Data and Decision Analytics

Data Analytics Assessment Overview

Problem: There is currently no standard way to implement and assess performance for data analytics

- Heterogeneous data sources/algorithms without ground truth
- Hard to know what capability is being purchased with few means to assess performance of service
- Dynamic mission space with changing requirements

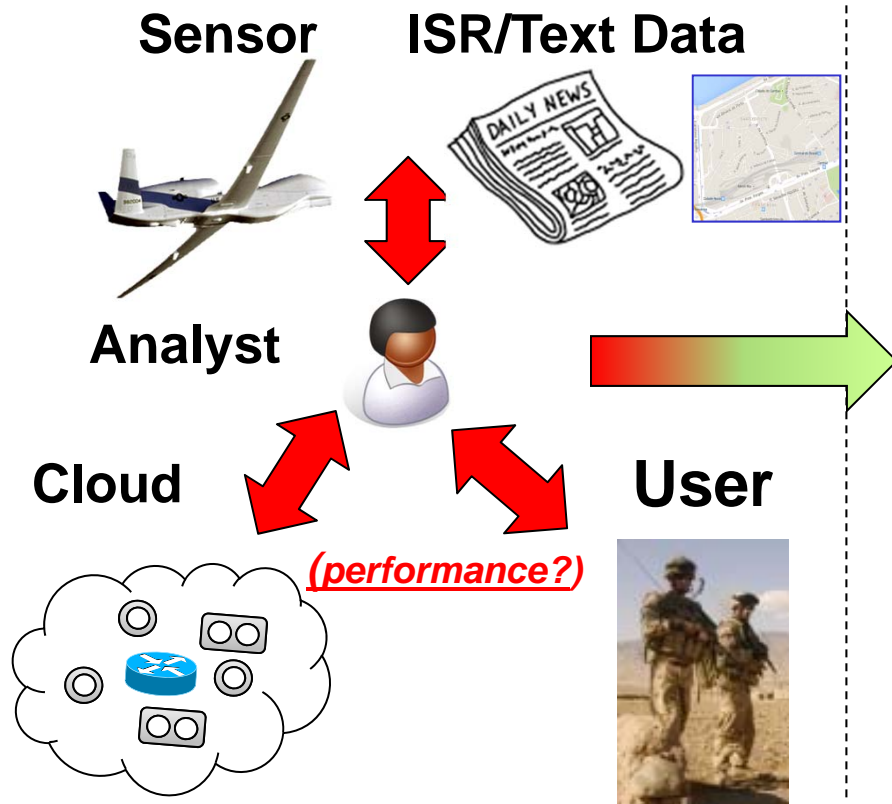**Solution: Data analytics framework**

- Standard data models with ground truth
- Development framework to standardize risk analytics on information sources, algorithms, and processing
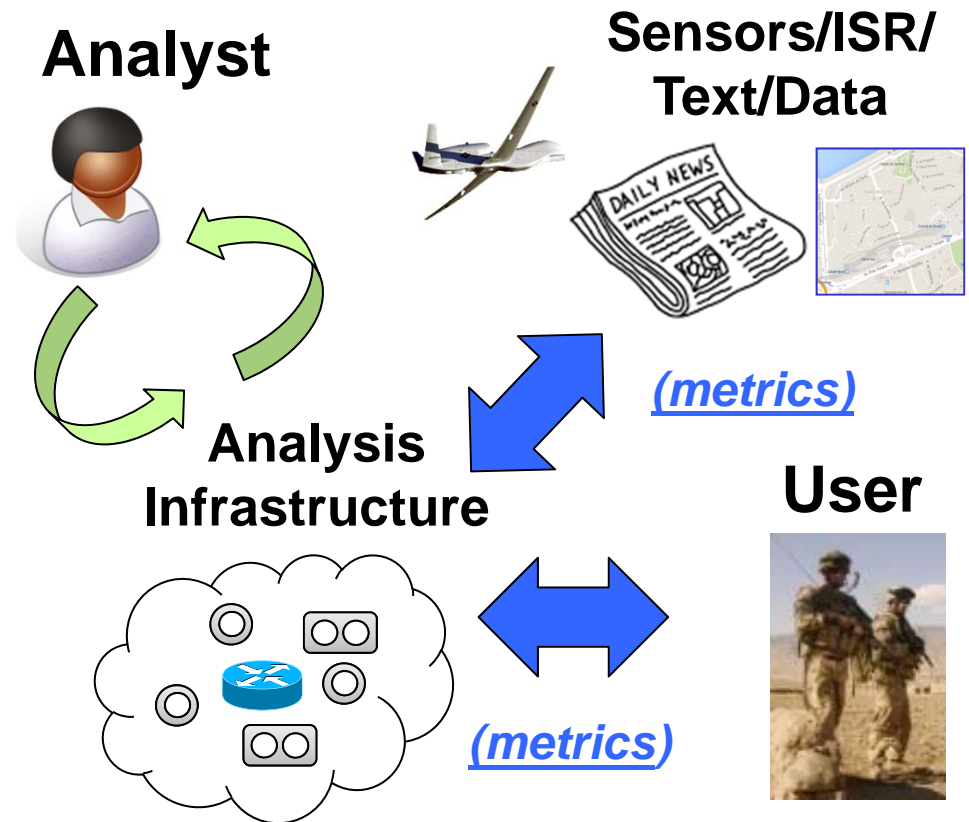- Adaptable framework that can change as mission requirements change

