Academic Profile

Amr Magdy

www.cs.ucr.edu/~amr

The content of my academic profile includes:

- Research Statement  (8 pages)
- Teaching Statement  (4 pages)
- Diversity Statement  (2 pages)
- Curriculum Vitae (CV)  (11 pages)
Research Statement

Amr Magdy

www.cs.ucr.edu/~amr

I joined the University of California, Riverside (UCR) as a tenure-track assistant professor of Computer Science and Engineering and a co-founder for the UCR Center for Geospatial Sciences (http://spatial.ucr.edu/people). This research center promotes interdisciplinary approaches for solving spatial-related problems leveraging spatial data analytics and geocomputation to pioneer new analytical techniques for computationally intensive geospatial planning and policy problems. This statement summarizes my research group contributions during my work at UCR (Sections 1 and 2) and my research philosophy (Section 3).

1 Research Overview and Accomplishments

1.1 Overview: My research endeavors focus on advancing spatial computing, aiming to develop innovative, scalable and expressive tools for spatial data science. I particularly emphasize enabling social scientists to analyze large-scale spatial data using advanced spatial statistical analysis techniques. Imagine a map with millions of data points, each representing different information like social media posts, environmental conditions, or population statistics. Social scientists, such as geographers and environmental experts, often need to analyze this kind of data to understand patterns and trends, like where a certain plant species is thriving or how people in different areas are discussing a topic on social media. Traditionally, the tools available for handling this type of data have been limited. They could perform basic tasks, like finding the nearest points of interest (like the closest hospitals to a given location) or identifying all points within a certain area. But when it comes to more complex analysis that social scientists need - like understanding how different areas compare in terms of environmental factors or how social media discussions vary across different regions - these traditional tools fall short, especially when dealing with vast amounts of data.

Traditional computational systems that handle large spatial datasets 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 in addressing the comprehensive needs of spatial statistical analysis. These methods are extensively utilized by social scientists. 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 indexed by Google Dataset Search, predominantly originating from geosciences and social sciences, constitute 45.2% of all datasets, accessed by a third of all users and primarily consisting of spatial data [9, 18, 12, 17]. However, current tools encounter scalability issues with substantial spatial data.

My work seeks to fill this void by empowering social scientists with scalable and expressive queries to support spatial statistical analysis on large datasets. Enhancing the query expressiveness is gifted by scalability improvements that enable previously unfeasible user-defined constraints to be considered. This work inherently includes interdisciplinary research projects, in addition to the regular computer science projects, developed in partnership with social scientists. This blend of interdisciplinary research is pivotal to the mission of the UCR Center for Geospatial Sciences, which thrives as an interdisciplinary cluster of faculty members.

1.2 Research funding: I got total of eight research grants, including the NSF CAREER award in 2023, a distinction that ranks among the most prestigious recognitions for assistant professors in USA. In particular, I have served as the Principal Investigator (PI) on four grants (two from NSF, one from Microsoft), co-PI on two NSF grants, and as a senior personnel on two NSF grants. The total fund raised in collaboration amounts to ~$4.67M, with ~$1.24M directly allocated to my research group.
1.3 Interdisciplinary research and networks: During my work at UCR, I devoted substantial time and effort to developing interdisciplinary networks and engaging in cross-disciplinary research, in line with the mission of UCR Center for Geospatial Sciences. Building such interdisciplinary connections presents several challenges. Firstly, establishing a common language of understanding among professionals from diverse disciplines requires lengthy cycles of meetings and discussions. Secondly, identifying mutually interesting research problems is time-consuming, as many proposed issues may lack novelty or relevance from a computing perspective. Thirdly, some identified research problems may not initially resonate with computer science conference reviewers, leading to extended periods of revisions and rejections before publication. Despite these hurdles, my research has greatly benefited from these networks, notably in undertaking impactful and practical research and securing diverse funding sources. Receiving funding from different grants supported by the USA National Science Foundation (NSF) underscores the significance and impact of my interdisciplinary research endeavors. I anticipate these networks will endure and continue to grow.

2 Research Highlights

As outlined in Section 1.1, my research focuses on advancing spatial computing, supporting scalable and expressive queries for spatial data analysis through interdisciplinary collaborations with social scientists. This is a broad umbrella that contains endless opportunities and challenges. Spatial data science performs diverse operations on different datasets and data types, depending on the underlying application.

The recent explosion in user-generated data has introduced significant computational challenges. Assume a social scientist analyzes 100M tweets in different city neighborhoods. This scenario includes a variety of spatial operations. First, tweets’ point locations are aggregated in different neighborhoods based on their source device, post language, or other attributes. Second, due to the low percentage of geotagged data, missing data will be estimated through spatial learning models, e.g., geographically-weighted regression. Then, the aggregated data could be used as derived attributes to group spatial neighborhoods in larger districts, i.e., spatial clustering of polygons. The city neighborhood themselves could be arbitrary defined based on user-defined constraints and aggregated by grouping line segments of road networks [20]. The tweets’ textual attributes will also be used with their temporal streaming nature to analyze textual content and estimate the stream behavior in incremental batches of new data as they arrive. So, analyzing such rich user-generated datasets involve a variety of learning, estimation, and grouping operations on different data types and attributes. However, the excessive data size along with the streaming nature makes it highly challenging to be supported through the existing techniques. My research addresses a variety of spatial learning, estimation, and grouping operations that empower social scientists to perform rich analysis on large-scale user-generated data.

2.1 Learning-assisted Spatial Analysis

My research exploits and boosts the power of learning in spatial applications through different spatial queries [15] that consider: (a) stable and streaming data, (b) spatial and textual aspects of the data, and (c) Euclidean and spatial network spaces.

2.1.1 Scalable Statistical Hotspots Detection on Spatial Networks

The spatial network is a popular spatial reference space, yet challenging due to the high computational cost of network distance that amplifies the runtime significantly compared to the regular Euclidean space. Due to its popularity and expressiveness as a spatial data model, there is no current replacement for spatial networks. Instead, existing spatial applications suffer from limited capabilities to support large-scale datasets. My group research explored learning hotspots in spatial networks to pinpoint areas within a network that exhibit a significantly higher concentration of objects than surrounding regions. This is part of a bigger picture that explores using learning as a way to avoid high-cost computations on spatial networks.

Existing hotspot detection methods are either clustering-based methods or statistical-based methods. While clustering methods have efficient runtime, they might result in false-positive results [22]. On the other hand, statistical methods offer rigorous statistical validation for the detected hotspots, e.g., Monte Carlo trials, and log-likelihood scores, ensuring the detected hotspots are statistically robust. However, such
methods do not scale to large-sized datasets due to the unacceptable execution time in statistical validation, which takes tens of minutes to run on tens of thousands of objects. This is particularly problematic in applications requiring fast response times, such as traffic hotspot detection. While social scientists prefer statistical-based hotspot detection methods due to their reliability, scalability becomes the limitation. For instance, the Chicago crime dataset includes 7 million objects, exceeding the capacity of existing statistical methods.

My group work scales up discovering hotspots using a localized expressive version of the network $K$-function that reveals the inherent statistical distribution of objects within a spatial network \[14\]. Unlike training an off-shelf machine-learning model, our learning paradigm fine-tunes the computation scalability of the $K$-function method to scale it up on large datasets. We address two problems: Hotspot Detection with Predefined Radius (HDPR) and Hotspot Detection Without Predefined Radius (HDWPR), ensuring a statistical confidence in both cases and offering automatic discovery of optimal radii of hotspots. Our techniques employ batch-processing strategies to optimize computations based on spatial proximity. They also use local maximum statistical confidence to discover hotspots on a localized scale. This work provides up to four orders of magnitude faster processing, and has been selected among best ACM SIGSPATIAL 2023 papers, and invited to ACM TSAS journal special issues.

