Mashaal Musleh

Ph.D. Computer Science & Engineering
University of Minnesota, MN, USA

Email: musle005@umn.edu
Mobile: +1 651-786-9855
Web Page: mashaal.me

Education

- **Ph.D. in Computer Science & Engineering** — MN, USA
  University of Minnesota; Research in Scalable Spatial Data Management and its applications. — *Sept. 2019 – May 2025*
  Advisor: Prof. Mohamed F. Mokbel. Thesis Title: Towards Highly Accurate Map Services.

- **M.S. in Computer Science & Engineering** — MN, USA
  University of Minnesota; GPA: 3.83/4.00. — *Sept. 2016 – May 2018*
  Advisor: Prof. Mohamed F. Mokbel.

- **B.S. in Computer Engineering** — Makkah, Saudi Arabia
  Umm Al-Qura University; GPA: 4.00/4.00 (First Class Honours). — *Sept. 2009 – Jan. 2015*

Selected Research and Professional Experience

- **Teradata - Query Optimizer R&D Team** — San Diego, California
  Senior Software Engineer — *Jun. – Aug. 2022 & May 2023 – Present*
  - **Code Generation:** Propose and Design LLVM-based high-performance frameworks for database query execution.

- **University of Minnesota** — MN, USA
  Research & Teaching Assistant — *Sept. 2016 – May 2023*
  - **KAMEL:** a novel system for GPS trajectory imputation. It accurately infers missing points along sparse trajectories without the need of the underlying map. It utilizes a BERT-like model trained on GPS data to generate dense, high-quality trajectories with a recall score of over 89% of missing points. KAMEL boosts the accuracy and reliablilty of map inference and any other downstream applications that uses trajectory data as input.
  - **TrajBERT:** A vision and undergoing project that treats vehicle trajectories as sentences and then uses the BERT NLP model to solve various trajectory problems after mapping them to their counterparts in the NLP domain (prediction, classification, etc.). TrajBERT does not only store and index trajectories but proposes techniques to tokenize the space and address the spatio-temporal constraints in order to natively supports trajectory analysis.
  - **RASED:** Developed a scalable interactive dashboard for monitoring road network updates in OpenStreetMap, which allows users to see the road updates worldwide and understand their continuous maintenance and development.
  - **QARTA:** Developed a map services system that utilizes machine learning algorithms to slash the error in the estimated time of arrival queries (ETA) by 25% over the state-of-the-art open-source methods.
  - **TurboReg and RegRocket:** Established a scalable spatial analysis framework that maps Auto Logistic Regression models into an equivalent Markov Logic Network which reduced learning time by three orders of magnitude.
  - **ST-Hadoop:** Planned and performed extensive experiments to evaluate the system performance which achieves orders of magnitude better performance than Hadoop and SpaitalHadoop when processing spatio-temporal data.

- **Qatar Computing Research Institute (QCRI)** — Doha, Qatar
  Research Associate — *Aug. 2018 – Aug. 2019*
  - **CoClean:** Proposed and created an open-source data science platform that allows multiple data scientists to collaboratively curate and prepare their datasets using their existing workflows on Jupyter notebooks.

- **KACST GIS Technology Innovation Center (GISTIC)** — Makkah, Saudi Arabia
  Research Software Engineer — *Jan. 2014 – Aug. 2016*
  - **Taghreed:** Implemented the recovery management, stress testing, and API layer for Taghreed, a scalable system for querying, analyzing, and visualizing billions of geotagged microblogs. The project was later generalized into an open-sourced system called Kite, and was the core of a successful startup named Lucidya that is currently serving hundreds of governmental and commercial customers.