D2D/Data Analytics Approach

Analyst oversees delivery of information products to customer with rigorous quality of service guarantees
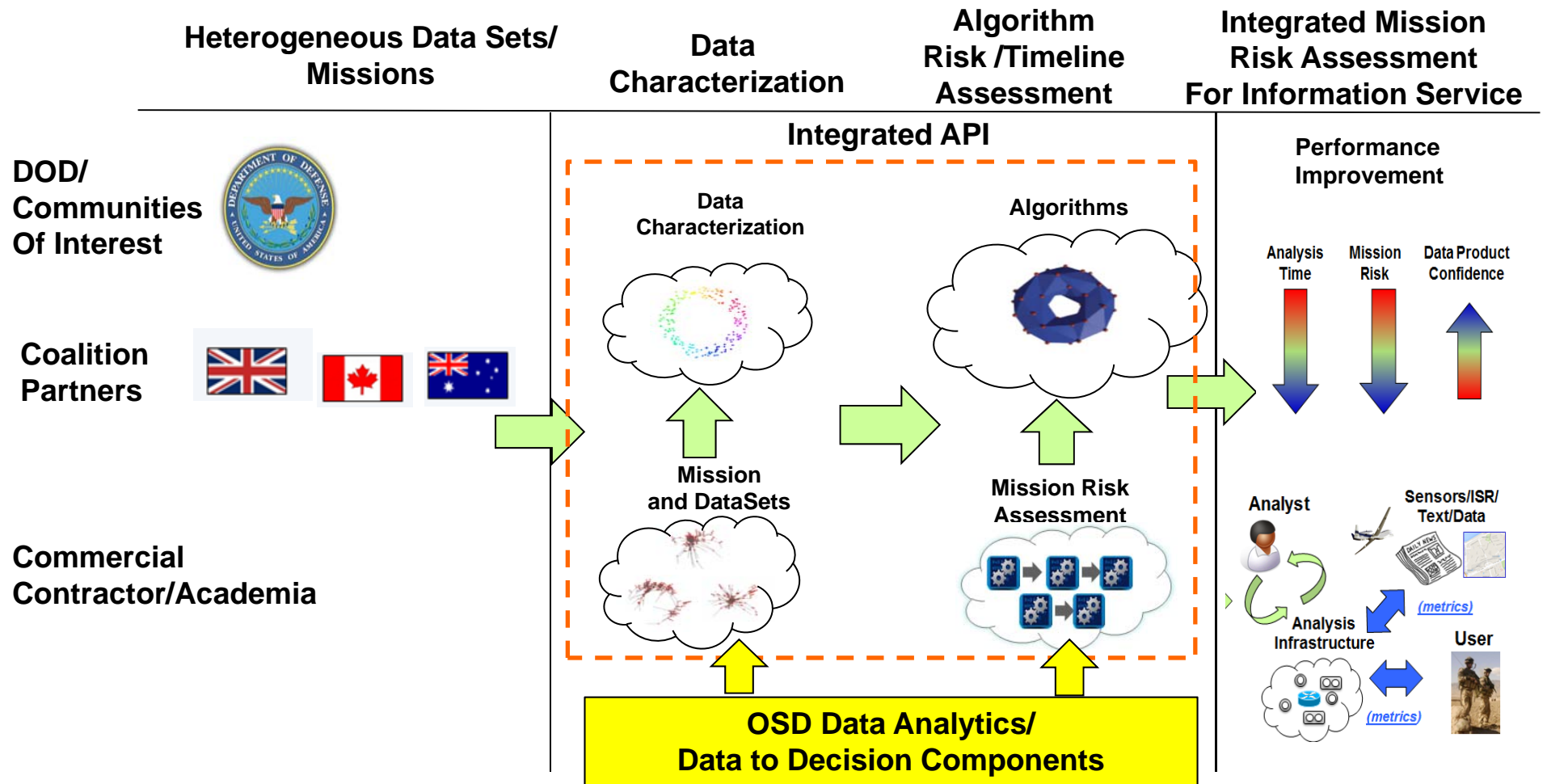
Current Approach

Sensor ISR/Text Data

Analyst

Cloud User

(performance?)

New Approach

Analyst Sensors/ISR/ Text/Data

Analysis Infrastructure User

(metrics)

(metrics)

Data Analytics Performance Assessment

Implementation and assessment of information service can be standardized to assess overall mission performance

Components Can Assess Multiple Mission Types

Incorporate a cloud based open standard for information services development and assessment so basic components can be used assess multiple types of missions
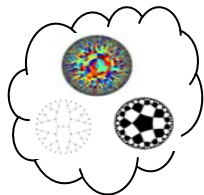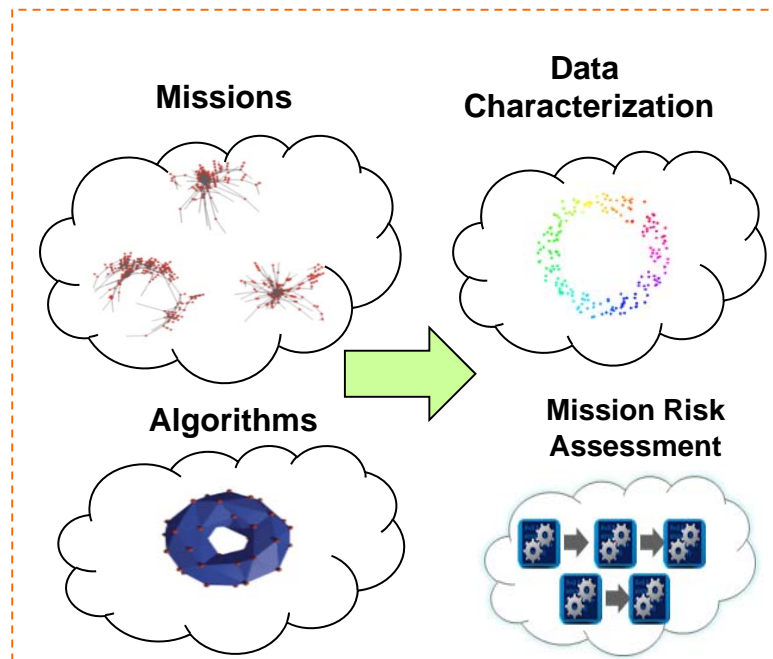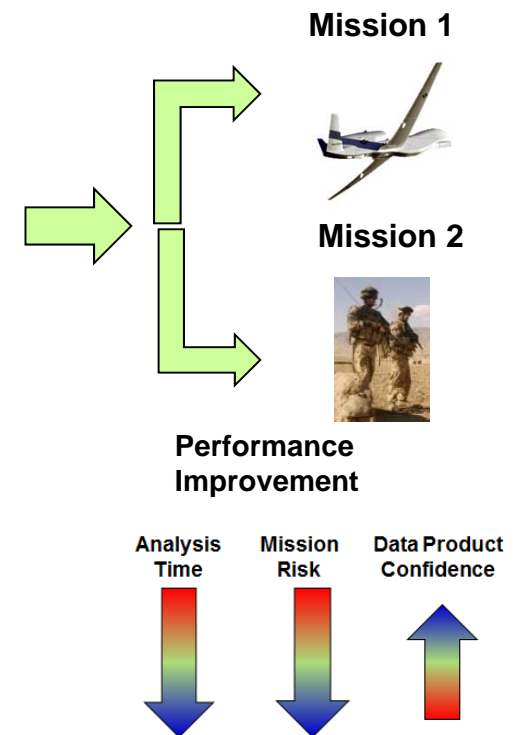
Data Analytics/ Data to Decision Elements