2.1.2 Scalable and Expressive Spatial Prediction on Large-scale Datasets

Geographically-weighted regression (GWR) extends traditional regression analysis by modeling the spatial variability in the relationships among variables. It has been widely applied in diverse areas including transportation science and urban planning. Nonetheless, GWR suffers from limited expressiveness and scalability. The limited expressiveness arises from employing a uniform geographical weighting factor (bandwidth) across all features, neglecting the fact that different features may exhibit unique spatial scales. On the scalability front, the efficiency of GWR is constrained by its quadratic time complexity in training, which restricts the applicability to large spatial datasets. An adaptive variation, called multiscale GWR (MGWR), allows different scales for different features, yet, it multiplies the scalability problem.

My group work introduces Augmented GWR (A-GWR) \[21\] as an advanced GWR variant to handle large-sized spatial datasets with more expressive capabilities. A-GWR improves the scalability by innovating the Stateless-MGWR (S-MGWR), an adaptation of the MGWR model to efficiently optimize the model parameters in parallel and eliminate the need to store historical values. This enables GWR variants for the first time to flexibly and efficiency use black-box parameter optimization techniques. In addition, it enables using divide-and-conquer strategies to parallelize training on smaller, more manageable chunks of data without losing the spatial relationships among the data points. Such optimizations enables A-GWR to achieve up to 14.4 times faster compared to the state-of-the-art models on large spatial datasets, even with limited computing resources. To improve expressiveness, A-GWR combines the S-MGWR model with general-purpose machine learning models, such as random forests, to analyze complex non-spatial relationships within the data. This seamless integration with traditional machine learning models allows the spatial models to make use of the new advances in machine learning without the need to tailor every new model for spatial data.

2.1.3 Learning-assisted Estimation on Spatial Streaming Data

Selectivity estimation has been a topic of study in databases for a long time to produce fast and accurate estimates as a crucial part of finding a cheap query plan for query optimization. My group work used learning to tune selectivity estimation accuracy and efficiency on spatial streaming data. I particularly support cardinality estimation of spatial-keyword queries on spatial streams, e.g., first responders estimate the number of tweets with keywords ‘fire’ that lie within ‘Downtown Thousand Oaks, California’ for resource allocations. As high dynamism in query workloads is a quantified phenomenon in streaming user-generated data \[13\], traditional cardinality estimators become inaccurate for changing query workloads in real time.

Traditionally, machine learning models replace the underlying estimation data structures, e.g., histograms or samplers. My group work developed LATEST model \[19\] that does not replace the underlying data structures but recommends the most efficient data structure that boosts the estimation performance at different points of time. Therefore, the model does not answer the estimation query directly but decides which estimator to use for query answering based on learning from the past streaming activities. LATEST
employs an initial estimator and continuously monitors the estimation accuracy and performance along the stream lifetime. Once the query accuracy and/or runtime falls under certain thresholds, the Estimator Adaptor switches to the most accurate and efficient estimator for this part of the stream lifetime. LATEST model employs six different estimators and has shown a great adaptivity to optimize both estimation accuracy and query latency.

2.2 Spatial Grouping Queries

My group research addresses spatial grouping on different spatial data types to boost both scalability and expressiveness of spatial grouping queries. Grouping is a major category of spatial operations that are performed on all spatial data types, points, lines, and polygons to serve many spatial applications, e.g., population studies, market analysis, crime analysis, or landscape planning. However, the high computational cost limits its scalability and flexibility to allow user-defined constraints that empower rich analysis capabilities. I briefly outline below my group contributions to fill this gap.

2.2.1 Spatial Regionalization

Grouping spatial polygons has been studied under different names in the literature, most commonly known as spatial regionalization. Regionalization groups spatial areas (polygons) to find homogeneous spatial regions based on spatial contiguity and constraints on non-spatial attributes, e.g., average income, level of quality of life, or population density. The recent literature has significantly limited expressiveness and scalability that limit the data analysis capabilities. I briefly outline below my group contributions [10] to boost their expressiveness and scalability.

**Scalable and expressive p-regions.** The recent literature addresses two main variations of the regionalization problems: the \( p \)-regions problem and the \( \text{max-}p \)-regions problem. The \( p \)-regions problem requires the user to input a pre-defined number of regions to be detected \( p \). \( \text{Max-}p \)-regions problem automatically discovers the maximum number of regions \( p \) that satisfy the user constraints. Existing variants of the \( p \)-regions problem lack support for user-defined constraints, limiting its applicability across various domains and a wide range of use cases. My group work introduced \( PRUC \) problem [10] that defines a more expressive version of the \( p \)-regions problem to support additional user-defined constraints. \( PRUC \) comes with scalability challenges as existing techniques have a significant probability, up to 80%, of generating regions that fall short of the user-specified constraints. This mandates extensive spatial rearrangements of areas that exacerbates scalability issues, making it more challenging to handle large datasets. We developed the Global Search with Local Optimization (GSLO) algorithm [16] to support \( PRUC \) efficiently on large data. GSLO innovates techniques to derive a region partitioning that aligns with the user-defined constraints, boasting a high success probability with low spatial rearrangements. GSLO is two orders of magnitude faster and achieves up to six times better quality solutions compared to state-of-the-art algorithms for both \( PRUC \) and the original \( p \)-regions problem.

**Scalable and expressive max-\( p \)-regions.** Unlike the \( p \)-regions problem, the \( \text{max-}p \)-regions (MP-regions) problem does not require explicit specification of the number of regions but requires users to specify a single constraint on an attribute with a lower bound threshold, e.g., each region has a total population \( \geq 30K \). This constraint-based approach allows for the meaningful definition of spatial regions based on attribute values, rather than requiring the explicit specification of the number of regions. Unfortunately, MP-regions is an NP-hard problem and scalability issues have historically hindered the effective use of MP-regions, particularly when dealing with a large number of geographical units. In addition to the scalability issue, the MP-regions formulation is also limited in expressiveness because it only supports a single constraint with a lower-bound threshold. However, many real-world analysis tasks require multiple constraints with different aggregate functions.

My group work boosts both scalability and expressiveness of MP-regions problem [11, 12]. For scalability, we addressed several challenges to parallelize building regions and propose spatial partitioning that respects region contiguity constraints. Such parallelization achieves 97% reduction in query time and extends its support to an order of magnitude larger datasets than state-of-the-art approaches. For expressiveness, we extended the MP-regions and introduced \( EMP \) formulation to support the five SQL aggregation operators: MIN, MAX, AVG, SUM, and COUNT. \( EMP \) supports multiple constraints within a singular
query, with a range operator, enabling representation of both lower and upper bounds. The new formulation introduce several challenges, including the non-monotonicity of MIN, MAX, AVG aggregates, that amplify the complexity of finding a solution for such an NP-hard problem. Our FaCT algorithm addresses these challenges through careful arrangement for handling constraints and exploiting the mathematical properties to eliminate infeasible solutions and optimize the objective value. FaCT algorithm processes datasets multiple times larger than the supported ones in the literature, demonstrating notable scalability while boosting the MP-regions expressiveness significantly.

**Statistical inference for spatial regionalization.** My group work built on the above contributions to enable for the first time statistical inference to assess the quality of regionalization algorithms. As NP-hard problem, all regionalization algorithms produce approximate solutions without any intuition about closeness from optimality. We developed an algorithm that efficiently produces reference distributions of solutions for a given instance of a regionalization problem, generating tens of solutions in much less runtime than the existing state-of-the-art. Such a reference distribution is essential for any statistical process that gauges randomness, and includes new challenges in both scalability and region-level constraints that are not addressed in the literature.

