



University Affiliate Centers to the Institute for Discrete Science

DHS Centers of Excellence

What is IDS?

The Institute for Discrete Sciences (IDS) is a partnership of the US Department of Homeland Security (DHS) and four University Affiliate Centers of Excellence. It was formed to study selected topics in the data sciences, machine learning, and discrete mathematics, that are collectively labeled “discrete sciences”.

What are the IDS Centers of Excellence?

In July 2006, DHS selected four universities to establish centers for research on advanced methods in information analysis focused on developing and extending computational technologies for deriving knowledge and understanding from the vast amounts of publicly available data. The IDS Centers of Excellence are led by Rutgers University and are based at the University of Pittsburgh, the University of Illinois at Urbana-Champaign, the Information Sciences Institute of the University of Southern California, and at Rutgers.

IDS Research Mission & Goals

Homeland Security critically depends on our ability to draw inferences from massive amounts of unstructured data contained in books, newspapers, reports, blogs, images, geospatial data, and audio & video streams. Some of the questions that our centers will help to answer include:

- Can we identify a common author writing under different aliases?
- Can we quickly distinguish between a bio-terror attack and a flu outbreak?
- Can we analyze text to ascribe author motives or beliefs?

Our broad mission is to develop fundamental theories, algorithms and tools to enable us to derive knowledge from data, including:

- Novel technologies to identify patterns in massive graphs & datasets that are changing rapidly
- Robust methods to extract, summarize, and track information about events and opinions from unstructured text

- New capabilities for knowledge discovery and information integration across media
- Computationally efficient methods to fuse and represent information in graphs
- Trainable learning algorithms to identify events like infectious disease outbreaks
- Trustworthy methods to preserve privacy

IDS Educational Programs

The IDS centers are committed to: building pioneering educational programs that nurture the homeland security workforce of the future; developing curricular materials for all academic levels; and facilitating participation by traditionally under-represented groups. Our programs include: intensive summer courses and research for undergraduates; programs for K-12 faculty, undergraduate faculty, and government & industry researchers; new courses and degree programs for graduate and undergraduate students; as well as focused workshops, seminars, and tutorials for students and researchers in related disciplines.



IDS University Affiliates: A “Center of Centers”

DyDAN: Center for Dynamic Data Analysis



DyDAN, led by Rutgers University, is the IDS coordinating center. DyDAN is developing novel technologies to find patterns and relationships buried in massive datasets that are rapidly changing through time. DyDAN is:

(1) Investigating new methods to represent, analyze, interrogate & navigate massive rapidly changing graphs that contain multiple types of information about the relationships between the entities associated with the nodes of the graph.

(2) Developing technology to run continuous queries to monitor incoming streams from multiple data sources and exploring new approaches to extract information across multiple data modalities, such as streams of text and video.

Website: www.dydan.rutgers.edu

CERATOPS: Center for Extraction and Summarization of Events and Opinions in Text



CERATOPS, led by the University of Pittsburgh, is developing robust methods to extract and summarize information on events and beliefs from free text. CERATOPS is:

(1) Creating trainable learning algorithms that automatically create domain-specific patterns to identify facts and relations associated with relevant events, such as disease outbreaks.

(2) Developing trainable learning algorithms that can distinguish factual assertions from subjective (non-factual) assertions, identify beliefs that are held by an entity, and assess the intensity, polarity, and motivation and attitude types of those beliefs.

(3) Creating methods for understanding event and belief progressions over time.

Website: www.cs.pitt.edu/mpqa/ceratops

Center for Multimodal Information Access and Synthesis



MIAS, led by the University of Illinois at Urbana-Champaign, is developing new computational algorithms and tools to allow analysts to access a variety of data formats and integrate them with existing resources and to transform raw data into useful and understandable information that enables productive and efficient analysis.

MIAS is extending the state-of-the-art and developing technologies for: (1) collecting focused data from multiple sources; (2) searching and navigating across disparate data modalities; (3) augmenting knowledge bases by inferring semantics from unstructured data and images; (4) relating real-world entities to existing institutional resources; and (5) constructing rich semantic structure and networks of entity linkages.

Website: www.mias.uiuc.edu

Center for Knowledge Integration and Discovery



CKID, led by the USC Information Sciences Institute, is developing new technologies for information integration, access, and knowledge discovery across media. CKID is:

(1) Expanding capabilities for information extraction from natural language text and speech

(2) Developing methods for cross-medium integration of geospatial data

(3) Using statistical and inference-based techniques within data analysis and pattern discovery algorithms to analyze extracted information, social networks, and other data

(4) Creating a centralized repository in a semantics-based knowledge representation system to facilitate general access

Website: www.ckid.org