Transitionable Components

Same Components Provide Assessment For Multiple Customer Missions

Text Analytics

Imagery Analytics

New Algorithm Research

Missions

Algorithms

Data Characterization

Mission Risk Assessment

Mission 1

Mission 2

Performance Improvement
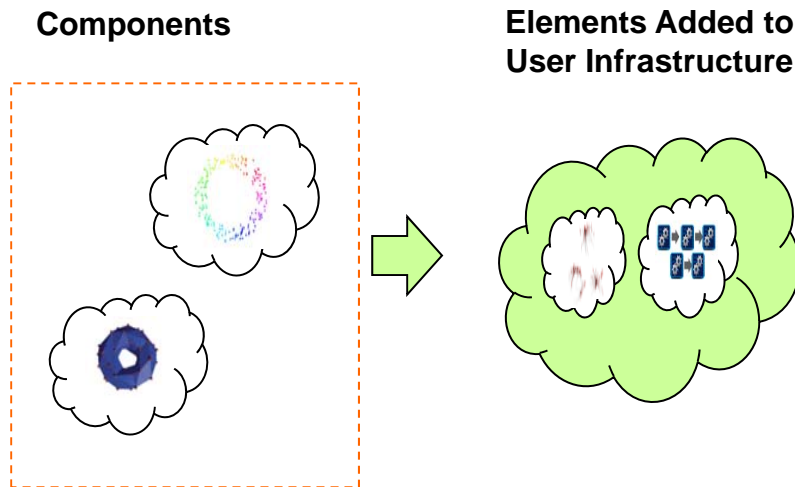
Analysis Time

Mission Risk

Data Product Confidence

Transition Models

Models can either be added to existing infrastructure or used by existing infrastructure as diagnostics for performance
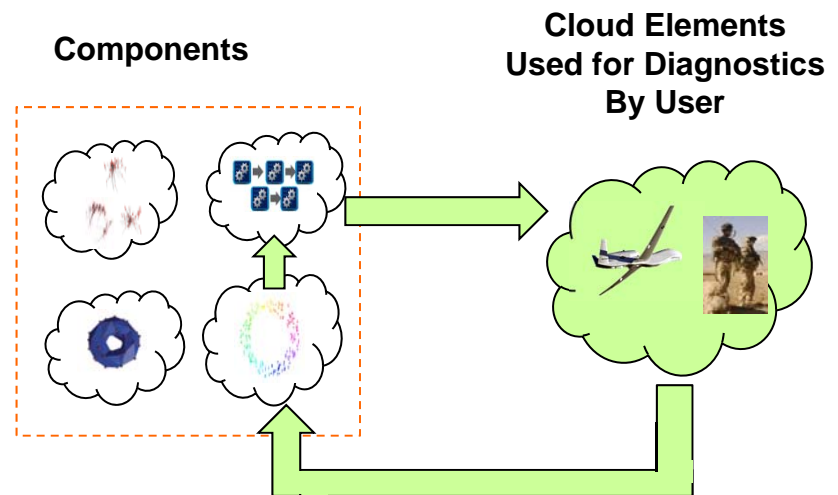
Model 1
(direct integration of components)

Model 2
(user integrates remote elements for their analysis)

Components

Elements Added to User Infrastructure

Components

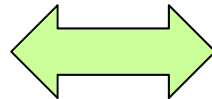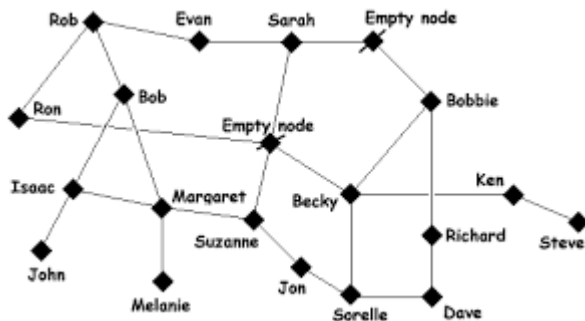**Cloud Elements Used for Diagnostics By User**

Mission and Data Set Components

Standard threat or mission graphs and the associated data needed to assess a particular threat are available for baseline assessment and design of future missions analysis

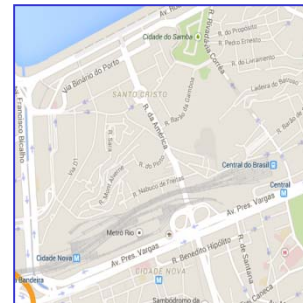## Standard Mission Graphs

**Scenario Graph Specifies
What Data Should Be Collected**

Standard Data Sets

Standard Data Sets Specify Ground Truth for Different Data Types & Provenance of Relevant Data

Imagery Truth Data

Text Analytic Data

Algorithm and Mission Risk Component

The algorithm and mission risk components can calculate
- **Provenance and risk of data + algorithm conclusion**
- **Timeline for output at given data risk level**
- **Overall mission risk and certainty of conclusion**

Algorithms Data Base

Algorithms data base specifies
risk incurred for different data types
and fidelities and processing time required
for actionable information
over a given architecture.

Assessment of text algorithm
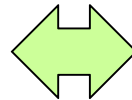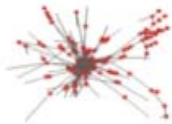
Assessment of track algorithm
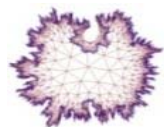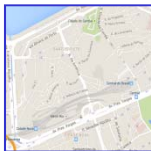
Mission Risk Analysis

Database of algorithm conclusions against
different scenarios with specified
truth data.

Overall risk to mission with truth
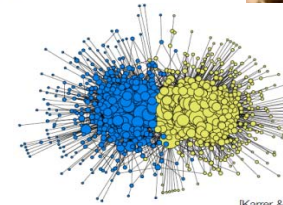
[Karrer & Newman, 2010]

Measure/Model/Manage

Integrated modeling, validation, verification, and management can characterize mission performance with advanced data models

