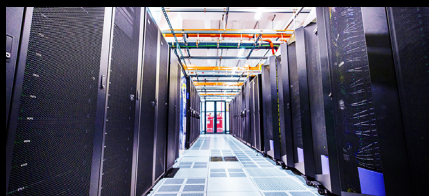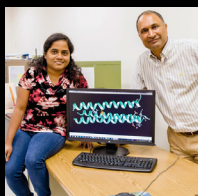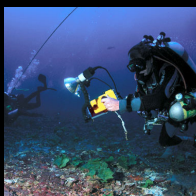# UNIVERSITY OF MIAMI
INSTITUTE for DATA SCIENCE & COMPUTING

SERVICES
+ RESOURCES

Who Can Access IDSC?

IDSC services and resources are available to everyone: Students, faculty, researchers, the public and private sector, nonprofits, industry, and beyond.

Our collaborative approach brings together talented minds at the interface of disciplines to harness the University's AI-ready supercomputer and 5G+ edge computing environment.

As a member of the University of Miami's Frost Institutes of Science and Engineering, IDSC is focused on utilizing the extraordinary potential of data science to tackle society's greatest challenges.

From creative student projects to innovative industry-shaping ideas, IDSC services and resources can help you achieve your research, training, or business goals.

idsc.miami.edu

IDSC
SERVICES
+ RESOURCES

Triton Specs

IBM Power9/Nvidia Volta (6 Racks)

IBM Declustered Storage (2 Racks)

96 IBM Power 9 Servers

30TB RAM (256/node)

1.2 Pflop Double Precision

240 Tflop Deep Learning

64 bit Scalar

100 GB/sec Storage

150TB Shared Flash Storage

400TB Shared Home

2 @ 1.99TB SSD Local Storage

Advanced Computing · · · · · · · · · ·

Through IDSC, the University of Miami maintains one of the country's largest centralized academic cyberinfrastructures. Rated one of the top 5 US academic institution supercomputers in 2019, Triton—the U's newest supercomputer acquisition—was designed to maximize data movement between the IBM POWER9 CPU and attached accelerators like GPUs. Triton represents a completely new approach to data science and computing.

IDSC's Advanced Computing core supports a large data storage system with well over 10PB of current "live" data, and more than 1,700 users. This cutting-edge high-performance environment enables researchers to simulate real-world experiments within a virtual lab, and makes computationally intensive tasks faster.

AI + Data Science Capabilities

- **Advance System Services** Self-Serve to Concierge Services, Development to 24x7 Non-Stop Operations, On-Premises and Cloud

- **AI/Machine Learning** Platform/Algorithm Selection and Implementation

- **Artificial Neural Networks** Specialized for different purposes, such as: Convolutional Neural Networks (image recognition), Standard and Recurrent Neural Networks (pattern detection, classification, prediction), Graph Neural Networks (for network optimization and prediction)

- **Cybersecurity** End-to-End Solution Design, Implementation, Support

- **Data Engineering** Identifying, Integrating, Cleaning, ETL, Processing

- **Data Management** Pipeline Services from Ingest to Presentation and Preservation

- **Fuzzy Logic** Applying logic to problems where the concepts of true/false are defined in terms of degrees and not as complementary alternatives. Python libraries will be used.

U MIAMI IDSC

AI + Data Science Capabilities (cont'd...)

- **Genetic Algorithms** The iterative application of heuristic techniques to simulate solutions of complex optimization problems. The Python language will be used.

- **Neuro-Fuzzy Systems** Applying neural networks to classification problems where logic elements are fuzzy in nature. The Python language will be used.

- **Research and Predictive Analytics** Audio Research, IoT, Environmental Monitoring, Security Services

- **Rule Extraction Algorithms** The process of extracting a rule base from empirical observations as an alternative to predicting outcomes. Coding will be done in Python.

- **Scientific Programming** Code Porting, Optimization and Parallelization Support Vector Machines

Services + Resources

» Machine Learning + Deep Learning Services

- **Distributed Deep Learning** When deep learning problems are too large for a single graphics card and separating jobs into discrete parts would affect the math adversely, distributed deep learning spreads calculations over multiple GPU cards. This process is analogous to MPI on traditional CPUs.