Publications

- **Conference Papers**
  - **Mashaal Musleh** and Mohamed Mokbel. KAMEL: A Scalable BERT-based System for Trajectory Imputation. *International Conference on Very Large Data Bases, VLDB*, 17(3), 2023. [PDF]
  - **Mashaal Musleh**, Mohamed F Mokbel, and Sofiane Abbar. Let's Speak Trajectories. In *ACM International Conference on Advances In Geographic Information Systems, SIGSPATIAL*, 2022. (Vision Paper). [PDF]
  - **Mashaal Musleh** and Mohamed F Mokbel. RASED: A Scalable Dashboard for Monitoring Road Network Updates in OSM. in *IEEE International Conference on Mobile Data Management, MDM*, 2022. [PDF]
  - **Mashaal Musleh**, Sofiane Abbar, Rade Stanojevic, and Mohamed Mokbel. QARTA: An ML-based System for Accurate Map Services. *International Conference on Very Large Data Bases, VLDB*, 14(11), 2021. [PDF]
  - Ibrahim Sabek, **Mashaal Musleh**, and Mohamed F Mokbel. TurboReg: A Framework For Scaling Up Spatial Logistic Regression Models. In *ACM International Conference on Advances In Geographic Information Systems, SIGSPATIAL*, 2018. [PDF]
  - Louai Alarabi, Mohamed F. Mokbel, and **Mashaal Musleh**. ST-Hadoop: A MapReduce Framework for Spatio-Temporal Data. In *International Symposium on Advances in Spatial and Temporal Databases, SSTD*, volume 10411. Springer, 2017. [PDF]
  - Amr Magdy, Thanaa M Ghanem, **Mashaal Musleh**, and Mohamed F Mokbel. Understanding Language Diversity In Local Twitter Communities. In *ACM Conference on Hypertext and Social Media*, 2016. [PDF]
  - Amr Magdy, **Mashaal Musleh**, Kareem Tarek, Louai Alarabi, Saif Al-Harthi, Hicham G Elmongui, Thanaa M Ghanem, Sohaib Ghani, and Mohamed F Mokbel. Taqreer: A System For Spatio-Temporal Analysis on Microblogs. *IEEE Data Engineering Bulletin*, 38(2), 2015. [PDF]
  - Amr Magdy, Louai Alarabi, Saif Al-Harthi, **Mashaal Musleh**, Thanaa M Ghanem, Sohaib Ghani, and Mohamed F Mokbel. Taghreed: A System For Querying, Analyzing, And Visualizing Geotagged Microblogs. In *ACM International Conference on Advances in Geographic Information Systems, SIGSPATIAL*, 2014. [PDF]
- **Demos and Short Papers**
  - **Mashaal Musleh** and Mohamed Mokbel. A Demonstration of KAMEL: A Scalable BERT-based System for Trajectory Imputation. In *ACM International Conference on Management of Data, SIGMOD*, 2023. **[Best Demo Award]** [PDF] [VIDEO]
  - **Mashaal Musleh**. Towards A Unified Deep Model For Trajectory Analysis. In *ACM International Conference on Advances In Geographic Information Systems, SIGSPATIAL*, 2022. **[SRC Winner 1st place]** [PDF]
  - **Mashaal Musleh** and Mohamed F Mokbel. A Demonstration of RASED: A Scalable Dashboard for Monitoring Road Network Updates in OSM. in *IEEE International Conference on Data Engineering, ICDE*, 2022. [PDF] [VIDEO]
  - Sofiane Abbar, Rade Stanojevic, **Mashaal Musleh**, Mohamed ElShrif, and Mohamed Mokbel. A Demonstration of QARTA: An ML-based System for Accurate Map Services. *International Conference on Very Large Data Bases, VLDB*, 14(12), 2021. [PDF]
  - **Mashaal Musleh**, Mourad Ouzzani, Nan Tang, and AnHai Doan. CoClean: Collaborative Data Cleaning. In *ACM International Conference on Management of Data, SIGMOD*, 2020. [PDF] [CODE] [VIDEO]
  - Ibrahim Sabek, **Mashaal Musleh**, and Mohamed F Mokbel. Flash In Action: Scalable Spatial Data Analysis Using Markov Logic Networks. *International Conference on Very Large Data Bases, VLDB*, 12(12), 2019. [PDF]
  - Ibrahim Sabek, **Mashaal Musleh**, and Mohamed Mokbel. A Demonstration of Sya: A Spatial Probabilistic Knowledge Base Construction System. In *ACM International Conference on Management of Data, SIGMOD*, 2018. [PDF]
  - Amr Magdy, Louai Alarabi, Saif Al-Harthi, **Mashaal Musleh**, Thanaa M Ghanem, Sohaib Ghani, Saleh Basalamah, and Mohamed F Mokbel. Demonstration of Taghreed: A System For Querying, Analyzing, And Visualizing Geotagged Microblogs. In *IEEE International Conference on Data Engineering, ICDE*, 2015. [PDF]
  - Thanaa M Ghanem, Amr Magdy, **Mashaal Musleh**, Sohaib Ghani, and Mohamed F Mokbel. VisCAT: Spatio-Temporal Visualization And Aggregation of Categorical Attributes In Twitter Data. In *ACM International Conference on Advances In Geographic Information Systems, SIGSPATIAL*, 2014. [PDF]
  - **Mashaal Musleh**. Spatio-Temporal Visual Analysis For Event-Specific Tweets. In *ACM International Conference on Management of Data, SIGMOD*, 2014. [PDF]

