Amr Magdy
Assistant Professor of Computer Science and Co-founder of the Center for Geospatial Sciences
University of California, Riverside (UCR)

Amr’s recent research focuses on scalability and expressiveness for spatial data science. He particularly emphasizes enabling social scientists to analyze large-scale spatial data using advanced spatial statistical analysis techniques. Traditional computational systems for spatial data have primarily addressed basic queries like spatial kNN, spatial range, and spatial join queries, including variations such as kNN-join queries. However, these queries fall short of addressing the comprehensive needs of spatial statistical analysis. Statistical methods are extensively utilized by social scientists, including geographers, environmental scientists, and ecologists. Consequently, over the past two decades, there has been a parallel development of computing tools, resulting in Python and R libraries like PySAL, GeoDa, GeoPy, and GeoPandas, catering to this broad user base. These users are well-positioned to conduct meaningful analyses on the plethora of available datasets. Notably, ~14 million datasets are indexed by Google Dataset Search, predominantly originating from geosciences and social sciences accessed by a third of all users and primarily dominated by spatial data. However, current tools encounter scalability issues with substantial spatial data, such as the vast amounts of social media data rich in insights for social scientists. Amr’s work seeks to fill this void by empowering social scientists with scalable computing systems for spatial statistical analysis on large datasets. Additionally, it enhances the expressiveness of queries, enabling previously unfeasible queries due to scalability constraints.

Amr's recent work spans various aspects of scalable and expressive spatial data science. He has significantly enhanced the scalability and expressiveness of spatial grouping queries, which are foundational for spatial statistical analyses. His improvements cover the grouping of spatial points, lines, and polygons across diverse settings and applications. Notably, he has developed advanced spatial regionalization queries, essentially sophisticated polygon grouping queries, that surpass existing methods with up to a 200x increase in runtime speed and the capacity to handle datasets up to 10 times larger. Additionally, he has facilitated the generation of random distributions, empowering the statistical assessment of spatial regionalization solution quality. Side by side, Amr's work integrates machine learning to augment both the scalability and accuracy of spatial queries. This integration encompasses a variety of learning techniques to refine spatio-textual selectivity estimation in dynamic streaming data, identify spatial hotspots along networks with statistical rigor, and enhance the efficiency of spatial joins in distributed settings.

Interdisciplinary projects, developed in partnership with social scientists, are at the heart of Amr’s research with collaborators at UCR, as well as the American Association of Geographers, Texas State University, San Diego State University, San Diego Mesa College, and the California Geographic Alliance. Amr’s research is published and recognized in top research venues for big data management and spatial data analysis, including ACM SIGSPATIAL, IEEE ICDE, VLDB, SSTD, IEEE MDM, VLDB Journal, ACM TSAS, IEEE TKDE, Springer GeoInformatica, in addition to interdisciplinary venues. His papers have been shortlisted for best paper awards at five major conferences for spatial and spatio-temporal data analysis, IEEE ICDE'14, twice at ACM SIGSPATIAL'19 and '23, IEEE MDM'23 and SSTD'23. He is also awarded the best paper at the 6th International Conference on Information and Communication Technologies for Disaster Management (ICT-DM'19). His research is being supported by eight research funding grants from the USA National Science Foundation (NSF), Google, and Microsoft, four as a Principal Investigator (PI), two as a co-PI, and two as a collaborator. This includes the NSF CAREER Award 2023, a distinction that ranks among the most prestigious recognitions for assistant professors from NSF. So far, he raised a total of $4.67M along with his collaborators, including $1.24M dedicated to supporting his research group.