### 2.2.2 Spatial Group-by Queries

Grouping spatial points and lines are widely used in different spatial analysis operations, e.g., polygonization, harmonization, segregation analysis, hot-spot and cold-spot analysis, or counting-based spatial statistics. However, involving very large dataset sizes, e.g., 100M social media points and 152M road network segments, demonstrates very inefficient runtime for traditional techniques even on distributed frameworks, e.g., Spark, that reaches tens of minutes or even breakdown.

To enable social scientists to perform large-scale studies of world-scale data, my group work built scalable distributed spatial group-by query modules for points and lines. The modules group scattered line segments into polygons (polygonization) and group points inside each polygon efficiently using low-overhead distributed data processing. Our modules significantly optimizes point-in-polygon checks on real complex polygons, with thousands of perimeter points, through two-level clipping to collectively optimize a set of point group-by queries. They also provides a cost-based query optimizer that handles the high skewness of real spatial data to switch between the traditional and new grouping techniques based on the data and queries distributions. This work provides up to two orders of magnitude faster processing, and has been selected among best SSTD 2023 papers, awarded IEEE MDM 2023 best paper runner-up, and invited to Springer GeoInformatica journal special issues.

### 2.3 Ongoing Research

My group ongoing research spans two directions. The first direction continues to build on our contributions supporting social scientists in analyzing large spatial datasets. This involves several pieces of work, including: (1) building a unified framework to support a wide variety of spatial regionalization queries using high-level declarative languages, (2) developing a two-level learning-based hotspot detection system to scale up statistical hotspot discovery on larger data, (3) scaling up spatial prediction techniques through the incorporation of stateless parameter optimization, (4) innovating a general scalable graph partitioner that imposes strict cardinality constraints on the graph partitions, and (5) integrating the developed technology into education to support underrepresented populations and improve their social mobility.

The second direction of my group ongoing research explores more use cases that exploit learning-assisted data analysis in enterprise applications. For this direction, I work with both Google and Microsoft on the following two use cases:

**Learning-assisted spatial join.** I work with one of the major Google teams to improve spatial join scalability using machine learning. In enterprises, it is common to run the same query workloads repeatedly. In such common cases, the data and query distributions either do not change or encounter minimal changes. The main idea of this work is to eliminate repetitive work by learning the right decisions from the historical runs of similar or identical workloads. This applies to two steps of spatial join, the spatial partitioning step and the indexing step, on distributed systems like Spark. Our approach incrementally maintains a gallery of partitioner and indexing templates that is updated over time based on new workloads. If the incoming
workload is similar to a previous workload, the partitioning and indexing times are cut down significantly by learning from the gallery to reuse a good enough partitioner and index that were effectively used before. We are currently modeling the query and data distributions as a set of features and similarity metrics to effectively and efficiently handle a variety of large-scale spatial join workloads.

**Edge-based spatial data analysis.** I work on a Microsoft-funded project in collaboration with computer vision researchers and domain experts in smart agricultural tools and aviation. The aim is to revolutionize smart surveillance and autonomous systems by enabling edge camera devices to run AI model inference directly on the edge devices using distilled models. The work involves developing data management features on edge devices to ensure quick access to relevant compact data, effective and efficient edge device coordination with both cloud and edge terminals, and the integration of multi-modal data for better resource utilization and higher inference accuracy on edge devices. The project includes several enterprise use cases in airport safety, fruit picking, and crowd management.

### 3 Research Philosophy

Reflecting on my journey since 2011, starting with my Ph.D. years at the University of Minnesota – Twin Cities and continuing as a principal investigator at the University of California, Riverside (UCR), my research has been ever motivated by (a) solving real problems and (b) educating the next generation of intellectual leaders in the field. These two elements contribute to forming my academic personality and my never-ending search for the next adventure in the student-educator journey and the scientific community. Below, I briefly outline a few core research beliefs that drive my innovations:

1. **Impactful and Practical Research:** I always search for a “customer” who is interested in using the solutions I am trying to innovate. So, my research problems have always been driven by a practical use case motivated and confirmed by this customer, who could be a significant end user or a collaborator from industry or another scientific discipline. This started in 2012, in the early years of my Ph.D., when I collaborated with the Microsoft Bing team through Microsoft Research to solve the problem of producing fresh search results from Twitter data as a side panel for Bing search results. This triggered all the research challenges I addressed to support Microblogs data management in my Ph.D. thesis. The outcomes of this thesis contributed to a well-established startup company for social media analysis called Lucidya, which has been running for over a decade. After joining UCR, I worked with a different type of customers, specifically social scientists, who analyze spatial data. Currently, I am crafting new collaborations with teams in the industry, forming new interesting research problems. Although this takes a longer time to converge, it is always rewarding in several ways, including having my research employed in real system deployments and crafting networks of real users.

   I developed such a core research belief early as a member of the data management lab at the University of Minnesota while pursuing my Ph.D. I realized how practical research that is realized inside real working systems makes the researcher aware of every potential problem that might prevent the proposed techniques from being useful to end users. In our research group, we were keen to conduct practical research and realize it inside real big data systems.

2. **Effective Academic-Industrial Collaborations:** A branch of conducting impactful and practical research is getting involved in fruitful industrial collaborations. In today’s world, a considerable portion of big data research problems are generated by the big industry players as they own the big datasets that motivate many of the new research problems. A big separation in the mindset between academic and industrial researchers is about the role of each in the innovation pipelines, which always puts questions on defining how an effective academic-industrial collaboration looks like.

   My position on that separation as an academic researcher is that academia is much easier to maneuver than industry, with cheaper resources and more agility of outcomes. Our talented Ph.D. students are high-quality engineers who can try out and realize different potential solutions for a given problem. Also, the nature of academic programs tolerates a high probability of failure if the solutions do not work realistically. Such a high probability of failure is very costly in the industry, has to get a stack of approvals, and usually cannot be tolerated within tight-budget teams. Still, young, talented Ph.D. students do not have access to the large-scale computing and data resources that are available in the industry. So, an effective academic-
industrial collaboration makes use of the lightness of academic partners with the large-scale resources of industrial partners to innovate proof-of-concept solutions in academia and deploy them at scale in industrial systems. Such collaborations can be accomplished through joint projects in which students work on them during the academic year and deploy the successful results in industry Summer internships.

(3) **Effective Interdisciplinary Collaborations:** Another branch of conducting impactful and practical research is getting involved in fruitful interdisciplinary collaborations. The big data era we are witnessing has changed the way scientists are doing their research in different disciplines, e.g., astronomy, medical sciences, social sciences, and food sciences. In many cases, scientists face new challenges that have not been addressed before by existing data analysis technology. For a data analysis researcher, those challenges are new opportunities to extend the impact of big data technologies to other scientific disciplines. This is valuable for the scientific community as a whole and helps to advance the cutting edge knowledge of humanity. I had core collaborations with social scientists around the USA during my work at UCR to analyze large-scale spatial data and was involved in collaboration with food scientists at the University of Minnesota to identify foodborne illnesses using social media data. Such rewarding experiences greatly advance the researcher’s mentality, maturity, and skills to manage such collaborations effectively and diversify the funding resources. Effective management includes tackling several challenges, such as establishing a common language among professionals from diverse disciplines, identifying mutually interesting research problems that are novel, and resonating these problems with computer scientists. Tackling such challenges comes with the reward of undertaking impactful research and securing diverse funding sources from different interested agencies.

(4) **Scalability Opens New Opportunities:** In big data analysis research, achieving scalability for new queries is not only a publishable research achievement but also an opener for new opportunities to support more expressive versions of such queries that were not possible before. Looking ahead for such opportunities is key in enriching the data analysis capabilities of data practitioners and scientists from different disciplines to deepen their understanding and potential to harness their data. The data science researcher should always be on top of these opportunities to open new lanes of collaboration and advance cutting-edge technology.