- **IBM PowerAI Vision** Offering intuitive tools to label, train, and deploy deep learning models for computer vision without coding or deep learning expertise.

- **Large Model Support** Allows the successful training of deep learning models that would otherwise exhaust GPU memory.

- **Watson APIs** Focused in finance, governance, and healthcare, enabling access to the full power of Watson from a local installation.

» Machine Learning + Deep Learning Frameworks

- **BVLC/caffe+ IBMCaffe** for vision, speech, and multimedia projects (2.0 brings focus to mobile and low power compute). As its full name "Convolutional Architecture for Fast Feature Embedding" indicates, Caffe is fast and suitable for conventional CNN applications.

- **PyTorch** is one of the latest deep learning frameworks. It is famous for simplicity and flexibility with dynamic tensor computations while maintaining efficiency. Great for rapid prototyping and research.

- **SnapML** An efficient, scalable, machine learning library for training linear models in finance and governance, developed at IBM Research Zurich specifically to address scaling issues in commercial clouds.

- **TensorFlow** is designed to provide an end-to-end platform for production and scalability. IDSC's optimized version includes visualization tool TensorBoard and the easy-to-use Keras API.

- **TensorFlow Keras** is a high-level API used for building and training deep learning models, and offers fast prototyping.

- **Time Series Analysis and Anomaly Detection** Time series data represents a collection of quantity that varies in time such as audio, video, financial indices, and metrics tracking data. The analysis of the internal struct of the time series data using techniques such as autocorrelation, autoregression, and recurrent neural networks could reveal the trend, provide the forecast and set alarm to the anomaly events.

» Data Engineering Services

- **Data Preparation and Cleaning** The process of reformatting, correcting and enriching data prior to analysis via Python libraries such as NumPy & Pandas, as well as Jupyter Notebooks.

- **Fully Secure Data Services** Blockchain attestation available.
  - IBM Mainframe hardware encryption exceeding US DoD standards
  - Extract, transform, and load data to the supercomputer graphically
  - Example Data Layers supported: Apache Drill, Storm, Spark, MQTT

Services + Resources (cont'd…)

- **Integrated Data Presentation/Persistence Services**
  - Common Data Access Services Supported: HTTP(s), FTP, SSH, bbcp, Aspera etc.
  - Common Web Platforms/Stacks Supported: LAMP, MEAN, Django, Node-JS, React, etc.
  - Common Data Stores: PostgreSQL (commercial support available), MongoDB, Maria/MySQL, Redis
- **Interactive Analytics** Familiar tools at scale—AutoHPC

» Data Pipeline Systems

- Faster, Cheaper & More Reliable than Cloud Technologies
- Flexible and Secure System Architecture
- User Interfaces on 2 Rockhopper II Mainframes (Hardware Encryption)
- Data store on ESS Parallel Systems (De-clustered RAID)
- AI Training and Inference on Triton
- Centralized Data Design (Multiple System Same Data)
  - No tiering costs, No transaction/transfer cost
- Globally Accessible Network (100GB/sec I2, 10GB/sec internet)

» High Performance Storage

The 7PB WADE Storage Cloud offers SaaS and PaaS architecture.

305.243.4962 | hpc@ccs.miami.edu
idsc.miami.edu/**platforms**/**advanced-computing**

Data Analytics

Making discoveries in oceans of data is possible only with the use of efficient algorithms. As data grows in volume, velocity, variety, and veracity, so does the demand for efficiency in data science applications. IDSC explores audio, bioinformatic data, images, spatial data, text, time series data, and video, and can offer insight into all your data projects.

Classification

Large-scale data analysis in data mining, pattern discovery, machine learning, and other algorithmic development is available to on- and off-campus researchers. Classification models for labeled data organizes your data into categories for the most effective and efficient use. (Jupyter Notebooks / Juptyerlab available; Python and R environments, Julia, and Erlang; R-Studio; Matlab)

Clustering Analysis

IDSC mines your data to cluster groups sharing similar patterns Common approaches include Self-organizing map (SOM) and k-means. Advanced techniques employed include:

- Concurrent Clustering of Data Points and Data Attributes
- Clustering with Constraints
- Subspace Clustering

Association Rule Mining

Association rules are if/then statements that help uncover relationships between seemingly unrelated data. IDSC uses techniques for discovering frequently occurring combinations of attributes to produce inferences based on the combinations thus discovered.

