



Panel Discussion S&T Direction (Big Issues)

Dr. Carey Schwartz
Panel Chair
25 June 2014

DISTRIBUTION STATEMENT A. Approved for Public Release



S&T Challenges for Data Science

Dr. Carey Schwarz

Overarching Problems:

- Find the information bearing component of a high dimensional signal or set of signals
- Find “regularities” in the data that enable warfighters to successfully complete their mission in a dynamic world within the timeline of the mission

Where we need scientific and technical advances

1. Representations
2. Data Alignment
3. Treating Processing and Analysis as a SYSTEM
4. Numerical methods including scaling and stability
5. Methods that address uncertain, incomplete, imprecise, and contradictory data
6. Analytics including imagery (all modalities), text, and graphs
7. How are good decisions made; what is the impact of an imperfect picture on decision making?



Data Science S&T Challenges

Mr. Jerry Desrosiers

Overarching Issue

- Fielding tactical warfighting capabilities to enable effective decision making across all Naval missions in a highly dynamic A2/AD environment

Needed scientific / technological advances

- Scalable data ecosystem for efficient & effective data ingest, indexing, pedigree, provenance, tagging, alerting . . .
- Effective, efficient mapping mechanisms to align semantics across mission areas/systems without requiring a common semantic model
 - Can't easily change the myriad of legacy data models
 - Must support cross warfare area analytics
- Abstraction mechanisms to hide ecosystem implementation details to enable rapid analytics development & changes to underlying data infrastructure
- Cloud to Cloud federation capabilities in challenging D-DIL environment
 - Federated query
 - Mission based subscription



Communications & Networks: Operational Constraints & Desired Capabilities

Dr. Santanu Das

- **Crowded spectrum**
 - Drives the need for increased spectral efficiency and diversity.
- **Limited platform real estate and power-constrained tactical platforms**
 - Drives the need for integrated, multi-functional, wideband apertures.
 - Drives the need for innovative apertures with reduced size, weight, and power.
- **Electronic attack and EMI**
 - Drives the need for anti-jamming capabilities and co-site mitigation techniques.
- **Electronic intercept**
 - Drives the need for techniques and approaches for low probability of intercept and low probability of detection.
- **Increased sensor resolution and shooter demand**
 - Drives the need for higher data rates.
- **Dispersed forces and intermittent connectivity**
 - Drives the need for increased range, loss tolerance, and recovery.
 - Robustness vs. efficiency, Distributed vs. centralized algorithms and protocols for Mobile Ad-Hoc networking protocols
- **Dynamic priorities and missions**
 - Drives the need for quality of service and autonomous network management/control.
 - Interoperability and security



S&T Challenges for Data Science

Gary Toth

Overarching Problems:

- Understanding Naval activities from incongruous large dispersed heterogeneous big data stores
 - Beyond Cassandra (Facebook, Twitter, etc) and Hbase
 - Dealing with the data hoarding paradox: If you create the storage space, humans will fill it
- Adaptive machine learning techniques to replicate human brain-like automated reasoning in lightning fast processing speeds.
 - Real-time big data analytics
 - Efficient and affordable high performance computing on sensors
- Real-time Big Data Predictive Analytics to support diverse users
 - Development of analytic layers that sit above the data layer for deploying real-time scoring and dynamic analytics
 - Development environment to enable data distillation, model development, validation, real time scoring and model refresh

Machine Intelligence: Automated Image Understanding

Goal: Develop principles, computational methods, and tools that enable vision systems to fully understand images/video in all modalities, in all environments, for decision-making, for performing all vision-based tasks.

Applications

- Automated surveillance, reconnaissance, search, monitoring, ...
- Targeting and damage assessment
- Decision-aids for image analysts
- Forensic investigations
- Perception for autonomous agents/robots
- ...

Interested in all

- **Modalities:** EO/IR, SAR, ISAR, Sonar, ...
- **Elevations:** Satellite to ground level
- **Illuminations:** Outdoor-Indoor, Night-Day, Passive-Active
- **Environments:** Coastal areas, harbors, urban, rural, ...

Information Dominance

- Overwhelming amount of imagery
 - Less than 1% of Argus WAMI data are looked at by analysts
- Example of analysts' products for decision-makers
 - **Tactical:** Brief messages: when, where, what, ...
 - **Strategic:** Well annotated with detailed descriptions

Image understanding is hard

- **Size of the World**
- **Clutter in the image:** 100s of objects, many activities and signal & noise are similar on local scales
- **Appearance:** Lighting, pose, scale, occlusion
- **Image data is underspecified**



We use knowledge & reasoning (intelligence)!

Problems in image understanding

What we really want to do is **Scene** understanding

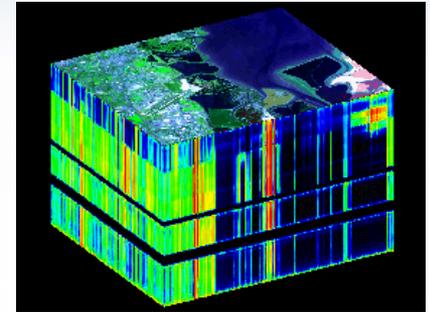
- Compute structure of the scene, recognize the scene type, infer functionality of spaces
- **Representation** of objects, activities, events, and methods for **learning** them (algorithms, datasets)
- Recognize all **specific** and **classes of** objects, persons, activities, events
- Models for integration of knowledge & image data (**Knowledge Bases, Reasoning**)
- **Infer intentions and threats, predict events**
- Describe the scene in natural language

Distribution A: approved for public release

Challenge: identify conditions for theoretical guarantees for the algorithm performance.