(5) **Research is an Impressive Educational Tool:** An academic researcher is also an educator who influences graduate students’ lives through mentoring and training to become independent thinkers and intellectual leaders in their field. This takes long shots of introducing and polishing several key skills, both technical and soft skills. This educational part of the research journey makes the academic setting much easier to maneuver than industry researchers because even failed research trials still contribute to student education and polishing their skills.

**References**


Teaching Statement

Amr Magdy

www.cs.ucr.edu/~amr

The academic life in my perspective is all about impacting students’ lives positively through high-quality education. Our students in turn will impact the society’s present and future. Thus, teaching and mentoring students are key motivations for me to join academia. This statement outlines my teaching philosophy, experience, and interests in computer science education.

Teaching Philosophy

This section briefly outlines the core values of my teaching philosophy. The subsequent sections highlight how these core values are reflected in my actual teaching and mentoring experiences, providing specific examples. My teaching philosophy is outlined by the following core values:

1. **Building a student-centered learning environment.** Teaching is about helping students to learn. Therefore, I always aim to maximize students learning outcomes by: building an active learning environment that supports diverse learning styles and inclusive teaching practices, incorporating students feedback, and adapting flexibly to students’ needs.

2. **Stimulating enthusiasm for learning.** Enthusiasm is essential for fruitful learning that maximizes students’ outcomes. For that, I make efforts to motivate students about the course content in innovative ways, including passion transfer, inquiry-based learning, continuous reinforcement of key motifs, and assigning achievable challenges and reflective activities.

3. **Engagement and adaptability with institutional mission.** Different institutions have different goals for their teaching, mainly based on the student population and the institution mission. Therefore, I carefully understand my institution mission to serve it effectively. I adapt my teaching style, methods, and strategies to serve the student population and the institution goals.

4. **Providing strong foundations for life-long learning.** Since Computer Science is a rapidly evolving and dynamic field, a computing technology learner has to acquire life-long knowledge to adapt with the new technologies. For that, I train students to develop self-learning abilities through acquiring knowledge independently and practicing critical thinking and digital literacy skills.

5. **Trading off fairness and opportunities.** In my teaching practice, I believe that an educator must be fair, provide equal opportunities to all students, and fairly evaluate and appreciate their efforts. However, fairness should not be exaggerated to strictness that prevents giving excuses and accommodations when needed. The ultimate goal is to give students opportunities for learning, which is the main objective.

6. **Personalized mentoring.** It is a common case that a subset of students often has different needs and/or capabilities. Therefore, I customize some of the course assignments and materials to flexibly accommodate their needs. This practice is more effective in reasonably-sized classes than large classes, typically less than 40 students.

7. **Innovative teaching methods.** In my teaching, I use innovative teaching methods, and continuously improve them, to create more engaging, effective, and enjoyable learning experiences and reduce the learning burden for students. Up to date, I use flipped classrooms, design problem-based learning (PBL) and collaborative learning activities, and mentor experiential learning assignments and projects.

Teaching and Mentorship Experiences

1. **Teaching experience and professional development.** Guided by the core values of my teaching philosophy, I have taught nineteen offerings of five different courses at the University of California, Riverside (UCR) as a tenure-track assistant professor (2017-Present), including: (1) three undergraduate courses covering an “Introduction to Data Structures and Algorithms”, “Design and Analysis of Algorithms”, and “Senior Project in Big Data Analysis”, and (2) two graduate courses covering “Spatial Computing” and “Database Management Systems”. Before joining UCR, I co-taught one offering of an undergraduate “Database Management Systems” course at the University of Minnesota (UMN) as a lecturer (Spring 2016). I also significantly contributed to teaching fourteen offerings of nine undergraduate courses as a tenure-track teaching assistant at Alexandria University, Egypt (2008-2011), including “Computer Networks”, “Pattern Recognition”, “Programming Languages and Compilers”, “Advanced Data Structures”, “C/C++ Programming”, “Introduction
Throughout my teaching experience at UCR, UMN, and Alexandria University, I have mentored several graduate and undergraduate students in different capacities during my work at UCR. I have advised seven Ph.D. students, and I chose the RIDLE track “Design for Multiple Modes of Learning” for Learning Engagement program, a year-long initiative offering workshops and activities centered on “Rethinking Instructional Design for Learning Engagement”. I chose the RIDLE track “Design for Multiple Modes of Learning” to enhance my skills in crafting online and hybrid courses. This track emphasizes classroom technologies and optimal pedagogical practices for diverse audiences, considering factors like room capacity, course waitlists, student persistence, and accessibility.

In addition to teaching students, I delivered tutorials in top conferences (IEEE ICDE and ACM SIGMOD), many invited talks, and professional presentations, tutoring the research community about my research. Conference tutorials had over 50 attendants ranging from graduate students to highly-recognized researchers, including ACM fellows and leading researchers. I received positive feedback from several audience members after my tutorials and talks. Most notably and recently, the chairman of ACM SIGSPATIAL approached me after my last presentation saying “I wonder if your paper presentation is this clear, how your teaching is”, a testimony that I am proud of among many others.

(2) **Curriculum development.** Throughout my teaching experience at UCR, UMN, and Alexandria University, I have significantly contributed to developing curricula for various courses and training programs, including regular semester-long courses, short summer courses, lab summer trainings, extended tutorials, and intensive programming contest trainings. At UCR, I have developed or contributed to developing the curriculum for all the courses I taught. As a tenure-track teaching assistant at Alexandria University, I have developed or contributed to developing content for the following courses: *Computer Networks, Pattern Recognition, C/C++ Programming*, and *Introduction to Web & HTML*. Additionally, as a TA, I developed assessments, including homework exercises, programming assignments, lab assignments, quizzes, and exams for all courses I contributed to at UMN and Alexandria University.

The most recent and notable curriculum development experience was creating a course from scratch at UCR covering spatial computing topics. I identified a gap: UCR is the closest major campus to ESRI headquarters, a leading company in GIS software, still, did not offer strong spatial-related education. I began by defining measurable learning objectives that embed spatial concepts in different layers of the data analysis stack. The course was divided into modules, starting with an overview, followed by in-depth modules on core spatial data models and operations, and breadth modules applying spatial concepts to various data analysis topics, such as data management, data mining, visualization, and network data. The final part of the course covers advanced and recent topics and engages students with innovative teaching methods like flipped classrooms, group collaborative work, and problem-based teaching. I use various teaching tools, including slides, videos, hands-on exercises, and in-class exercises, to engage students fully. I have been recently approached by ACM Books to write a book covering the material of this course.

The spatial computing curriculum has been improved over the years and has become more popular, with the highest student enrollment for a graduate course in my department. The content, assessments, and delivery methods of different modules have evolved based on student feedback and the teaching staff’s experience. At some point, we tried offering it as an interdisciplinary course to serve other relevant departments, such as Earth Science, though we still face challenges and are working on them. On a positive note, the curriculum modules balance breadth and depth without overwhelming students and keep up with new advances in the field. I always invite external and industry speakers to communicate with the students.

Introducing the spatial computing course has served my institution’s mission and goals to strengthen this area at UCR, which was the main trigger behind founding the UCR Center for Geospatial Sciences. The course contributes to different data science programs at UCR, serving both undergraduate and graduate levels. Creating this course was a unique opportunity to provide student-centered learning, stimulate their enthusiasm, and apply all my core teaching values.

