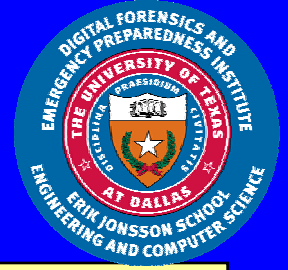


# Secure Sensor Data/Information Management and Mining

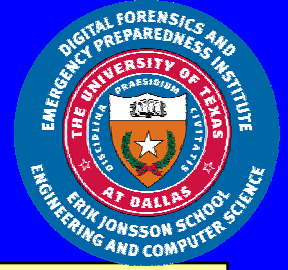
Bhavani Thuraisingham  
The University of Texas at Dallas

October 2005



# Outline

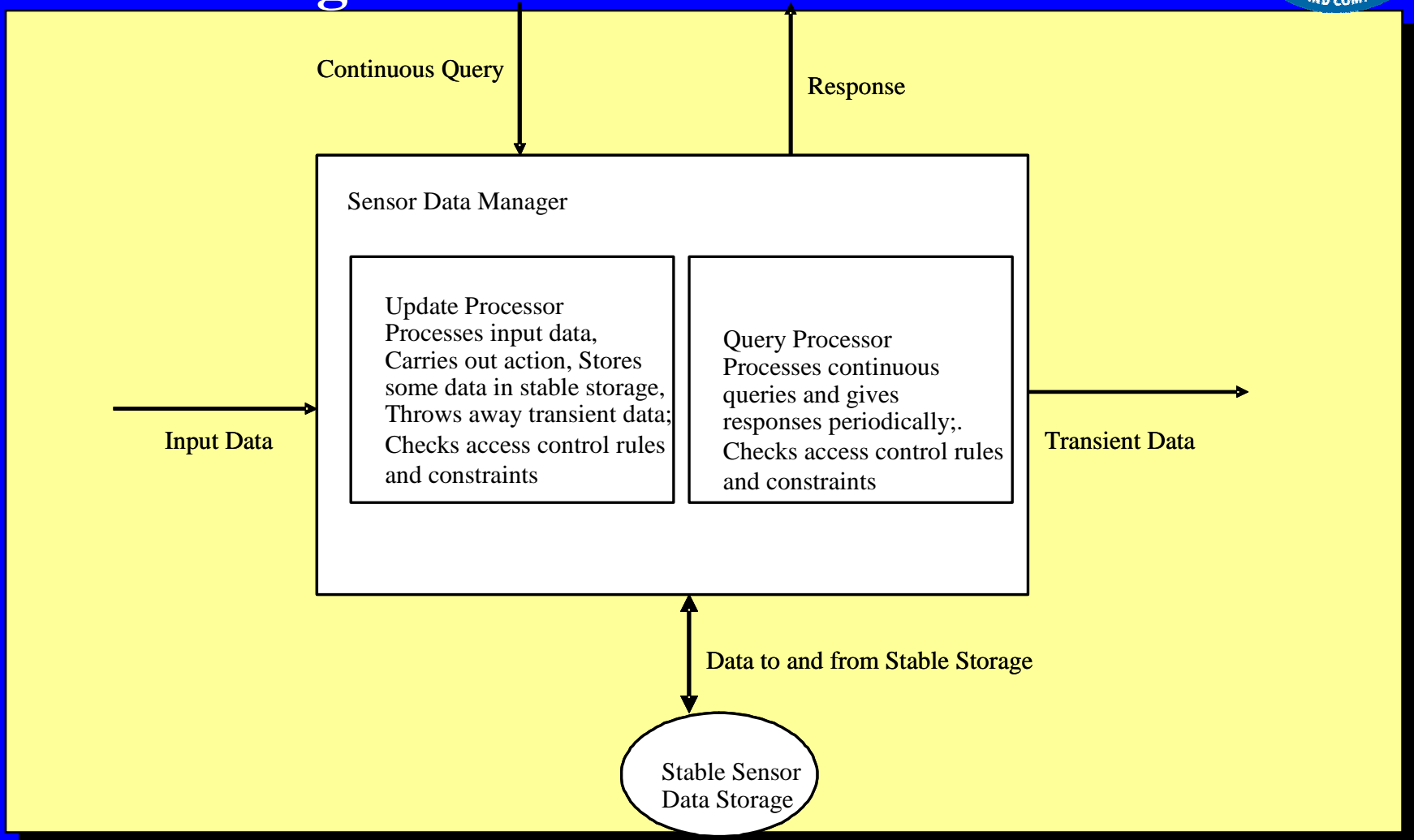
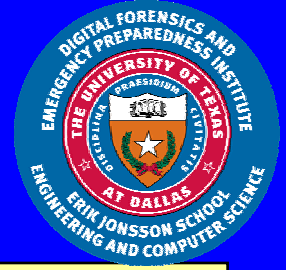
- **Secure Sensor Data/Information Management**
  - Overview, Data Manager, Security Policy, Directions
- **Dependable Sensor Data/Information Management**
  - Overview, Architecture, Directions
- **Dependable Sensor Data Mining**
  - Overview, Real-time data mining, Security and Data Mining, Directions



