RFP MIDAS-HS: **Health Data Science Challenge Thrust**

One of the primary missions of the University of Michigan Institute for Data Science (MIDAS) is to foster innovative and groundbreaking multi-disciplinary research in data science across the three U-M campuses. MIDAS will launch major research initiatives in data science in the following four thrust areas:

1) Transportation Science;

2) Learning Analytics;

3) Social Science;

4) Health Science.

Up to $10M of funding will be awarded to proposals in these Challenge Thrusts based on a competitive process involving submission of a white paper, followed by invitation to submit a full proposal. A total of eight (8) proposals will be funded at a level of approximately $1.25M each. A successful research proposal will involve a multi-disciplinary team engaged in research that will both have disruptive impact on a relevant thrust application and significantly advance the methodological foundations of data science. The ultimate intent of the MIDAS challenge initiatives is to stimulate research activities that can be leveraged into successful external funding proposals from government, private foundations, or industry.

Examples of topics that could be addressed in a **Health Data Science Challenge** proposal might include one or more of the following:

- Integrated personalized omics and profiling (genome, epigenome, proteome, metabolome, etc.) and molecular phenotyping applied to health, wellness, and disease at the individual and population levels.

- Microbiome ecological modeling, metagenomics, and population health.

- New breakthroughs in basic biomedical, clinical and translational research enabled by Data Science.

- Novel data-driven modeling and analytics to leverage the Federal Government's Precision Medicine, Cancer "Moon Shot", and BRIAN Neuroscience Initiatives.

- Physiological phenotyping in Critical Care and/or in the field using wearable health sensor networks to support clinical decision making.

- Longitudinal analysis of Electronic Health Record (EHR) data, including unstructured text, imaging, environmental, and demographic data using novel data science methodologies (Machine Learning, NLP, etc.).

- Building a data-driven and adaptive Learning Health System demonstration project incorporating novel analytics and decision support.

- Mining of health outcomes data from a variety of sources such as CMS records, OptumLabs, Federal Government Drug Safety Records (FDA, CDC), Framingham and Jackson Heart Studies, ISR Health and Retirement Study, etc.

- Pharmacogenomics and Toxicogenomics studies at population scale.

- Bio-behavioral assessment and epidemiologic outcomes.

- Data mining and predictive analytics of self-reported health and behavioral information from social network resources such as Twitter, PatientsLikeMe, etc.

- Health disparity studies in underserved populations using heath, econometric and demographic big data.

- Applied to: Cancer, Cardiovascular Disease, Obesity, Diabetes, Alzheimer's Disease, Depression/Bipolar Disease, Autism Spectrum Disorders (ASD), etc., including co-morbidities.

- Breakthrough Allied Health Data Analytics (in Nursing, Dental, Pharmacy, Kinesiology, others).

- Ethical issues surrounding the use of Biomedical Big Data, including patient and research participant communication and consent aspects.

A successful proposal must develop methodological approaches complementing a compelling **Health Science** application, leveraging heterogeneous data at scale for novel date-enabled Public Heath and Medicine research.

Data Science methodological examples are as follows:

- Massive integration and harmonization of complex heterogeneous data.

- Novel health data de-identification and privacy methods.

- Scalable active learning and causal inference.

- Human-in-the-loop learning and analytics.

- Analytics enabled by statistical randomization models.

- Adaptive anomaly detection.

- Dimensionality reduction for visualization of complex data.

- Embedded systems for data mining and statistical inference.

- Distributed, cloud-enabled and interoperable algorithms.

- Bayesian approaches for big data.

These are only a sample of possible approaches that could be considered.

Timeline:

- Mar. - Apr. 2016:   information and team-forming sessions (Time and location TBD)
- June 30, 2016:   white papers due
- July 22, 2016:   full proposal solicitations communicated
- Oct. 17, 2016:   full proposal due
- Nov. 18, 2016:   awards announced

White Paper Format & Requirements:

White papers should not exceed five (5) pages (excluding budget and biographical sketches) and be structured as follows:

Title page (1 page). Proposed title, abstract, names and affiliations of co-PI's and senior investigators, and contact information for the PI.