Measure

Model/Validate

Large Data Analytics

Risk Analysis & Modeling

Manage/Verify

Automated Infrastructure

Local Network Analysis

Truth/Verification
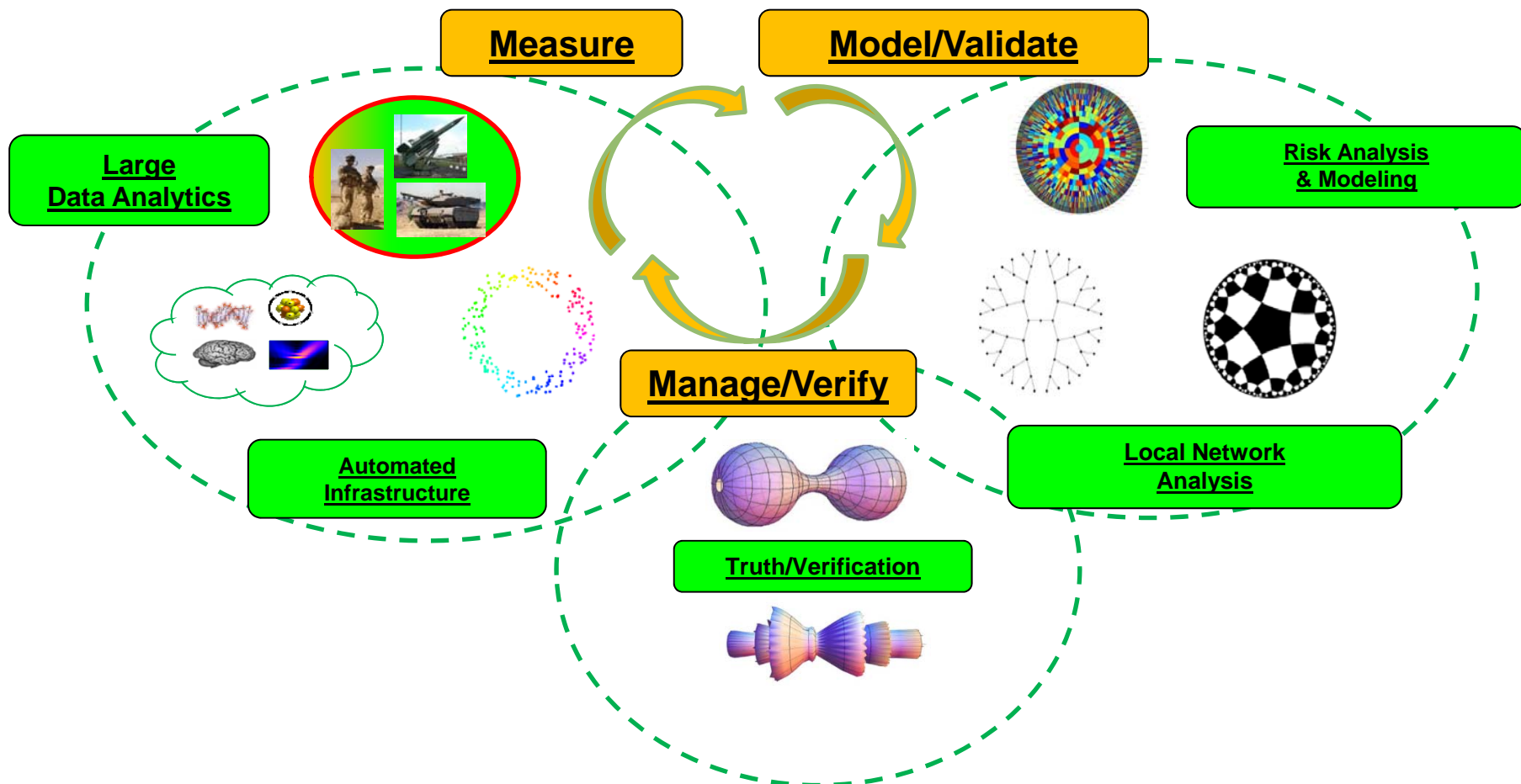
Measurement

We wish to understand how to measure the state of a mission on an infrastructure

What to measure?
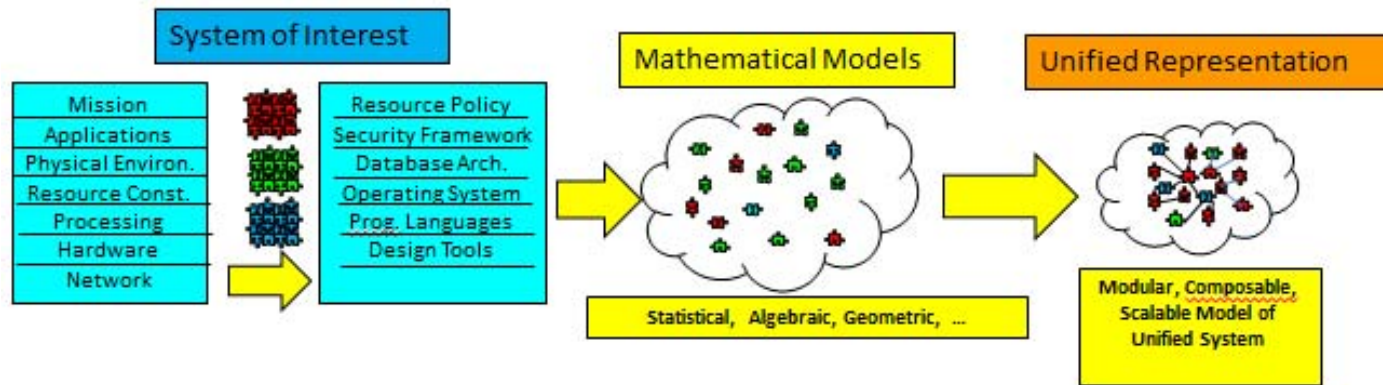
How to measure?
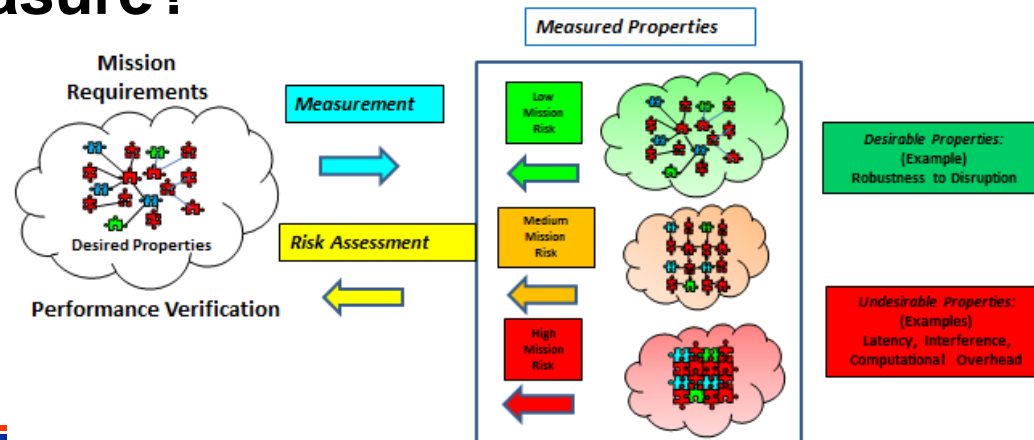
Modeling

We must have validated models of mission performance which can come from known models or empirical data

Mission Operation Trade-space
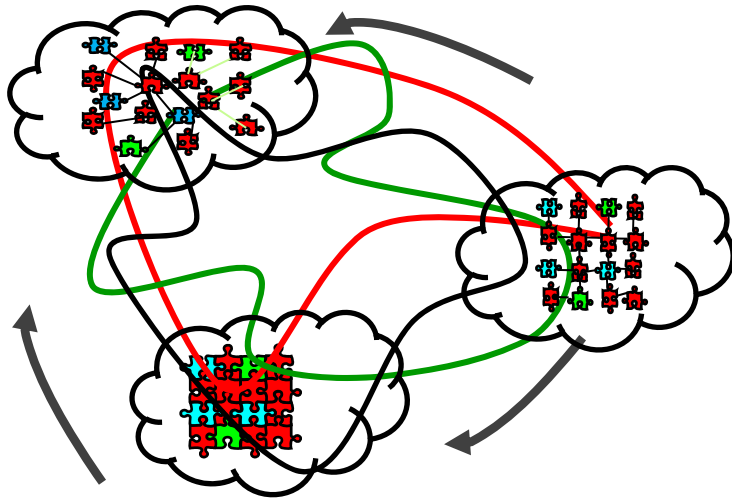
Un-validated Modalities
(high mission risk)
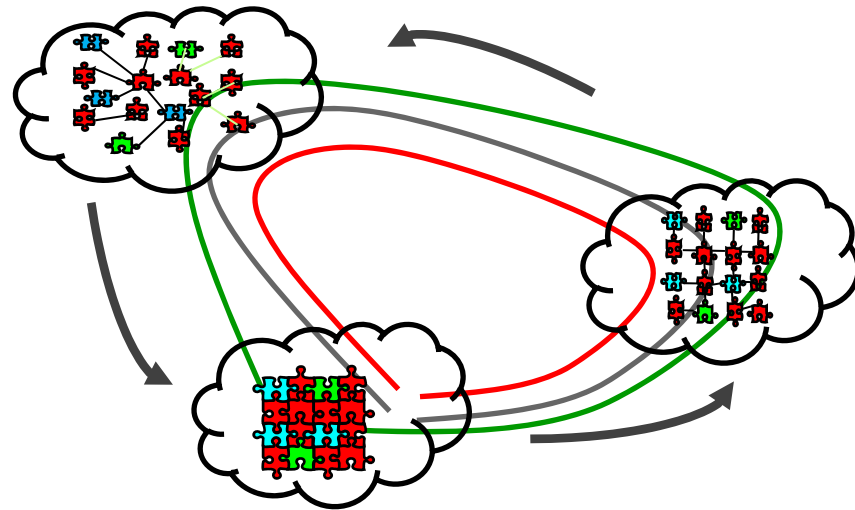
Validated Modalities
(low mission risk)

Management

How do we close the loop at multiple architectural layers to assure mission performance and verify system policy/protocol is working?