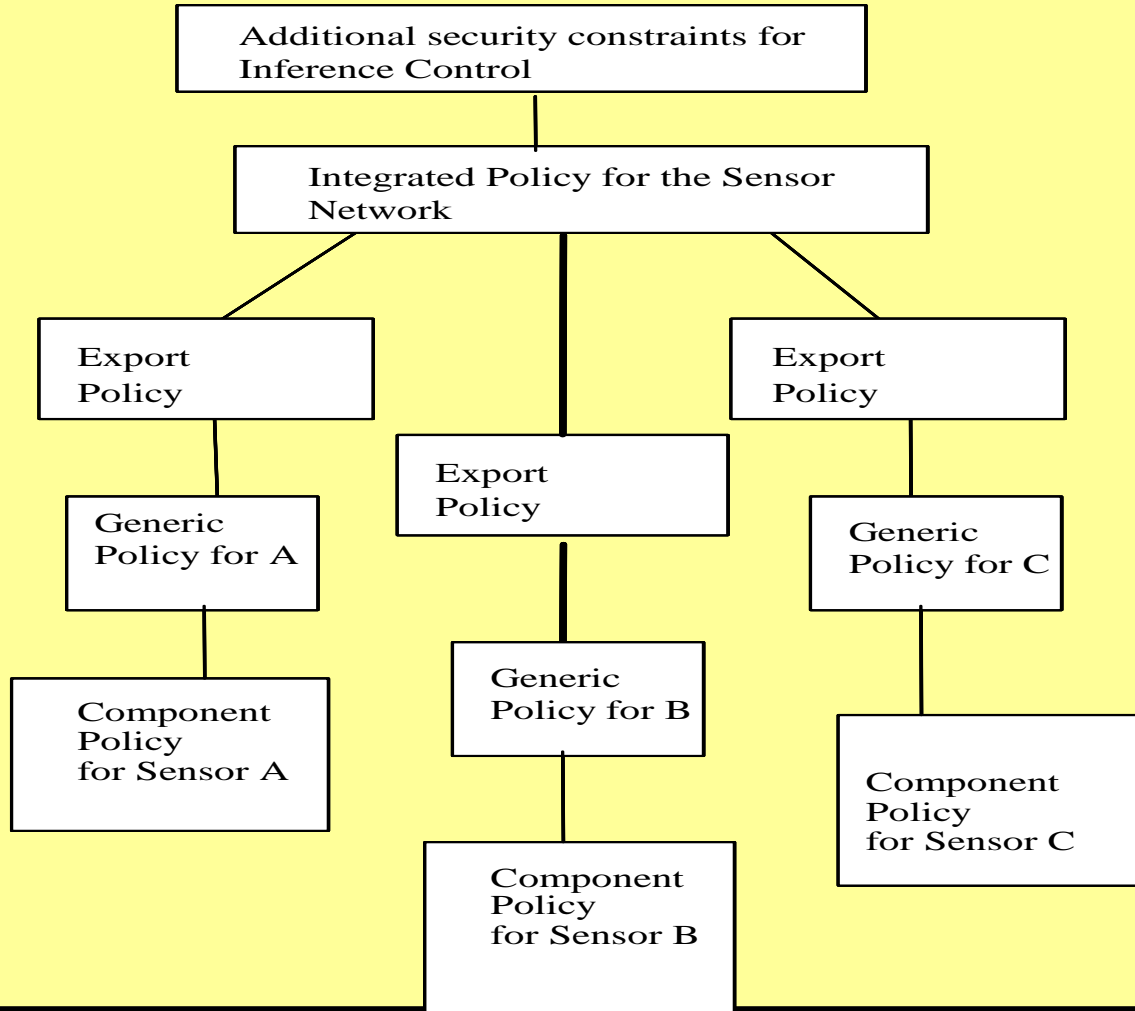
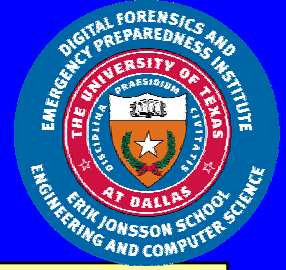
# Secure Sensor Information Management