- **Journal Papers**
  - **Mashaal Musleh** and Mohamed F Mokbel. Let's Speak Trajectories: A Vision To Use NLP Models For Trajectory Analysis Tasks. *ACM Transactions on Spatial Algorithms and Systems, TSAS*, 10(2), 2024. [PDF]
  - Ibrahim Sabek, **Mashaal Musleh**, and Mohamed F Mokbel. RegRocket: Scalable Multinomial Autologistic Regression With Unordered Categorical Variables Using Markov Logic Networks. *ACM Transactions on Spatial Algorithms and Systems, TSAS*, 5(4), 2019. [PDF]
  - Louai Alarabi, Mohamed F Mokbel, and **Mashaal Musleh**. ST-Hadoop: A Mapreduce Framework For Spatio-Temporal Data. *GeoInformatica*, 22(4), 2018. [PDF]
- **Workshop Papers**
  - Amr Magdy, Thanaa M Ghanem, **Mashaal Musleh**, and Mohamed F Mokbel. Exploiting Geo-Tagged Tweets To Understand Localized Language Diversity. In *Workshop on Managing and Mining Enriched Geo-Spatial Data, GeoRich@SIGMOD*, 2014. [PDF]

Selected Honors and Awards

- **ACM SIGMOD:** Best Demo Award, Seattle, USA, 2023.
- **ACM SIGSPPATIAL:** First-place award in the Student Research Competition (SRC), USA, 2022.
- **University of Minnesota:** Dedication to Student Learning Certificate, USA, 2021.
- **University of Minnesota:** Graduate Fellowship (USD $3600), USA, 2016.
- **Conference Travel Awards:** SIGSPATIAL 2023, SIGMOD 2023, SIGSPATIAL 2022, ICDE 2014, SIGMOD 2014.
- **Umm Al-Qura University:** First Rank Student Award (USD $2600), Saudi Arabia, 2014.
- **ACM SIGMOD:** Undergraduate Poster Finalist, USA, 2014.

Technical Skills

- **Languages & Frameworks:** Python, Java, Ruby, C#, C, C++, LLVM, VB.net, Ruby on Rails, JavaScript.
- **Databases & Data Science:** Sci-Kit Learn, Pandas, GeoPandas, PostgreSQL, MS SQL Server, MongoDB.
- **Platforms & Tools:** Linux, Docker, Ray, Hadoop, Redis, Git, Cloud (AWS, Azure, Google Cloud).

[*Last Updated on May 23, 2025*]