Louiqa Raschid

Louiqa Raschid has made pioneering contributions towards meeting data integration and data management challenges in multiple non-traditional domains including the life sciences, health information systems, Web data delivery, humanitarian IT applications, social media monitoring, and the next generation of data science for finance research. Her multi-disciplinary research spans the fields of computer science, information systems and data science.

She is a Professor at the University of Maryland, College Park. She received her primary and secondary education at Bishop's College and St. Bridget's Convent in Sri Lanka, and was ranked first in the 1973 island-wide General Certificate of Education Ordinary Level (Year 10) examination. She received a Bachelor of Technology degree from the Indian Institute of Technology, Madras in 1980, and an MS and Ph.D. degree from the University of Florida, Gainesville in 1982 and 1987, respectively, in electrical engineering.

Her research on tools for semi-automatic information extraction and mediation from Web accessible sources in the late 1990's was one of the first to recognize the importance of these non-traditional information sources. Today, following the wave that she partly created, such sources are ubiquitous. She initiated work, as early as in 2000, on cost-based semantic query rewriting for Web sources with limited query answering capabilities and developed wide area cost models for efficient query evaluation on the Web. These projects extended traditional database query optimization and evaluation technology to the Web and the cost model was the first to show that time and day variations, as well as network topological features, could be learned using query feedback, and exploited to construct a Catalog and prediction tool to predict access latencies.   Not only are both these contributions highly cited (among the most cited papers of the VLDB Journal), they were also in the vanguard of almost a decade of research as well as system prototype development on wide area optimization and heterogeneous cost models.   Ideas from this stream have eventually been incorporated into industrial products such as the IBM DB2 II information integration solution.

Her research on exploiting semantic knowledge to find patterns in graph datasets illustrates the breakthrough potential of the Linked Data initiative for biomedical applications. The research exploits algorithmic solutions including dense subgraphs, minimum description length (MDL) graph summarization methods and semantics based edge partitioning (semEP). Methods have been developed for link prediction and to identify communities and graph summaries that have both statistical significance and biological meaning. A key contribution is that a scientist can intuitively exploit semEP communities to build hypotheses, e.g., drugs with a specific function that can effectively target a gene family. semEP makes more novel predictions compared to competitors for drug target identification. semEP can create Drug-Pathway-Gene-Phenotype profiles that combine pharmaco-genomic communities with clinical data. Preliminary results have been promising for drug induced liver injury (DILI) scenarios and the research was a highlight of a recent 5 Million Euro EU H2020 grant.   This multi-disciplinary research has the distinction of being published in the premier conferences in multiple disciplines, including algorithms, information retrieval, bioinformatics, data mining and the semantic Web.

Raschid is a world expert in disaster information management. When disasters occur, multiple diverse organizations providing relief and recovery operations require immediate access to information from multiple complex dynamic heterogeneous sources. Raschid has led the Sahana FOSS project for disaster information management, serving as founding Board Chair, database architect and research

mentor. Sahana was initiated in Sri Lanka, in the aftermath of the 2003 tsunami. It is today the only comprehensive product for disaster information management that supports sharing of disaster data using open ontologies, standards and protocols. It was deployed for multiple disasters including the 2010 Haiti earthquake, hurricane Sandy and for refugee registration in Germany in 2015. UNESCAP and Sahana have teamed to develop SAMBRO: Sahana Alerting and Message BROker using the Common Alerting Protocol (CAP) standard to share situational awareness updates.

Turning from natural to financial disasters and the 2008 US Recession, Raschid has led an effort to identify data management and computational challenges of monitoring and modeling financial supply chains and global markets to better manage systemic risk. Because of her vision and intellectual leadership, her deep understanding of the computational, data management and modeling challenges, and the identified pressing need in light of the Dodd-Frank act, she was invited by the National Science Foundation to develop a data science for finance (DSfin) research agenda. Since 2015, she has teamed with the Office of Financial Research (OFR) and NIST to develop and run a very successful Financial Entity Identification and Information Integration (FEIII) Challenge. She is the founding chair for the ACM SIGMOD Workshop for Data Science and Macro-Modeling with Financial and Economic Datasets. She is co-PI of an NSF CASCADE Industry/University Center to develop DSfin datasets and tools. She has also been a prominent educator and evangelist in creating the DSfin multi-disciplinary community.

Her service to the ACM and IT community include the following:
- Editor in Chief, ACM Journal of Data and Information Quality, 2013 --. Recruited to increase outreach and awareness, to improve the management of the journal, and to elevate the journal impact factor. Introduced best practices to reduce the review cycle and worked closely with Associate Editors to improve the rigor and completeness of reviews. Expanded the base of submitting authors by reaching out across multiple venues to develop a strong multi-disciplinary group of authors and target audience. Introduced a very successful series of short Challenge papers that raised significant awareness of the journal across multiple ACM communities.
- Research envisioned by Raschid and colleagues (2003 Joint NSF / NIH Report, "Data Management for the Biosciences") was influential in the formation of NIH research agenda on this topic and paved the way for decades of advancement in biomedical informatics.
- Contributor to Grace Hopper Open Source Day, 2012-2014. Lead a Sahana team that assisted participants to install and explore Sahana code. Coordinated a team, together with the Google Crisis Response team to develop a blueprint for a Sahana CAP message editor to share data with Google Public Alerts services.
- Founding Co-Chair, Program Committee, ACM SIGMOD Workshop on Data Science for Macro-Modeling with Economic and Financial Datasets, 2014 and 2016. This workshop series has become the locus of an interdisciplinary research agenda that reaches multiple ACM communities.
- Invited Co-Chair of the 2010 NSF Workshop / Report on Knowledge Representation and Information Management for Financial Risk Management and the 2012 CRA/CCC Workshop / Report on the Next Generation Financial Cyberinfrastructure.
- Invited by the Office of Financial Research and the National Institutes of Standards and Technology to organize the Financial Entity Identification and Information Integration (FEIII) Challenge, 2015 to present.