(3) **Mentorship experience.** In addition to my teaching experience, I have mentored several graduate and undergraduate students in different capacities during my work at UCR. I have advised seven Ph.D. students,
twenty MSc students, and six undergraduate researchers. Three new Ph.D. students are joining my group in Fall 2024. Among these, I have already graduated three Ph.D. students and fifteen MSc students. All of my Ph.D. graduates have pursued academic careers post-graduation: one as a postdoctoral researcher at the University of Illinois at Urbana-Champaign (UIUC) CyberGIS Center for Advanced Digital and Spatial Studies, and two assistant professors in their home country. A handful of my MSc students have embarked on careers with major software companies. Notably, I played a key role in supporting Zacharias Chasparis, one of my MSc students, toward a position at ORACLE Corporation, a global leader in database technology. Furthermore, I mentored three individuals and one team of undergraduate and high school students on science projects, undergraduate research, AP capstone projects, and the FIRST LEGO League project. One of my undergraduate Hispanic mentees, Aldo Estrada, secured a summer internship at Amazon.com, Inc. based on our work, and another, Eric Ong, achieved the distinction of publishing in the VLDB journal, a premier venue for database research. A high school mentee secured third place in the Riverside County Science Fair competition. This mentorship demonstrates my commitment to nurturing young talent and fostering an interest in technology and research from an early age.

Before joining UCR, during my Ph.D. at the University of Minnesota (UMN), I mentored Mashaal Musleh from being an undergraduate student in Saudi Arabia until he received several Ph.D. offers from prestigious US universities, including the University of Wisconsin, Madison, Duke University, Brown University, the University of Maryland at College Park, Purdue University, the University of Minnesota, and Arizona State University. Mashaal’s work was published in the challenging ACM SIGMOD’14 undergraduate poster competition and awarded 4000 USD as a travel grant to the conference from Saudi Arabia. Mashaal also continued his excellent work and co-authored several publications with me in ACM SIGSPATIAL’14, IEEE ICDE’15, and ACM SIGMOD’14 workshops. I mentored another undergraduate student at UMN, in collaboration with food scientists, who was awarded a research fellowship for a full academic year. I have also had several mentoring experiences in Egypt before my Ph.D. I directed the ACM Alexandria Student Chapter and mentored 17 programming contestants who received 11 medals in the fourth Egyptian Olympiad in Informatics (EOI 2006), and advised Mohamed Elfeki to join his dream research group for the Ph.D. program, currently, he is a Senior Research Manager at Microsoft.

(4) Diversity and inclusion. I have dedicated significant efforts to supporting diversity and inclusion in computer science education, with a focus on serving female and Hispanic underrepresented minorities. I have mentored nine female students, including two Ph.D. students, six master’s students, and one high school student. Additionally, I have mentored two Hispanic undergraduate researchers and co-mentored a Hispanic Ph.D. candidate. Currently, I am holding a Google-funded grant to support a Hispanic undergraduate researcher. Since 2018, my diversity contributions have extended to two CSforAll projects focusing on broadening participation in computing, serving the Hispanic community, particularly in the Sweetwater Union High School District located in Chula Vista, California. The projects aimed at enhancing geospatial course instruction for high school and college students, integrating computer science technology. In 2019, I presented a segment of my work on teaching development at RESPECT 2019, the 4th International Conference on Research in Equity and Sustained Participation in Engineering, Computing, and Technology, concurrently held with ACM SIGCSE, the leading conference in computer science education.

(5) Reflections of core teaching values. During my teaching and mentorship experiences, I have had the opportunity to apply my core teaching values as outlined in my teaching philosophy. My classes provide student-centered education. I conduct an entry survey in the first week to understand the background and expectations of my students and adjust the course syllabus or class discussions accordingly. This is not a one-time process; instead, I release an anonymous feedback Google form that continuously receives students’ feedback throughout the semester without requiring them to reveal their identity. I also support students with different learning styles by diversifying the delivery methods and types of in-class activities. I engage students with group work and think-pair-share activities, giving a problem or a question and asking small groups to interact and solve them collaboratively, e.g., in database design, data structures, and algorithms. Additionally, I use slides, videos, and in-class exercises to make lectures more interactive. For example, I have used a comedy video about Mathematics to motivate Relational Algebra in databases and an illustrative video for map projections and orientations. This approach has been particularly successful in the spatial computing course due to the various available resources. I asynchronously share videos with students to visualize different concepts, such as sorting, map projection, insertions, and deletions in data structures.
accommodate students’ needs, I allow resubmission policies for partial credit and handle their requests and legitimate excuses flexibly, particularly students with special medical needs.

To stimulate students’ enthusiasm for learning, I employ various practices. First, I build my own passion for the topic to reflect it in the class. In many cases, I naturally have this passion, and in some cases, I make efforts to build it in different ways. Several times, students note in their evaluations that “Prof. Magdy is very enthusiastic about the topic” or similar comments. I also incorporate an inquiry-based learning strategy. I prepare and formulate the key questions that should be addressed by the curriculum, gradually rolling them out as interactive in-class exercises, repeating and emphasizing them to stick in students’ heads. This continuous reinforcement of key motifs maximizes learning outcomes. Additionally, I relate key motifs to real-life applications and career objectives. For example, I relate database transactions to many interesting applications that students use every day and algorithms to job interview questions. I also stimulate more enthusiasm by assigning achievable challenges and reflective activities, helping students recognize their growing understanding of the key motifs using hands-on exercises. Furthermore, I regularly invite industry experts, when appropriate, to share their professional experience with the students. These have been very successful experiences based on students’ testimonials.

I always serve my institution’s mission and goals. At UCR, social mobility and supporting underrepresented populations are key goals. The Hispanic population represents the major underrepresented group at UCR and in the Southern California region. I actively support and engage with this group, as well as female students, in different capacities, as elaborated in the diversity section. My course design takes into consideration their lack of easy access to high-speed internet, so I limit heavy online activities, such as virtual lectures, unless necessary. I also limit mandating marginalized students to purchase materials as much as possible and point out free or cheap alternatives utilizing university resources. For transfer students from community colleges, I provide extra revision material and help them catch up on what they missed, ensuring personalized mentoring. Based on our work, my Hispanic mentees, Aldo Estrada and Eric Ong, secured a summer internship at Amazon.com, Inc. and published in the VLDB journal, a prestigious research venue.

As life-long learning skills is essential for our students, I train my students to acquire knowledge independently through open exercises for which they are responsible for finding answers outside the course material. I also work on improving their critical thinking and digital literacy skills by, for instance, organizing the lesson as a story that embeds a chain of “why?” and “what are the alternatives?” questions. This helps them assess the quality and credibility of online resources and reason about decisions, e.g., database system module design or algorithm design.

I maintain a fair environment for all students to have equal opportunities to learn, yet I understand that fairness is not equivalent to strictness and some students need special attention. I provide flexibility for assignment deadlines when needed for personal or documented medical reasons. I also offer extra instruction, either sessions or material, for underprepared students, such as transfer students from community colleges, or extraordinary students who can afford more learning loads. This reflects the student-centered environment and personalized mentoring that I am keen to maintain and practice. I consistently receive positive student feedback regarding maintaining both fair and flexible courses.

Teaching Interests
My teaching interests are diverse and span both core computer science and engineering courses, as well as advanced and specialized courses. My interests include courses that cover programming at different levels, data structures, design and analysis of algorithms, database systems, data science courses of different types, spatial computing, distributed data analysis, computer networks, compilers, operating systems, digital logic design, capstone projects, and seminar courses that introduce new topics. For both undergraduate and graduate teaching, I am open to teaching new courses based on the needs of the department and college, whether they fit within my expertise or I have an interest in preparing for them even if I have not taught them before. This is one of my core values in engaging with the institutional mission and goals. I am also eager to create new courses, if feasible, that open new horizons for students, connect them with the latest technology, and train them for effective job hunting after graduation. I have already created a new course at UCR, namely “CS225: Spatial Computing,” which has become one of the most popular courses in terms of student enrollment in a few years. Through newly created courses, I have the opportunity to serve and apply several of my core teaching values, including providing student-centered learning, fostering life-long learning skills, and stimulating students’ enthusiasm.
Diversity Statement

Amr Magdy
www.cs.ucr.edu/~amr

I am a professor of Computer Science and Engineering at the University of California, Riverside (UCR). I have dedicated significant efforts to supporting diversity and inclusion in computer science education, with a focus on serving female and Hispanic underrepresented minorities. My work has included mentoring students, securing grants for underrepresented groups, and leading projects to broaden participation in computing.