- **Sensor network consists of a collection of autonomous and interconnected sensors that continuously sense and store information about some local phenomena**
  - **May be employed in battle fields, seismic zones, pavements**
- **Data streams emanate from sensors; for geospatial applications these data streams could contain continuous data of maps, images, etc. Data has to be fused and aggregated**
- **Continuous queries are posed, responses analyzed possibly in real-time, some streams discarded while rest may be stored**
- **Recent developments in sensor information management include sensor database systems, sensor data mining, distributed data management, layered architectures for sensor nets, storage methods, data fusion and aggregation**
- **Secure sensor data/information management has received very little attention; need a research agenda**

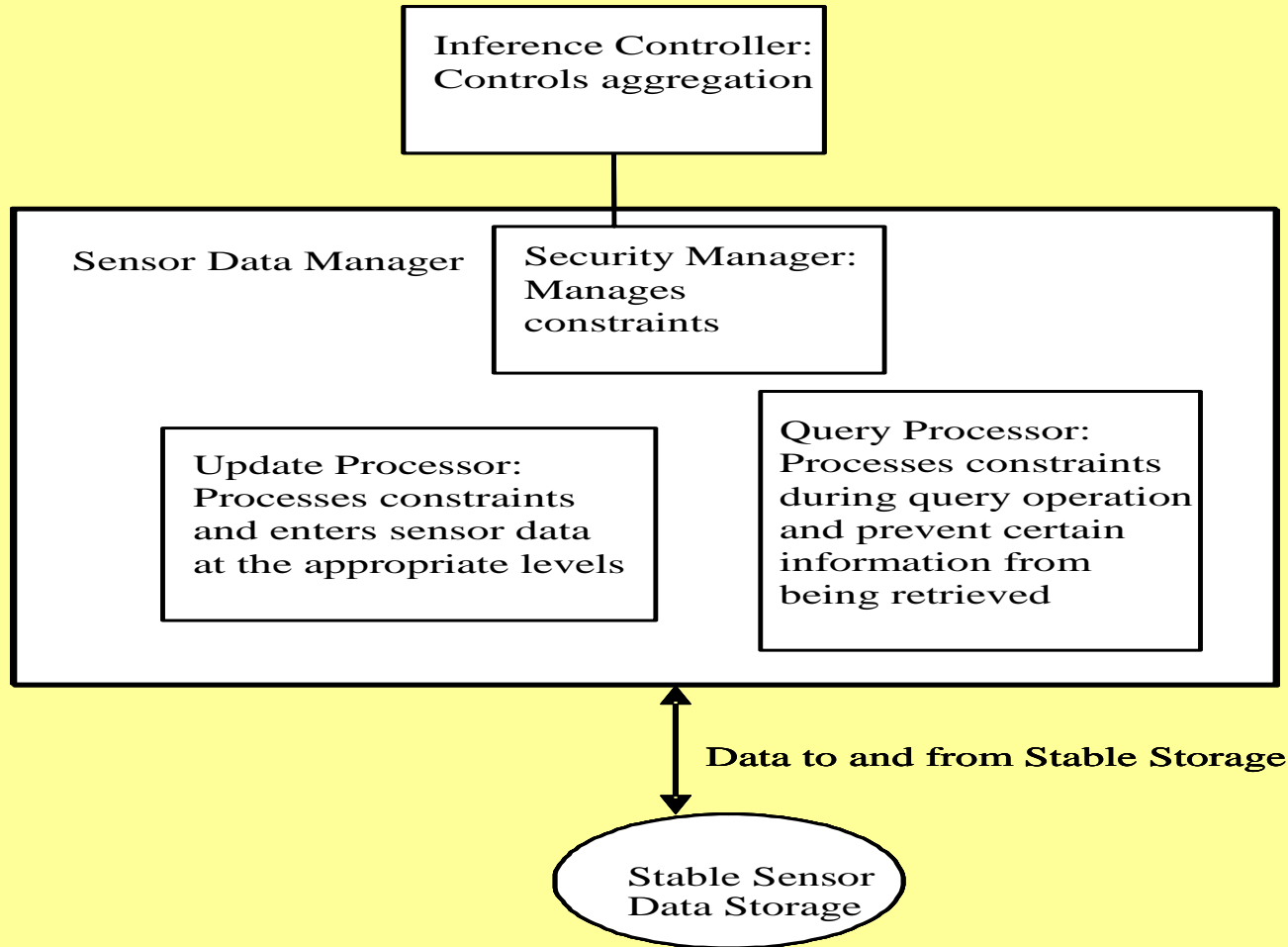
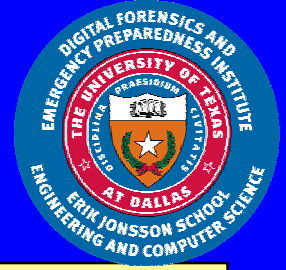
# Secure Sensor Information Management: Data Manager