Technical description (maximum four pages). Problem to be addressed and technical approach to solve problem. Nature of data to be collected/analyzed/managed. Methodology to be applied and analytical tools to be used or developed. Data Science Services and computational infrastructure to be used, e.g. CSCAR and UM Flux or cloud resources. Description and justification of team, including partners from industry or other institutions (cannot be part of budget). Expected impact of research resulting from the project.

Draft budget. Approximately $1.25M total for three years can be requested and should be broken down yearly. The Associate Deans for Research (ADR) of all colleges or schools in which the co-PIs and senior investigators hold their primary appointments should be sent a copy of this white paper.

Biographical sketches. Two page bios of each of the co-PIs and senior investigators should be attached. These bios should have a paragraph stating the strengths that the co-PI brings to the proposed project, a reverse chronological list of the co-PI's education and employment history, honors and awards, and a list of the 5 most relevant publications to the effort, a list of other related publications, and other support. While there is no specific format requirement, either NSF or NIH formatted biosketches would be acceptable.

To submit white papers: midas.umich.edu/rfp

Full Proposal Format & Requirements: (by solicitation only)

Full proposals should not exceed 11 pages (excluding references, budget, biographical sketches, and letters confirming cost sharing) and be structured as follows:

Title page (1 page). Proposed title, abstract, names and affiliations of co-PI's and key investigators, and contact information for the PI.

Technical volume (maximum 10 pages). Sec. 1 Technical description. Sec. 1.1 Problem to be addressed and the challenges faced. Sec 1.2 Nature of data to be collected/managed/analyzed. Sec. 1.3 Technical approach proposed to solve problem, including methodology to be applied and analytical tools to be used or developed. Sec. 1.4 Expected impact on technology, science, and society. Sec. 2 Resources. Sec. 2.1 Databases or data collections to be used, and IRB, FERPA, and HIPAA issues to be addressed, if applicable. Sec 2.2 Data science services and computational resources to be used, including CSCAR and UM Flux or cloud resources. Sec 3 Data management and dissemination plan. Sec. 4 Description and justification of team, including partners from industry or other institutions (cannot be part of budget).

References. (no page limit).

Budget. A detailed budget (totaling approximately $1.25M total for three years) broken down yearly and showing the cumulative budget, in addition to a budget justification. Note that there will be no indirect costs.

Biographical sketches. Two page bios of each of the co-PIs and senior investigators should be attached. These can be updated versions of those included in the white paper.

Letters confirming cost sharing.

For Ann Arbor participants: ADRs of the schools and colleges in which the co-PI's and senior personnel hold primary appointments should provide letters confirming a commitment to cost sharing, collectively covering 50% of the total requested budget for the Ann Arbor component of the proposed effort.

For Dearborn and Flint participants: No letters are required; UM will work with representatives from these campuses' research office and colleges to assess the availability of funding.

Review process and review criteria

The white papers and full proposals will be reviewed by a panel of experts. The panel will review each proposal according to the following criteria: 1) relevance to the stated thrust area; 2) impact or likelihood of the project to result in innovative creation and/or application of data science methodology; 3) complementarity to existing projects at UM and contribution to the development of reusable and lasting data science capacity on campus; 4) multi-disciplinary coherence of team; 5) likelihood that proposed work will lead to significant and sustained extramural funding. 6) substantial involvement of students.

Post-award expectations

Projects participants are expected to be strongly engaged with MIDAS and all co-PI's and senior investigators are expected to become active affiliate members of MIDAS. In addition, all teams will be expected to: 1) submit yearly reports on progress towards the aims of their grant; 2) participate in a yearly review, organized as a workshop for all co-PI's on all projects funded under the MIDAS Challenge Thrusts; 3) maintain an active project website; 4) actively work with MIDAS to enhance data science at U-M, e.g., through hosting DS student interns, and participating in targeted industry outreach. It is expected that all grant-derived publications acknowledge MIDAS.

Contact Information:

midas-rfp@umich.edu
MIDAS: http://midas.umich.edu
RFPs: http://midas.umich.edu/rfp