Information System

Software States   Hardware States

Network States

System Measurement

System Stability/ Security Analysis

System Taxonomy/ Model

System Policy/Protocol Update

Model Check/ Verification
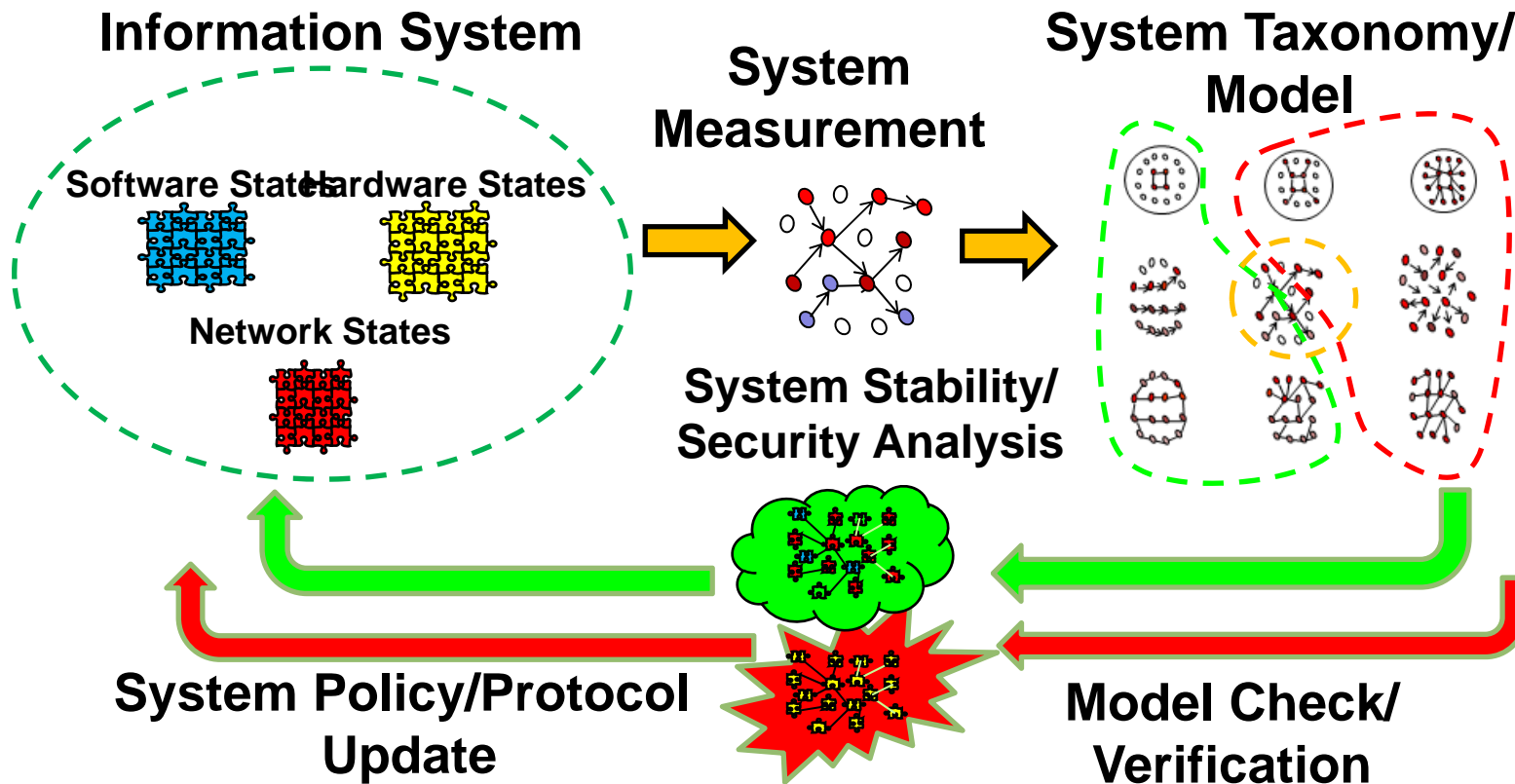
Metrics of Performance

Metrics of performance allow timelines, tracking, and mission performance to be rigorously assessed analyst/commander in real time.

Example Metrics   ## Desired Outcome

Timeline Reduction

Rigorous Mission Threat/Risk Assessment

Rigorous Data Product Confidence Analysis
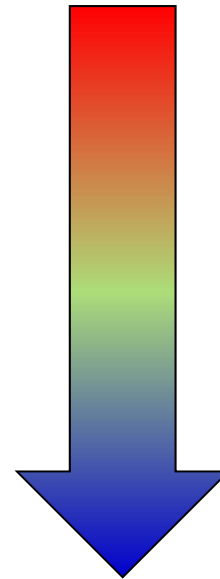
Analysis Time   **Mission Risk**   **Data Product Confidence**
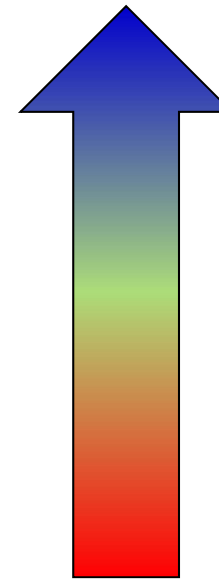
Risk Analysis and Modeling