305.243.4962 | idsc@miami.edu
idsc.miami.edu/**platforms/services**

Traffic Signal
status: OK

Humans

Moving Vehicle
22 mph

Moving Vehicle
31 mph

Moving Vehicle
35 mph

Moving Vehicle
30 mph

Humans

Humans

Moving Vehicle
10 mph

Humans

Humans

Software Engineering

The IDSC Software Engineering team are professional software engineers who actively seek collaborative partners for new and innovative software application and systems development projects. The team has developed applications and software systems to support work in clinical research, drug discovery, genomics, mapping and urban planning, observational biology/ecology, and the digital humanities.

Core Competencies

- Software Application Conceptualization, Design, Implementation, and Project Management
- AI, Machine Learning
- Computer Vision, Photogrammetry, and 3D Reconstruction
- Drones and Aerial Survey with a Focus on Urban Spaces
- Extended Reality (XR) (augmeneted AR, Virtual, Mixed MR)
- GIS and Geospatial Application Development
- Multi-Tier Web-Based Application Development
- Participatory Mapping and Community-Centered GIS Development
- Search/Complex Data
- Semantic Web
- Spatial Surveying and Architectural Documentation

Software Engineering offers expertise for collaborators who need to include novel software systems as part of their projects, including expert personnel for inclusion in grant proposals, and the development of prototypes or initial analysis in preparation for proposal submission.

305.243.4962 | idsc@miami.edu
idsc.miami.edu/**platforms/sweng**

Visualization + Infographics ● ● ● ● ● ● ● ● ● ● ● ●

Data Visualization augments cognition to enable the exploration of complex data sets and illuminate scientific findings. IDSC's Visualization, Data Communication, and Information Design program regularly collaborates with scientists in communicating their results to stakeholders and to the general public using multimodal media— from static infographics to interactive web presentations, and other products that take advantage of technologies like augmented, mixed, and virtual reality.

Visualization and Infographics Consulting offers guidance on visual design, user experience, and which tools or tutorials are best to bring your design to life. Visit: idsc.miami.edu/vizconsult

Visualization Laboratory

The Viz Lab is a resource available by appointment to the UM community. Capable of accommodating 14 guests, the space offers an impressive venue to host a class or your next executive meeting. Located in the Arthur A. Ungar building Room #330D, with a direct connection to all UM computing resources, the Viz Lab features a Mechdyne 3D-Display and a Cyviz 2D-Wall. IDSC regularly offers demonstrations and orientation sessions for new users.

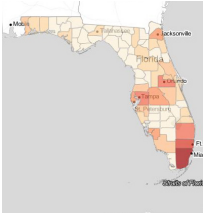To book a tour, or for more information, email vizlab@miami.edu.

305.243.4962 | idsc@miami.edu
idsc.miami.edu/**research/information-design**

GC IPCR -> IPFCDL, lag=20, r=0.008

GC IPFCDL -> IV1, lag=20, r=0.008

GC IPFCDL -> ... lag=20, r=0.008

GC IV1 -> IPFCDL, lag=20, r=0.008

Screen Shot 2020-03-06 at 12.00.53 PM

Region II

R II

R I

R III

IDSC
Collaborative Projects
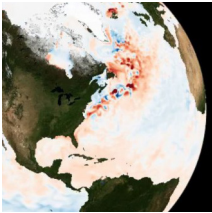

XSAT Covid-19/Outbreak Dashboard
xsat.idsc.miami.edu

IDSC researchers developed the xSAT eXperimental Situational Awareness tool, a first-of-its-kind early-detection platform that enables local residents to interact with infectious disease outbreak data and identify hot spots.

PUBLIC HEALTH APPLICATION | Advanced Computing, Software Engineering

High-Resolution Global Climate Modeling
idsc.miami.edu/wp-content/uploads/2019/10/hr_historical.mp4

A RSMAS research team employed a hierarchy of models with differing resolutions (from 100km to 50km in the atmosphere, and 1 degree to 1/10 degree in the ocean) to improve both the modeling and understanding of important fine-scale processes.

