bring new insights through the analysis of these experiments. Results of Experiments and Analysis Papers can be, for example, showing the benefits of well-known approaches in new settings and environments, opening new research problems by demonstrating unexpected behavior or phenomena or comparing a set of traditional approaches in an experimental survey.

Application Papers report practical experiences on applications of quantum data science and management. Application Papers might describe how to apply quantum technologies to specific application domains.

Vision Papers identify emerging new or future research issues and directions and describe new research visions for quantum data science and management. The new visions will potentially have significant impacts on society.

Demo Papers deal with innovative approaches and applications for quantum data science and management. These papers describe a showcase of the proposed approach/application. We are especially interested in demonstrations having a WOW effect.

3. Topics of Interest

We are interested in all topics concerning quantum computing for data science and management, such as the following:

- Quantum computing, quantum algorithms and quantum software tools for problems related to data science and management
- Quantum machine learning for data science, data management and database optimization
- Post-quantum cryptography and security for databases and data management
- Classical data science and management for quantum computing and quantum machine learning

4. Review process

We have enforced a rigorous peer and single-anonymous review process with the option for authors of a double-anonymous review process. All manuscripts submitted to our workshop have been reviewed by at least three PC members. To verify the originality of submissions, we have used Plagiarism Detection Tools to check the content of the submitted manuscripts against previous publications.

The articles have been evaluated according to the following aspects:

- Relevance to the workshop
- Novelty and practical impact
- Technical soundness
- Appropriateness and adequacy of literature review, background discussion, and analysis of issues
- Presentation, including overall organization, English, and readability

5. Rationale about Recruiting the Chairs and Program Committee with special regard to Diversity Considerations

The PC chairs of the Quantum Data Science and Management workshop are coming from two continents, Europe and North America, which would attract an international community. One of the PC chairs is female (25%). The h-index of the PC chairs ranges between 19 and 40[1].

We have currently recruited 20 PC members (inclusive chairs) listed in the previous section who are experts in the topics of interest of our proposed workshop. All PC members have already confirmed their membership. Our PC represents a good mixture of different experiences, not only in terms of research areas but also in terms of levels of research experience. Although most PC members are from academia, we also have two experts from the industry and one expert from a national research laboratory. The chairs and PC members are listed in the Appendix.

6. Accepted Papers

The accepted papers include four research papers and one short survey paper.

Quantum Optimisation of General Join Trees [1] by Scönberger, Trummer, and Mauerer is a continuation of their previous work on optimizing left-deep join trees using quantum computing [2]. In this work, they introduce a native QUBO encoding for the general join ordering problem, which selects an optimal plan among bushy join trees.

An Evolutionary Algorithm Design for Pauli-based Quantum Kernel Classification [3] by Tjandra and Sugiarto tackles the problem of designing a high-performance quantum feature map that could work as an alternative kernel for Support Vector Classifiers. They represent a method to automatically generate Pauli feature maps using the genetic algorithm. Based on the results of their evaluation, the Pauli feature maps generated by

[1] according to Google Scholar and Scopus

the genetic algorithm perform better than several other classical and quantum kernel baselines.

Empirical evaluation of a quantum accelerated approach for the central path method in linear programming [4] by Adoni, Hafshejani and Gaur study if the central path method benefits from replacing the equation-solving step with the quantum algorithm for linear systems of equations (HHL-algorithm). They use numerical simulations and multiple instances of the proposed algorithm to evaluate the effectiveness of the quantum computing-based approach.

Index Tuning with Machine Learning on Quantum Computers for Large-Scale Database Applications [5] by Gruenwald, Winker, Çalıkyılmaz, Groppe, and Groppe provides a vision of a quantum machine learning algorithm to optimize the divergent design index tuning problem for replicated databases to minimize query processing costs. They outline the previous work and discuss the challenges and issues of designing such a quantum algorithm.

Quantum Computing for Databases: A Short Survey and Vision [6] by Yuan, Lu, Chen, Wu, Yao, Yan, and Chen represents a short survey and a visionary outline for the future of applying quantum computing for data management and databases.

7. Proceedings

We publish the accepted articles in a joint workshop proceedings called Proceedings of VLDB Workshops (VLDBW23) published by the CEUR Workshop Proceedings (CEUR-WS.org).

CEUR papers are indexed in Scopus, DBLP, SJR and other bibliographic databases. Altogether, this ensures maximum visibility to all who are interested in the topics of our workshop.

8. Stipend

Quantum Brilliance is a proud sponsor of the 2023 edition of Quantum Data Science and Management. Quantum Brilliance develops room-temperature diamond quantum accelerators for massively parallelized, edge and ubiquitous quantum computing. Our industry sponsor Quantum Brilliance offers to supplement the registration fees for young researchers, as well as researchers from underrepresented counties and communities.

9. Summary and Conclusions

Quantum computing hardware and software are at an early stage. Although we do not currently have real-world applications showing a quantum advantage over the corresponding classical algorithms, we are still positive about the future of quantum computing. We are convinced that it is the right moment to start researching quantum computing possibilities for data science and data management. We are looking forward to motivating, insightful, and enthusiastic discussions at the workshop. The submitted papers and the expertise of the keynote speakers and the authors are among the first to propose quantum computing solutions to data science and data management problems. We are confident that our workshop will foster the collaboration of researchers and practitioners and support networking for long-lasting professional connections after the workshop.

References

[1] M. Schönberger, I. Trummer, W. Mauerer, Quantum optimisation of general join trees, Joint Workshops at 49th International Conference on Very Large Data Bases (VLDBW'23) — International Workshop on Quantum Data Science and Management (QDSM'23) (2023).