Unified methods for data modeling require a rigorous risk assessment in order to assure commanders, analysts, and system operators of performance.
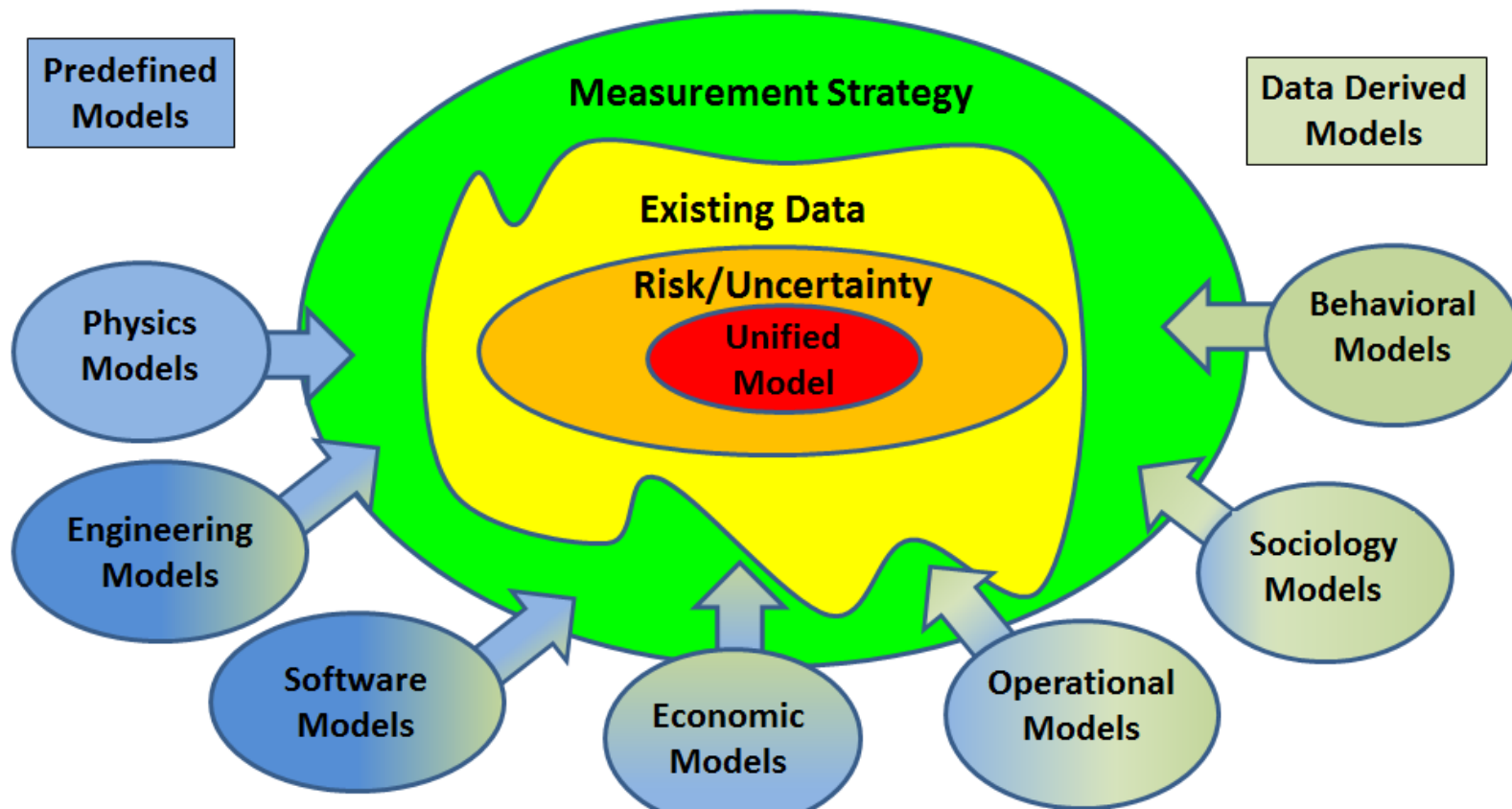
Risk and Autonomy

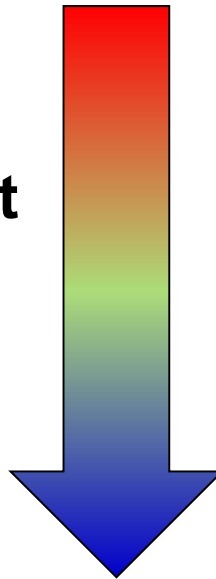For automated system performance to be trusted and effective, a strategy for autonomy that enables the lowest mission risk in balancing human workload with automation should be followed

Partially Autonomous Infrastructure

Real Time Data Analytics

Human Modeling & Sensing

Autonomous Infrastructure

Natural Language
User Applications
Physical Environ.
Resource Const.
Cultural Influence
Social Media
Social Network

Machine Lower Mission Risk

Online Mission Risk Analysis

Human Lower Mission Risk
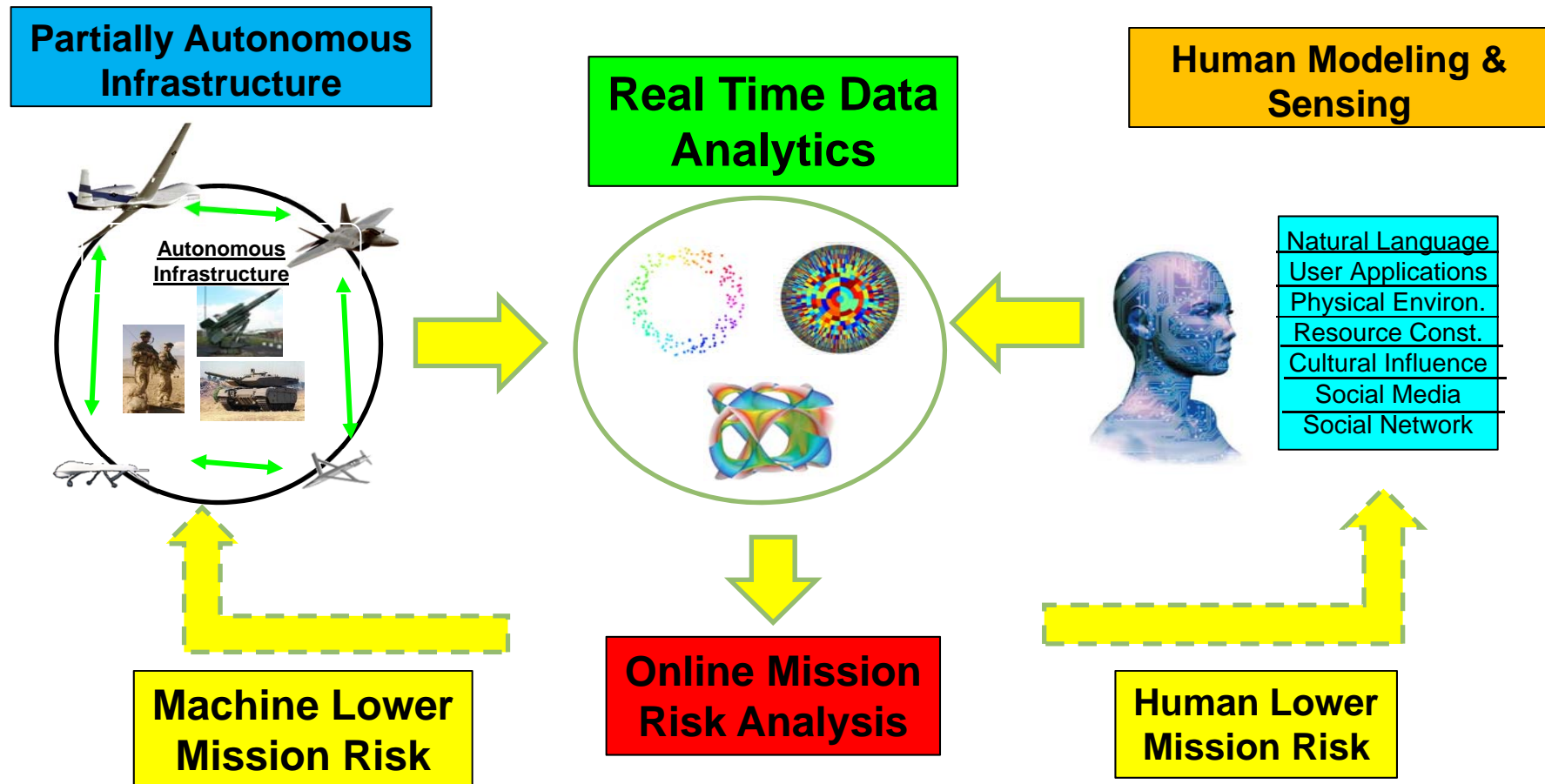
Mission Stack

Measurement, modeling, and management of mission stack must have rigorous performance and risk metrics associated with them

Risk Assessment

Mission Stack

User/Leader high level objectives

Software applications

Hardware infrastructure

Network infrastructure

Physical resources (platforms/spectrum/etc.)