EARTH SYSTEMS MODELING | RSMAS, CIMAS

Santa Cruz del Islote, Colombia

Mapping the world's most crowded island was part of the documentation of informal settlements in Southern Hemisphere cities by UMSoA Professors Adib Cure and Carie Penabad, architects who seek to help transform unplanned cities into stable and legitimate neighborhoods.

DRONE MAPPING INFORMAL CITIES | School of Architecture, Cure and Penabad
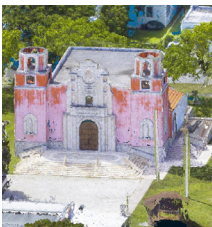
Nuestra Señora de la Merced Chapel Tour

Located on the campus of Corpus Christi Catholic Church in Allapattah, La Merced Chapel recreates the spirit of an old Peruvian church. IDSC worked with a team of art history students to create an immersive Magic Leap AR experience offering detailed information on the Chapel's interior, exterior architecture, and five of its Spanish Colonial paintings.

AUGMENTED REALITY | School of Art & Art History, College of Arts and Sciences

The U Experience, Magic Leap Tour

uexperience.wixsite.com/uexperience

Adding a digital dimension to real-world video, this Magic Leap campus tour combined real and virtual environments in an interactive XR application that complements freshman orientation. This capstone project was part of an MFA in Interactive Media.

EXTENDED REALITY (XR) | Office of Student Affairs, School of Communication

The Residency, Harbour Island, Bahamas

A group of UM students, faculty, and researchers evaluated the historical significance, potentional restoration, and future preservation of the historic Harbour Island Governor's residence in Dunmore Town. The team took aerial photos with drone mapping, and careated a point cloud, graphs, and drawings.

HISTORIC PRESERVATION | School of Architecture, Antiquities Monuments Museum Corp.
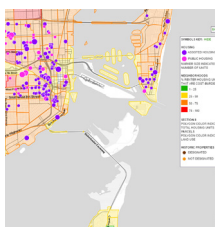
MAP The Miami Affordability Project

comte.ccs.miami.edu/housing/map

As part of the University's "Focus on Affordable Housing" initiative, the interactive MAP tool (a multi-tier web-based application) was designed to explore Miami's housing landscape, address needs, and promotes informed decisions on housing policy.

SEARCHABLE MAPS | Office of Civic and Community Engagement (OCCE)

URIDE

IDSC helped develop the University Research Informatics Data Environment (URIDE), a web-based platform that aggregates and visualizes de-identified patient data from the UHealth/UChart system, enabling the exploration of demographics, diagnoses, procedures, vitals, medications, labs, notes, allergies, comorbidities, locations, physicians, and more.

FREE TEXT SEARCHING | UHealth, Miami CTSI Biomedical Informatics

Mikvé Israel-Emanuel Synagogue, Curaçao

Located in the capital city of Willemstad on the island of Curaçao, Lesser Antilles, Mikvé Israel-Emanuel (the oldest continuously operating synagogue in the Western Hemisphere) was documented as part of a project to create an architectural record of the Caribbean's Jewish synagogues and temples.

COMPUTER VISION | School of Architecture ,Institute for Advanced Study of the Americas, Miller Center for Contemporary Judaic Studies, College of Arts and Sciences

NEW DEGREE.......

Master of Science in Data Science

This degree provides interdisciplinary connections and experiential learning opportunities across all aspects of data science and computing, from machine learning to marine science, city planning, or communications.

Students from any academic discipline are invited to explore advancing their careers with foundational knowledge on the applications and implications of data in a variety of fields.

For more information, please contact the Office of Interdisciplinary and Professional Studies at 305.284.8783.

Data Scientist

"The sexiest job of the 21st century."
—HARVARD BUSINESS REVIEW

UNIVERSITY OF MIAMI

msdatascience.miami.edu

UNIVERSITY OF MIAMI
INSTITUTE for DATA SCIENCE & COMPUTING

1320 S Dixie Hwy, Suite 600
Coral Gables, FL 33146-2930

305.243.4962 | idsc@miami.edu