Social Mobility at UCR

As a professor at UCR, I am proud to be part of an institution that is a national leader in social mobility and a major Hispanic Serving Institution (HSI). UCR’s commitment to diversity, equity, and inclusion (DEI) is reflected in its significant support for low-income students, first-generation college students, and underrepresented minorities, particularly within the Hispanic/Latinx community. UCR has received a No. 3 national ranking and No. 1 among public universities on a Chronicle list of universities with the highest graduation rates for institutions where at least 50% of the students receive Pell Grants; a US federal financial aid for college tuition of students with exceptional financial need. This highlights our exceptional performance in supporting low-income students and ensuring their academic success. Our Pell Grant recipients graduate at a slightly higher rate than non-Pell Grant recipients, showcasing the effectiveness of our support systems and dedication to equitable outcomes.

As a Hispanic Serving Institution, UCR plays a crucial role in enabling Hispanic/Latinx students to pursue and succeed in STEM fields. The Inland Empire region, where UCR is located, is larger and more populous than almost half of the U.S. states but has historically faced educational disparities. UCR’s efforts are pivotal in addressing these challenges and providing educational opportunities to underrepresented minorities. Currently, Hispanic/Latinx students make up 41.5% of the campus-wide population and 30.4% in the College of Engineering. This increasing representation over the past five years is a testament to our inclusive and supportive environment. In addition to Hispanic/Latinx students, UCR also serves a diverse student body with 33.8% of students identifying as Asian campus-wide and 40.8% within the College of Engineering. First-generation and low-income students respectively account for 56.6% and 44.7% of the student body. This diversity enriches the educational experience and fosters a multicultural and inclusive campus environment.

UCR’s strategic initiatives, including academic advising, financial aid, mentorship programs, and career services, are designed to support our diverse student population. These programs are crucial in helping students navigate and overcome the challenges they face, ensuring that they not only enter college but also graduate and succeed in their careers. By being at the forefront of social mobility and diversity, UCR exemplifies how a public university can transform lives and communities. This commitment aligns perfectly with my philosophy as an educator and researcher, aiming to contribute to an inclusive academic environment where every student has the opportunity to succeed and make a meaningful impact on society.

My Contributions to Diversity at UCR and Southern California

At UCR, I have actively contributed to social mobility and supporting underrepresented populations through both existing programs and new initiatives in education and research. One of my primary contributions has been through mentoring female and Hispanic students in computer science. I have worked closely with nine female students, including two Ph.D. candidates, six master’s students, and one high school student, providing guidance and support to help them achieve their academic and professional goals. I have also mentored two Hispanic undergraduate researchers and co-mentored a Hispanic Ph.D. candidate. In fact, I
Amr Magdy Diversity Statement, 2024

I am actively participating in programs to support my Hispanic students and elevate their profile. Currently, I am holding a Google-funded grant to support a Hispanic undergraduate researcher. One of my undergraduate Hispanic mentees, Aldo Estrada, secured a summer internship at Amazon.com, Inc. based on our project, and another, Eric Ong, achieved the distinction of publishing in the VLDB journal, a premier venue for database research.

In addition to my involvement in existing programs to support UCR students, I have also taken the initiative to create new projects and collaborations to further support social mobility outside UCR and within the broader scope of Southern California region. Since 2018, I have been actively involved in securing funds for two CSforAll projects funded by the US National Science Foundation (NSF) focusing on broadening participation in computing for the Hispanic community. Particularly, the projects are in collaboration with educators from San Diego regions, from San Diego State University (SDSU), San Diego Mesa College, and the Sweetwater Union High School District in Chula Vista, California. These projects aim to integrate computer science technology to enhance geospatial course instruction for high school and college students. This work not only helps to increase the representation of Hispanic students in STEM fields but also provides them with the skills and knowledge necessary to succeed in their future careers. In 2019, I presented a segment of my work in these projects at RESPECT 2019, the 4th International Conference on Research in Equity and Sustained Participation in Engineering, Computing, and Technology, concurrently held with ACM SIGCSE, the leading conference in computer science education.

In 2024, I started crafting a new collaboration with experts in digital accessibility to broaden my diversity contributions by supporting people with disabilities in accessing geospatial maps effectively. This initiative aims to create inclusive geospatial tools that accommodate users with various disabilities, ensuring that everyone can benefit from advancements in geospatial technology. By integrating accessibility features such as screen readers, alternative text, and simplified navigation, this project will make geospatial data more usable for individuals with visual, cognitive, and motor impairments. This collaboration reflects my commitment to expanding the reach of my diversity efforts and ensuring that technological advancements are inclusive and accessible to all.

Through these collaborations, I have worked on developing and implementing innovative teaching practices and geospatial tools that make computing technology more accessible and engaging for diverse students and community members. These efforts have been instrumental in increasing the representation of Hispanic and female students in STEM fields at UCR and beyond. By actively contributing to UCR’s mission and goals, I strive to create a supportive and inclusive academic environment that empowers all students to succeed and make a positive impact on society.

Earlier Stage and Continuously Supporting Diversity

Before joining UCR, a major diversity experience in my life began in 2010 when I moved to the USA as a Software Engineer Intern at Google Inc. and pursuing my Ph.D. degree at the University of Minnesota, which is a significantly more diverse environment compared to my hometown. The USA values diversity and embraces the virtue of accepting different people as a core value. Interacting with such a community enriched my personality and exposed me to many people from various backgrounds and attitudes. At Google, I interacted with interns from around the globe, which helped me to understand different mentalities and cultures, allowing me to learn the best from each. At the University of Minnesota, I lived in a diverse community of thousands of students from all parts of the world had a profound impact on my personal development.

In my teaching and research, I incorporate these diversity experiences to enhance cultural diversity within my community. Teaching in diverse environments has enabled me to effectively handle students from various backgrounds, transferring cross-cultural knowledge and raising diversity awareness among them. In research, I aim to build a strong, diverse research group that attracts talented students from around the globe, promoting collaboration to solve global problems. This long-term interaction provides an excellent opportunity to transfer cross-cultural knowledge and understanding, offering distinguished students from underserved and poor communities the chance to pursue graduate studies and improve their lives through high-quality education.
Amr Magdy
Department of Computer Science and Engineering
University of California, Riverside, USA
Office: 328 Winston Chung Hall
Office Phone: (+1) 951-827-9223
Email: amr@cs.ucr.edu
Webpage: http://www.cs.ucr.edu/~amr
Cell Phone: (+1) 651-605-5039
Fax: (+1) 951-827-4643

Research Interests
Big data management, spatial computing, database systems, spatio-temporal data management, large-scale data analytics, indexing, and main-memory management.