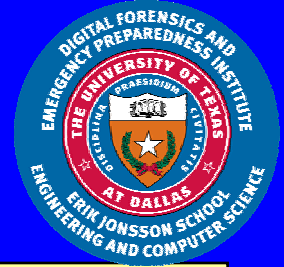
# Secure Sensor Information Management: Security Policy Integration



# Secure Sensor Information Management: Inference/Aggregation Control

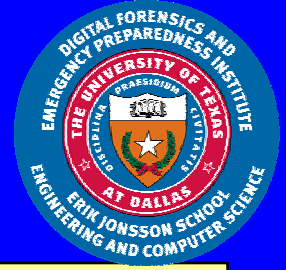


# Secure Sensor Information Management: Directions



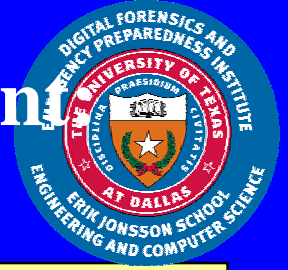
- Individual sensors may be compromised and attacked; need techniques for detecting, managing and recovering from such attacks
- Aggregated sensor data may be sensitive; need secure storage sites for aggregated data; variation of the inference and aggregation problem?
- Security has to be incorporated into sensor database management
  - Policies, models, architectures, queries, etc.
- Evaluate costs for incorporating security especially when the sensor data has to be fused, aggregated and perhaps mined in real-time
- Data may be emanating from sensors and other devices at multiple locations
  - Data may pertain to individuals (e.g. video information, images, surveillance information, etc.); Data may be mined to extract useful information; Need to maintain privacy

# Dependable Sensor Information Management

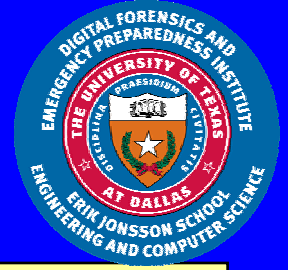


- **Dependable sensor information management includes**
  - **secure sensor information management**
  - **fault tolerant sensor information**
  - **High integrity and high assurance computing**
  - **Real-time computing**
- **Conflicts between different features**
  - **Security, Integrity, Fault Tolerance, Real-time Processing**
  - **E.g., A process may miss real-time deadlines when access control checks are made**
  - **Trade-offs between real-time processing and security**
  - **Need flexible security policies; real-time processing may be critical during a mission while security may be critical during non-operational times**