[2] M. Schönberger, S. Scherzinger, W. Mauerer, Ready to leap (by co-design)? join order optimisation on quantum hardware, Proc. ACM Manag. Data 1 (2023). URL: https://doi.org/10.1145/3588946. doi:10.1145/3588946.

[3] Y. Tjandra, H. S. Sugiarto, An evolutionary algorithm design for pauli-based quantum kernel classification, Joint Workshops at 49th International Conference on Very Large Data Bases (VLDBW'23) — International Workshop on Quantum Data Science and Management (QDSM'23) (2023).

[4] V. Adoni, S. F. Hafshejani, D. Gaur, Empirical evaluation of a quantum accelerated approach for the central path method in linear programming, Joint Workshops at 49th International Conference on Very Large Data Bases (VLDBW'23) — International Workshop on Quantum Data Science and Management (QDSM'23) (2023).

[5] L. Gruenwald, T. Winker, U. Çalıkyılmaz, J. Groppe, S. Groppe, Index tuning with machine learning on quantum computers for large-scale database applications, Joint Workshops at 49th International Conference on Very Large Data Bases (VLDBW'23) — International Workshop on Quantum Data Science and Management (QDSM'23) (2023).

[6] G. Yuan, J. Lu, Y. Chen, S. Wu, C. Yao, Z. Yan, T. Li, G. Chen, Quantum computing for databases: A short survey and vision, Joint Workshops at 49th International Conference on Very Large Data Bases (VLDBW'23) — International Workshop on Quantum Data Science and Management (QDSM'23) (2023).

A. Workshop Chairs and Organizers

Sven Groppe is a Professor at the University of Lübeck, Germany. He received 7 project grants from DFG, BMBF, and BMWi in the area of data management. He is the project coordinator of the BMBF-funded QC4DB project about accelerating relational database management systems via quantum computing. He published more than 150 journal, conference, and workshop papers at top-ranked publication venues including SIGMOD, VLDB, and ICPP with over 125 co-authors from 17 countries worldwide. He is a member of over 110 program committees of various conferences and workshops and a reviewer of over 35 journals. He is a workshop chair of SBD@SIGMOD (2016-2020), BiDEDE@SIGMOD (2021-2023), VLIoT@VLDB (2017-2022) and QDSM@VLDB 2023. He is a general chair of the International Semantic Intelligence Conference (ISIC) (2021-2022), International Health Informatics Conference (IHIC) (2022- 2023), and the International Conference on Applied Machine Learning and Data Analytics (AMLDA) in 2023. More information is available on https://www.ifis.uni-luebeck.de/~groppe/.

Jiaheng Lu is a Professor at the University of Helsinki, Finland. His current research interests focus on multi-model databases and quantum computing for database applications. He has written four books on Hadoop and NoSQL databases, and more than 130 journal and conference papers published in SIGMOD, VLDB, TODS, etc. He was the workshop co-chair of Keyword search and data exploratory with ICDE 2016, Keyword search on structured data with SIGMOD 2012, and Cloud databases with CIKM 2010.

Wolfgang Mauerer is a Professor at the Technical University of Applied Sciences Regensburg, and a Senior Research Scientist at Siemens Technology. His interests focus on software/systems engineering, and quantum computing. He has published strongly multi-disciplinary work in venues and journals from Nature Photonics and PRL via ICSE and TSE to SIGMOD, and is the author of three books. His organization experience includes industrial and open-source conferences with hundreds of attendees. For more details, see https://lfdr.de/.

Le Gruenwald is a Professor, Dr. David W. Franke Professor, and Samuel Roberts Noble Foundation Presidential Professor in the School of Computer Science at The University of Oklahoma, U.S.A. She also worked for National Science Foundation (NSF) in the Directorate for Computer & Information Science & Engineering (CISE) as a Cluster Lead and Program Director of the Information Integration and Informatics cluster and a Program Director of the Cyber Trust program. Her major research interests include Machine Learning Enabled Cloud Data Management, Quantum Data Management and Mining, Semantic Web, Data Stream Management and Mining, and Mobile Data Management. She has published more than 200 technical articles in journals, books, and conference proceedings.

B. Proceedings Chair

Valter Uotila is a Ph.D. student in the Unified Database Management Systems research group at the University of Helsinki. He received a master's degree in mathematics. His research interests are in the intersection of databases, quantum computing, and category theory.

C. Publicity Chairs

- Srinjoy Ganguly, Woxsen University, India
- Sanjay Vishwakarma, IBM Quantum, IBM Research - Almaden, USA

D. Stipend Chair

- Ghanshyam Singh, Malaviya National Institute of Technology Jaipur (MNIT), India

E. Program Committee

Our program committee consists of the following experts:

- Umut Çalikyilmaz, University of Lübeck, Germany
- Prasanna Date, Oak Ridge National Laboratory, USA
- Maja Franz, OTH Regensburg, Germany
- Srinjoy Ganguly, Woxsen University, India
- Jan Lellmann, University of Lübeck, Germany
- Jukka Nurminen, University of Helsinki, Finland
- Natacha Kuete Meli, University of Lübeck, Germany
- Nitin Nayak, University of Lübeck, Germany
- Stefan Prestel, Quantum Brilliance GmbH, Germany
- Manuel Schönberger, OTH Regensburg, Germany
- Ilya Safro, University of Delaware, USA
- Ghanshyam Singh, Malaviya National Institute of Technology Jaipur (MNIT), India
- Valter Uotila, University of Helsinki, Finland
- Sanjay Vishwakarma, IBM Quantum, IBM Research - Almaden, USA
- Tobias Winker, University of Lübeck, Germany
- Zhengtong Yan, University of Helsinki, Finland