Education
Ph.D. in Computer Science
Computer Science and Engineering Dept., University of Minnesota – Twin Cities
July 2017
Minneapolis, USA
• Thesis focus: Scalable Microblogs Data Management
• Thesis advisor: Mohamed F. Mokbel

M.Sc. in Computer Science
Computer Science and Engineering Dept., University of Minnesota – Twin Cities
April 2013
Minneapolis, USA

M.Sc. in Computer and Systems Engineering
Computer and Systems Engineering Dept., Alexandria University
July 2011
Alexandria, Egypt
• Thesis focus: Density-based Clustering of Data Streams with Heterogeneous Densities
• Thesis advisors: Nagwa El-Makky and Noha A. Yousri

B.Sc. in Computer and Systems Engineering
Computer and Systems Engineering Dept., Alexandria University
July 2008
Alexandria, Egypt
• Cumulative grade: Distinction with Degree of Honor (92.67%)
• Thesis focus: Highly Interactive Flash-based Online Classroom
• Thesis advisors: Magdy Nagi and Khaled Nagi

Research and Professional Experience
• University of California, Riverside, Computer Science and Eng. Dept., CA, USA
  Tenure-track Assistant Professor
  Jul 2017 - Present

• University of Minnesota, Computer Science and Engineering Dept., MN, USA
  Research Assistant
  Sep 2011 – Jun 2017
  Advisor: Mohamed F. Mokbel

• Microsoft Research, Redmond, WA, USA
  Research Intern
  Jun 2014 – Aug 2014
  Mentors: Sameh Elnikety, Yuxiong He

• GIS Technology Innovation Center, Makkah, Saudi Arabia
  Research Assistant
  Sep 2013 – May 2014
  Mentor: Mohamed F. Mokbel

• Microsoft Research, Redmond, WA, USA
  Research Intern
  May 2012 – Aug 2012
  Mentors: Sameh Elnikety, Yuxiong He

• Google Inc., Mountain View, CA, USA
  Software Engineer Intern
  Jun 2010 – Sep 2010
  Mentor: Mohamed G. Elfeky

• Microsoft Research, Cairo Advanced Technology Lab, Cairo, Egypt
  Research Assistant
  Feb 2009 – Jun 2010
  Mentor: Nayer Wanas

Awards and Recognitions
• Shortlist of ACM SIGSPATIAL 2024 10-Year Impact Award (equivalent to Test-of-time Award).
• ACM SIGMOD Distinguished PC Member Award 2024.
• Microsoft Research Award 2024.
• NSF CAREER Award 2023, by USA National Science Foundation.
• Google-CAHSI Research Award 2023.
• Shortlist of best ACM SIGSPATIAL 2023 research papers, invited to ACM TSAS journal special issue.
• Shortlist of best of SSTD 2023 research papers, invited to Springer GeoInformatica journal special issue.
• Best paper runner-up in IEEE MDM 2023.
• Best paper award in ICT-DM 2019.
• Shortlist of best ACM SIGSPATIAL 2019 research papers, invited to ACM TSAS journal special issue.
• NSF CRII Award 2019, by USA National Science Foundation.
• Doctoral Dissertation Fellowship, University of Minnesota, 2015.
• Shortlist of best of IEEE ICDE 2014 research papers, invited to IEEE TKDE journal special issue.
• Finalist in Microsoft Research Ph.D. Fellowship 2014.
• Best demonstration award in UMN U-Spatial Symposium 2012.
• Prof. Naem Aboutaleb Award for academic excellence – Alexandria University – 2008 (awarded to the top two students among all the department’s 67 students).
• Bronze medals in the 3rd and 4th Egyptian Olympiad in Informatics (EOI 2005 and 2006).
• National Scholarship for Academic Excellence, Egyptian Ministry of Higher Education – 2003 (awarded to the top 0.1% of students nationwide).

Research Fund
• PI: Google CAHSI Award: SUDeC: A Scalable Framework for User-defined Spatial Clustering Queries on S2 Spherical Coordinate System. PI: Amr Magdy, $99,993, 9/1/2023-8/31/2024.
• Co-PI: NSF Award SES-1831615: RIDIR: Scalable Geospatial Analytics for Social Science Research. PI: Sergio Rey, Co-PIs: Ran Wei, Amr Magdy, Vassilis Tsotras, $1,000,000, 10/1/2018-9/30/2023.
• Co-PI: NSF Award CNS-2031418: Collaborative Research: Encoding Geography - Scaling up an RPP to Achieve Inclusive GeoComputational Education. PI: Coline Dony, Co-PIs: Atsushi Nara, Sergio Rey, Amr Magdy, Michael Solem, Rachel Russell, $999,979, 1/1/2021-12/31/2024.
• Senior Personnel: NSF Award IIS-2123444: DS-PATH: Data Science Career Pathways in the Inland Empire. PI: Mariam Salloum, Co-PIs: Analisa Flores, Xinping Cui, Paea LePendu, Vassilis Tsotras, $1,025,000, 10/1/2021-9/30/2024.

Teaching Experience
• University of California, Riverside, Computer Science and Eng. Dept., CA, USA Fall 2017 - Present
  Tenure-track Assistant Professor
• University of Minnesota, Computer Science and Engineering Dept., MN, USA Spring 2016
  Instructor
  Responsibilities: Delivering lectures in database systems for the course “CSCI 5708: Architecture and Implementation of DBMS”.
• University of Minnesota, Computer Science and Engineering Dept., MN, USA Spring 2012
  Teaching Assistant
  Course: Introduction to Operating Systems.
• Alexandria University, Computer and Systems Engineering Dept., Egypt Fall 2008 - Summer 2011
  Tenure-track Teaching Assistant

Page 2 of 11
Last updated July 15, 2024

Selected Invited Talks
- Geospatial Data Science Series, University of Wisconsin, Madison, February 2024.
- Geospatial Talks, Google, Madison, Wisconsin, February 2024.
- CEMSE Seminar, King Abdullah University of Science and Technology (KAUST), Saudi Arabia, October 2023.
- Qatar Computing Research Institute (QCRI) Talk, Qatar, October 2023.
- Doha Data Forum for Inclusive Data Ecosystems and Effective Decision Making, Qatar, October 2023.
- Colloquium, School of Computer Science and Engineering, Nanyang Technological University (NTU), Singapore, July 2023.
- Guest Lecture, Department of Earth and Planetary Sciences, University of California, Riverside, April 2023.
- Geographical Information Systems and Spatial Databases Lectures, University of Maryland, College Park, March 2023.
- Colloquium, Computer Science Department, Iowa State University, March 2023.
- Lighting Talk, University of California GIS Week, November 2021.
- Colloquium, Department of Statistics, University of California, Riverside, October 2021.
- Colloquium, Computer Science Department, University of California, Santa Barbara, October 2019.
- SoCal Social Analytics Workshop, UC Irvine, May 2018.
- Data Science Seminar, UC Riverside, October 2017.

Professional Service and Activities

Professional Memberships
- IEEE Member:
  - Senior Member: 2024-Present
  - Professional Member: 2017-2024
  - Student Member: 2012-2017
  - Member of the IEEE Computer Society
- ACM Member:
  - Professional Member: 2017-Present
  - Student Member: 2007-2017
  - Senior Member application under review
  - Member of the ACM SIGSPATIAL and SIGMOD

Research Community Service
- Founding Chair: ACM SIGSPATIAL Recognition Committee. The committee is responsible for SIGSPATIAL awards. The committee reports to the SIGSPATIAL executive committee.
- Guest Editor: Springer Geoinformatica Journal, Frontiers in Big Data Journal
- Organizer/Co-organizer:
  - Student Research Competition (SRC) chair: ACM SIGSPATIAL 2023 and 2022.
  - Registration chair: SSTD 2019.

• **Proceedings chair**: ACM SIGSPATIAL Workshops LENS 2017 and GeoEd 2019.

• **ESRI Spatial Data Science Workshop at UCR**.