# Secure Dependable Information Management Directions



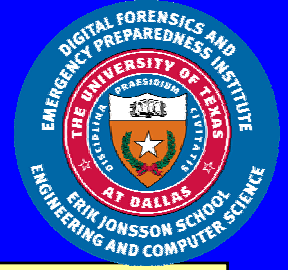
- **Challenge: How does a system ensure integrity, security, fault tolerant processing, and still meet timing constraints?**
- **Develop flexible security policies; when is it more important to ensure real-time processing and ensure security?**
- **Secure dependable models and architectures for the policies; Examine real-time algorithms – e.g., query and transaction processing**
- **Research for databases as well as for applications; what assumptions do we need to make about operating systems, networks and middleware?**
- **Developing dependable sensor objects**
- **How do we integrate sensor applications with dependable semantic webs?**



# Dependable Data Mining

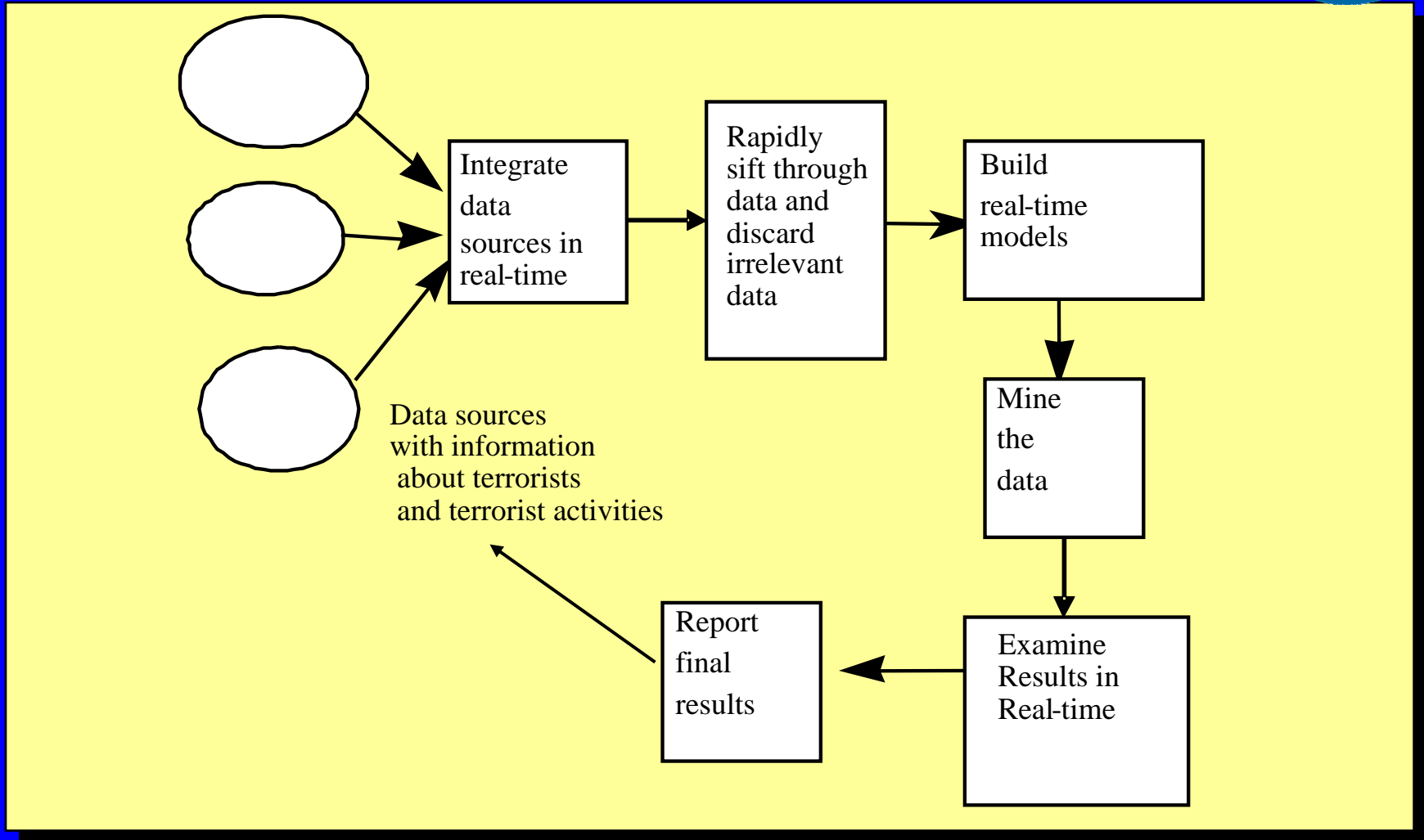
- **Data Mining techniques need to meet timing constraints for many applications including sensor information management**
  - **Counter-terrorism, Financial quotes, Military**
- **Data mining techniques need to maintain security and privacy**
  - **E.g., Prevent inference problem due to data mining**
- **Need high quality data to extract meaningful results from data mining**