Measure/Model/Manage

Mission Layer

Application Layer

Compute Layer

Network Layer

Physical Layer
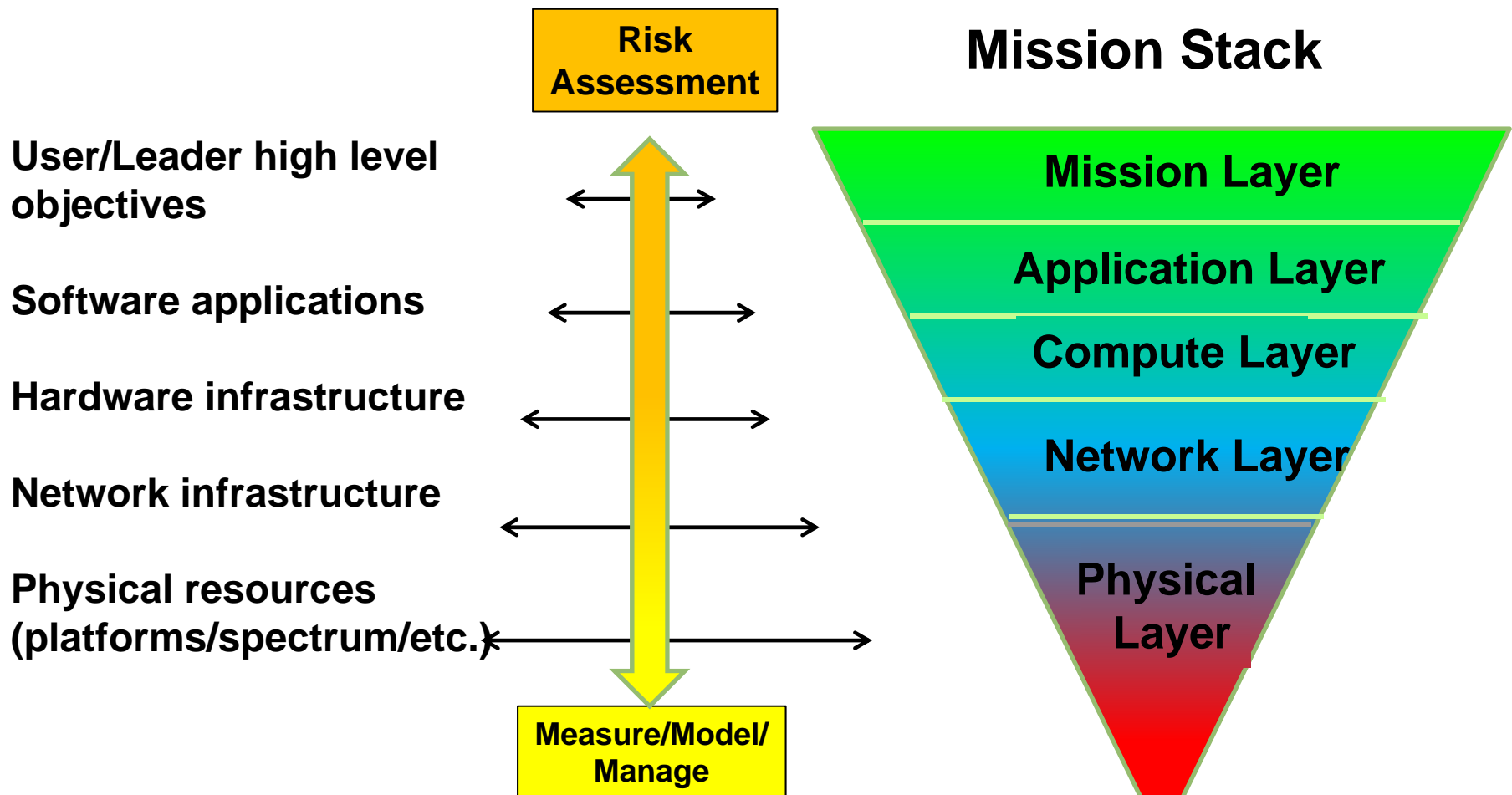
Application Layer

The mission layer may be made up of multiple applications such as sensing, communication, tracking, situational awareness, command and control, etc.

-These methods must be integrated with one unified representation for validation and verification.

Application of Interest

| Sensing |
| Communications |
| Tracking |
| Situational Aw. |
| Text Analytics. |
| Cyber Methods |

| Resource Policy |
| Security |
| Framework |
| Database Arch. |
| Operating System |
| Prog. Languages |
| Design Tools |

Heterogeneous Models

Mathematical Model

Unified Representation

Modular, Composable, Scalable Model of Unified System
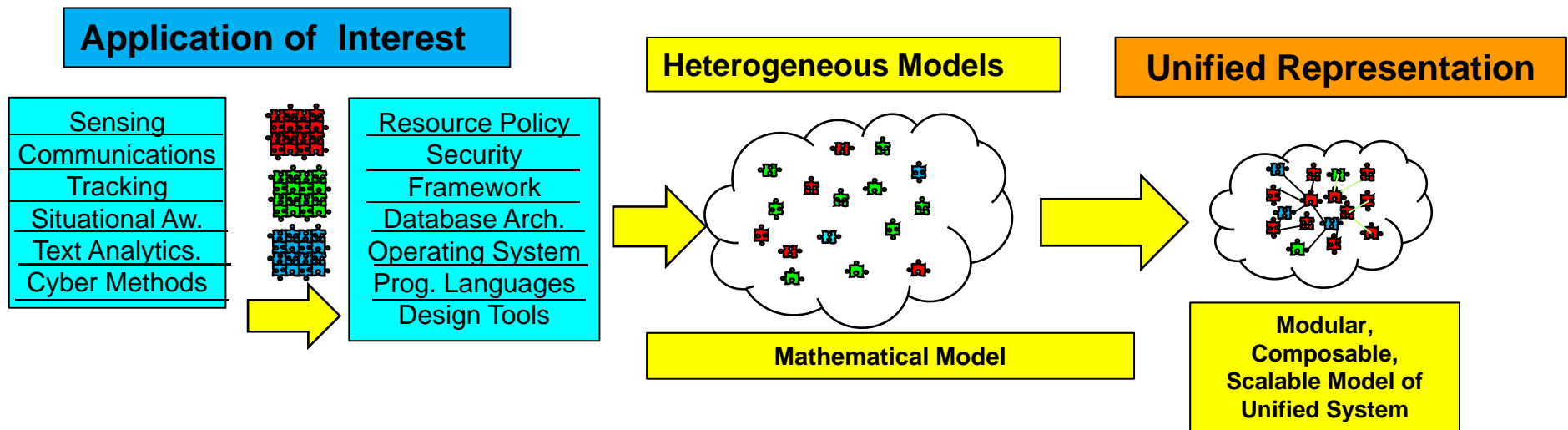
Compute Layer

Current computational infrastructures (cloud resources) are currently high distributed and resource allocation is static. Making this process more dynamic will resilient system performance.