- ***Small Data Problem:*** high dimensional data, small number of observations.

- What are the right constraints?
- What observations carry most information?

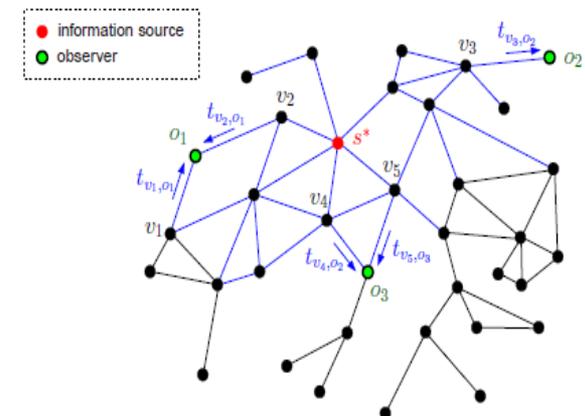


hyperspectral data cube
dimensionality > 1 million

- ***Identifying causal determinants***

- of diffusion of information in response to specific events
- of the patterns of interaction among groups

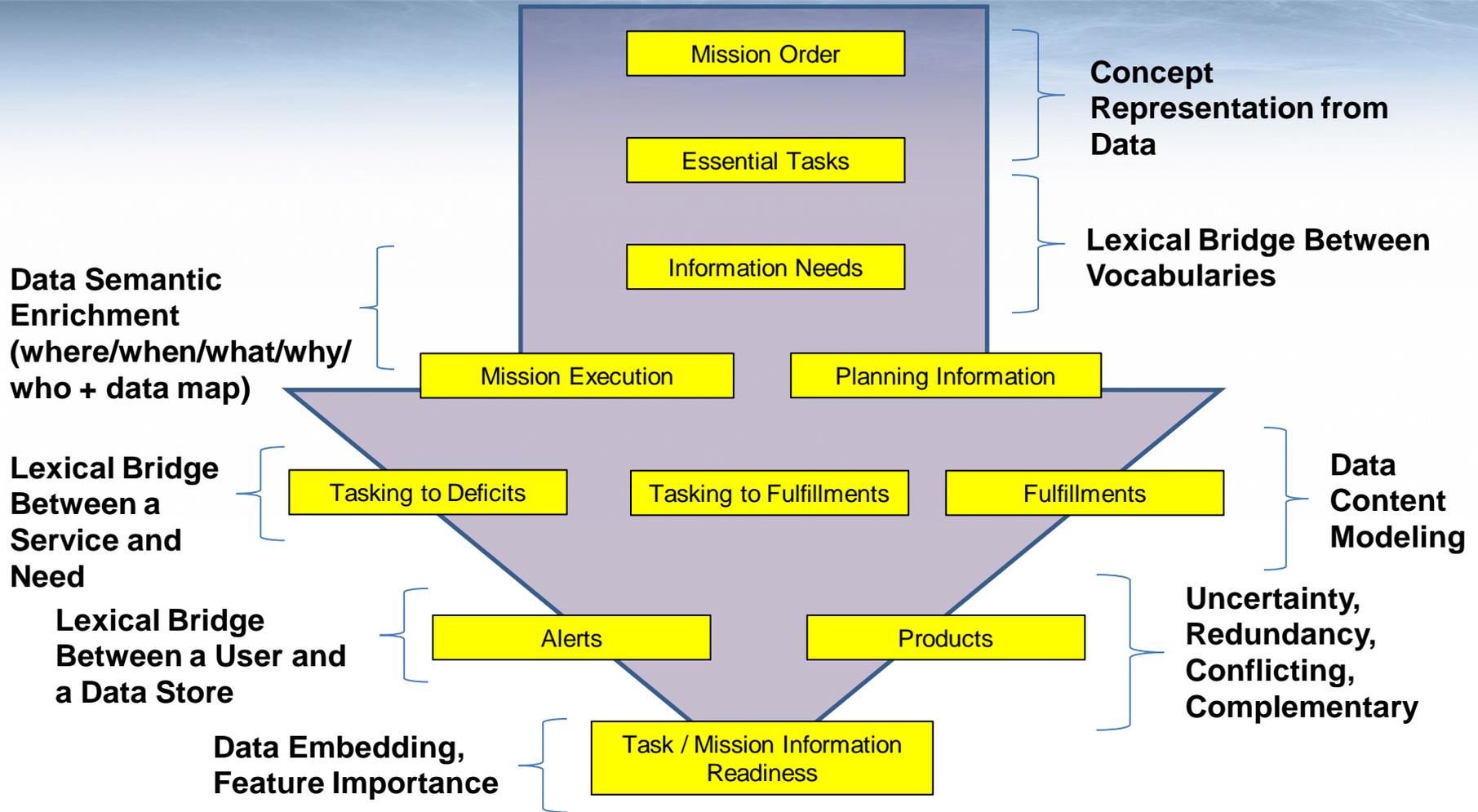
- Who influences whom? How to identify the main sources of influence, individuals or groups?
- Modelling real-world networks. New computational algorithms that scale to large networks with performance guarantees.





Data Science - Strategic Goal

Martin Kruger



Enable Automated Information Support Across All Missions Areas and Warfighting Functions