• **Peer-Reviewer and Program Committees**:

  • **Grant Proposals**:
    - USA NSF (2018, 2023)
    - Canada NSERC (Discovery Grants 2017, New Frontiers in Research Fund, NFRF, Exploration 2019, 2021, and 2022)
    - Puerto Rico Science, Technology & Research Trust (2019)

  • **Journals**:
    - VLDB Journal (VLDBJ, 2016-Present)
    - ACM Transactions on Database Systems (ACM TODS, 2017-Present)
    - IEEE Transactions on Knowledge and Data Engineering (IEEE TKDE, 2018-Present)
    - ACM Transactions on Spatial Algorithms and Systems (ACM TSAS, 2018-Present)
    - Springer GeoInformatica (2017-Present)
    - Springer Journal on Distributed and Parallel Databases (Springer DAPD, 2018-Present)
    - Elsevier Journal on Future Generation Computer Systems (FGCS, 2018-Present)
    - Frontiers in Big Data (2018-Present)
    - ISPRS International Journal of Geo-Information (ISPRS IJGI, 2018-Present)

  • **Conferences**:
    - ACM International Conference on Advances in Geographic Information Systems (ACM SIGSPATIAL, PC Member 2018-2022, Senior PC Member 2023-Present)
    - ACM International Conference on Management of Data (ACM SIGMOD, 2022-Present)
    - The International Conference on Very Large Databases (VLDB, 2023-Present)
    - IEEE International Conference on Data Engineering (IEEE ICDE, 2020-Present)
    - International Symposium on Spatio-Temporal Databases (SSTD, 2019-Present)
    - IEEE International Conference on Mobile Data Management (IEEE MDM, PC Member 2019-2023, Senior PC Member 2024-Present)
    - IEEE International Conference on Big Data (IEEE Big Data, 2024-Present)
    - International Conference on Scientific and Statistical Database Management (SSDBM, 2018-Present)
    - IEEE BigData Congress (2018-Present)

**University and Departmental Service**

- Graduate Admission Committee, CS Department, 2022-Present.
- Faculty Hiring Committee, external member, Department of Statistics, 2021-2022.
- TA Allocation Committee, CS Department, 2020.
- Data Science Faculty Committee, CS Department, 2020-Present.
- Committee for Spatial Analysis minor, cross-colleges, 2019.
- Reviewer, contributor, and presenter in several research and educational programs and activities, including women in engineering activities, UCR mini-grants, Chancellor's research fellowship, Cal-Bridge, CS4All code camp, ABET evaluation, grad open house, student orientations, workshops, panels, and department colloquium.

**Community and Public Service**
• Tutorial for K-12 teachers and college and university geography professors from San Diego State University and Mesa College, San Diego, on Using ChatGPT for Generating Web-based Interactive Maps, July 2023.

• Lesson segments to incorporate Geospatial Concepts in CS teaching, in collaboration with middle computer science teachers from Sweetwater Union High School District, San Diego, July 2023.

• Lesson segments to incorporate Computational Thinking in Geography Education, in collaboration with geography teachers from Sweetwater Union High School District, San Diego, June 2023.

• Tutorial for K-12 teachers in Riverside, on Using ChatGPT for Improving K-12 Teaching and Tutoring, June 2023.

• Introduction to Computational Thinking to K-12 Teachers, California Geographic Alliance, April 2022.


Pre-tenure Service

• **External Reviewer:** VLDB’12-’17, ICDE’12-’16, SIGMOD’12-’17, SIGSPATIAL’14-’16, SSTD’13, GeoInformatica, TKDE, EDBT’17, MDM’12-’13, SSDBM’12-’13, and ICDCS’12.

• **Student Volunteer:**
  - NSF Workshop on Social Networks and Mobility in the Cloud, Washington DC, 2012.

• **Webmaster:** PersDB 2012 Workshop

• **Advising Mentor** for three years in Alexandria ACM Chapter Mentorship Program. Mentee has joined a Ph.D. program in a distinguished US research group. (2013-2016)

• Microsoft Student Partner in Alexandria University during the academic year 2007/2008.

• Alexandria ACM Student Chapter, 2005-2008 (program committee, treasurer, and course coordinator).

• Mentored 17 students in the 4th Egyptian Olympiad in Informatics (EOI 2006); 11 got medals.

Professional Meetings


• NSF Workshop on GeoSpatial Software: Connecting Big Data with Geospatial Discovery and Innovation. NSF Geospatial Software Institute, University of Southern California, Los Angeles, CA, January 2018.

Selected Research Publications

The complete list of major journal and conference publications is available at: [https://dblp.dagstuhl.de/pid/05/8073-1.html](https://dblp.dagstuhl.de/pid/05/8073-1.html)

Journal Papers

1. Abdulaziz Almaslukh, Yongyi Liu, **Amr Magdy**. Scalable Spatio-temporal Top-k Interaction Queries on Dynamic Communities. In the ACM Transactions on Spatial Algorithms and Systems (TSAS), 2024.


Conference Papers


**Tutorials**


**System Demonstrations**


**Book Chapters**


**Patents**


**Edited Proceedings**


**Technical Reports**

Workshop Papers and Others


Student Advising

Ph.D. Students

- Laila Abdelhafeez, 2018-Present (Co-advised by Vassilis Tsotras).
- Yongyi Liu, 2020-Present.
- Alhassan Satii Alshareedah, 2023-Present.
- Yiyang Bian, 2024-Present.
- Ahmed Abdelmaguid, 2024-Present.
- Yunhan Chang, 2024-Present.
• Dr. Yunfan Kang, 2018-2023, Postdoc Researcher at the University of Illinois, Urbana-Champaign (UIUC)
• Dr. Hussah Alrashid, 2018-2023, Assistant Professor at Majmaah University
  Thesis: Towards a Scalable and Qualitative Spatial Regionalization.
• Dr. Abdelaziz Almaslukh, 2017-2020, Assistant Professor at King Saud University
  (Advised 2015-2017 by Vagelis Hristidis)
  Thesis: Scalable Temporal Queries on User-Generated Data.

M.Sc. Students
• Vishal Chinnam, 2024-Present.
• Tushar Baid, 2024-Present.
• Nikhil Mahendrakar, 2024-Present.
• Jatin Patwa, 2021.
• Yash Deshpande, 2021.
• Kexin Wang, 2021.
• Pranshu Shrivastava, 2021.
• Shamali Shinde, 2021.
• Anish Sekar, 2020.
• Ziang Zhao, 2019.
• Aparna Vivek Sarawadekar, 2024-Present.
• Harsha Devarapalli, 2024-Present.
• Priyanka Jadli, 2023.
• Rithika Ramasesha, 2021.
• Siddharth Shenoy, 2021.
• Zhuocheng Shang, 2021.
• Shashank Dahiya, 2021.
• Faisal Almaarik, 2021.
• Suchitra Pithavath, 2020.
• Zacharias Chasparis, 2018.

Undergraduate Students
• Aldo Estrado, 2023-2024, Intern at Amazon.
• Azka Khan, NSF Data Science Fellow, 2023.
• Dallas Carlson, NSF Data Science Fellow, 2023.
• Eric Ong, 2018-2019, Published in VLDBJ.
• Tiffany Cedeno, 2018.
• Abdelrahman Solyman, 2018.
• Samarth Srinivasa, NSF Data Science Fellow, 2023.
• Makena Grigsby, NSF Data Science Fellow, 2023.
• Celvin Lizama Pena, exploreCSR, 2021.
• Bai Carl, 2018.
• Hau Chen, 2018.

High and Middle School Students
• Marc Encarnacion, 2019-2020, AP Capstone project, Valley View High School, Moreno Valley, CA.
• Emma Shah, 2019-2020, Martin Luther King High School, won third place in RCOE Science Fair Expo.
• Dragon Rider Sean, 2019, Robosub team, FIRST LEGO League.