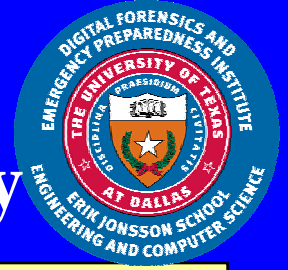
# Need for Real-time Sensor Data Mining: Counterterrorism



- **Nature of data**
  - **Data arriving from sensors and other devices**
    - **Continuous data streams**
  - **Breaking news, video releases, satellite images**
  - **Some critical data may also reside in caches**
- **Rapidly sift through the data and discard unwanted data for later use and analysis (non-real-time data mining)**
- **Data mining techniques need to meet timing constraints**
- **Quality of service (QoS) tradeoffs among timeliness, precision and accuracy**
- **Presentation of results, visualization, real-time alerts and triggers**

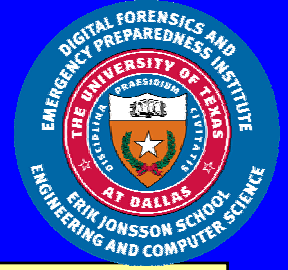
# Sensor Data Mining for Real-time Threats





# Data Mining as a Threat to Security/Privacy

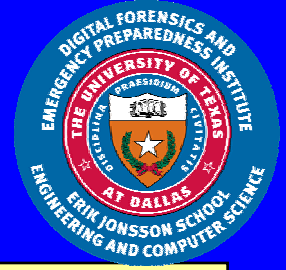
- Data mining gives us “facts” that are not obvious to human analysts of the data
- Can general trends across individuals be determined without revealing information about individuals?
- Possible threats:
  - Combine collections of data and infer information that is private
    - Disease information from prescription data
    - Military Action from Pizza delivery to pentagon
- Need to protect the associations and correlations between the data that are sensitive or private



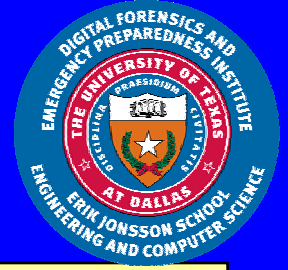
# Solutions and Directions

- **Randomization**
  - Introduce random values into the data and/or results
  - Challenge is to introduce random values without significantly affecting the data mining results
  - Give range of values for results instead of exact values
- **Secure Multi-party Computation**
  - Each party knows its own inputs; encryption techniques used to compute final results
- **Directions**
  - Build models in real-time
  - Associate timing constraints with data mining techniques?
  - Mining sensor and stream data
  - Integrate security and real-time processing for data mining

# Some Relevant Research at the University of Texas at Dallas



- **Privacy Preserving Surveillance**
  - Sensor Data Management and Data Mining
  - Motion capture animation databases
- **Dependable semantic web**
  - Ensure TPC: Trust, Privacy and Confidentiality
- **Directions:**
  - Build sensor information management applications that utilize dependable semantic webs



# Ideas and Directions?

**Prof. Bhavani Thuraisingham**

- Director Cyber Security Center
- Department of Computer Science
- Erik Jonsson School of Engineering and Computer Science
- The University of Texas at Dallas
- Richardson, Texas
- [bhavani.thuraisingham@utdallas.edu](mailto:bhavani.thuraisingham@utdallas.edu)

<http://www.utdallas.edu/~bxt043000/>

**President**

**Dr-Bhavani Security Consulting**

**Dallas, TX**

[www.dr-bhavani.org](http://www.dr-bhavani.org)