Critical DOD Apps on MAP-Reduce Cloud   Computing Engine

Measurement Based Graph Analytics
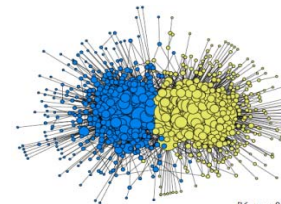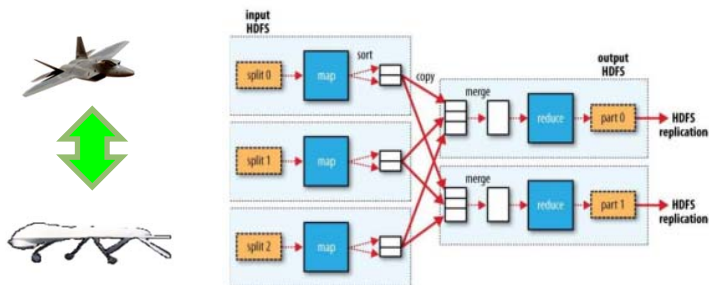
[Karrer & Newman, 2010]

System Performance Verification

Computed System State Representation

Compute   Compute   Compute   Compute
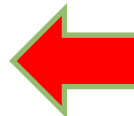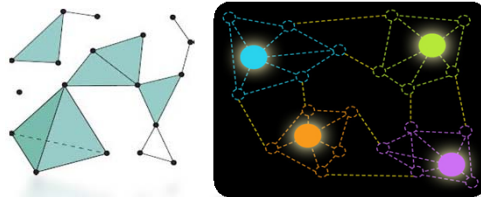
Network Layer

Advances such as software defined networks are changing stove piped network management to a heterogeneous management problem which requires dynamic assessment
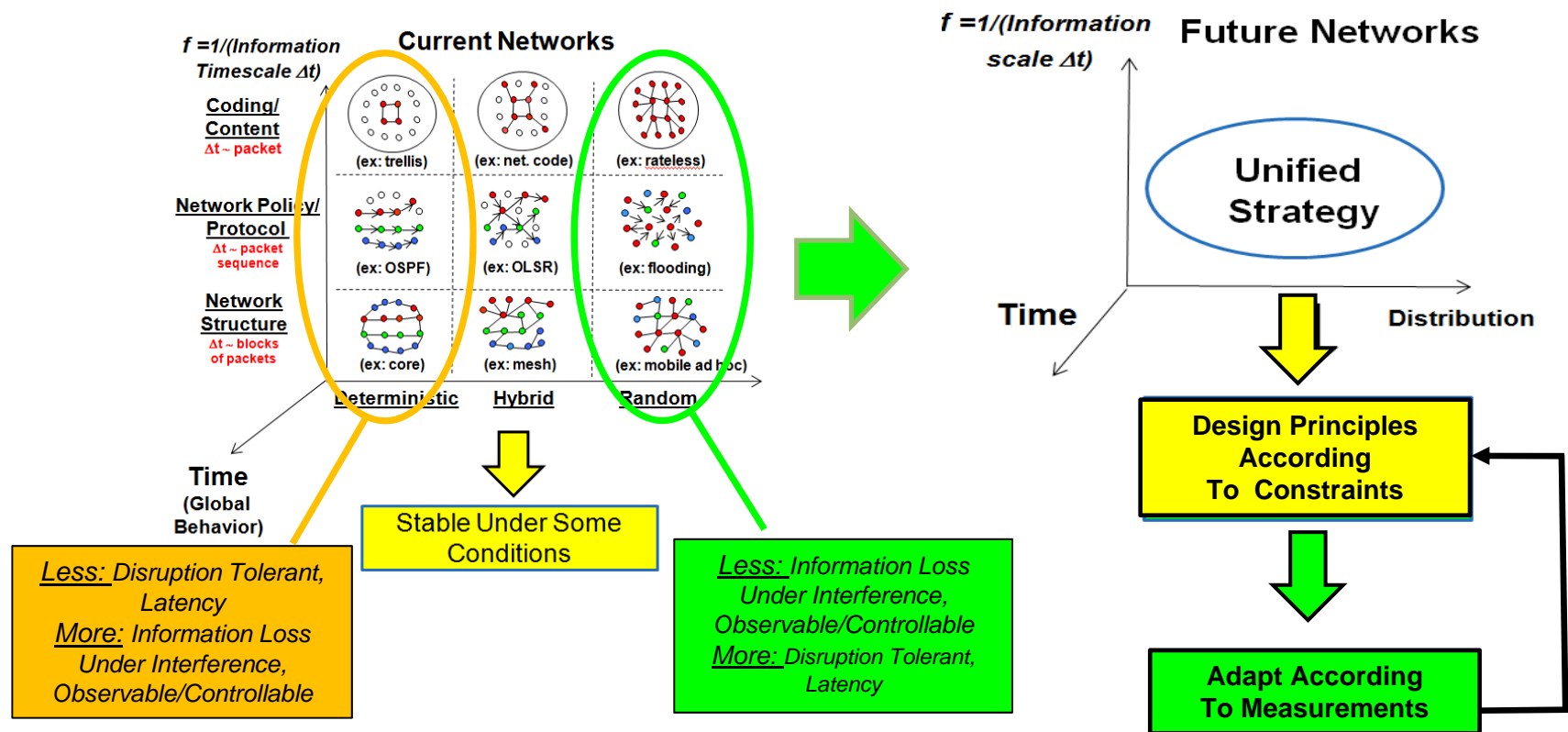
Physical Layer
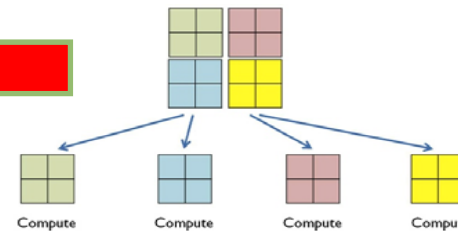
Commercial pressure on spectrum is changing the static and highly segregated assumptions about physical layer performance.

Current State – Static/stove-piped

Future State – Highly coordinated/ & dynamic

Integrated Mission Performance

Measure

Model

Manage

Predict/ Design

Sensor Performance

Network Performance

Software Computing

Under-resourced Unstable

Resourced Stable
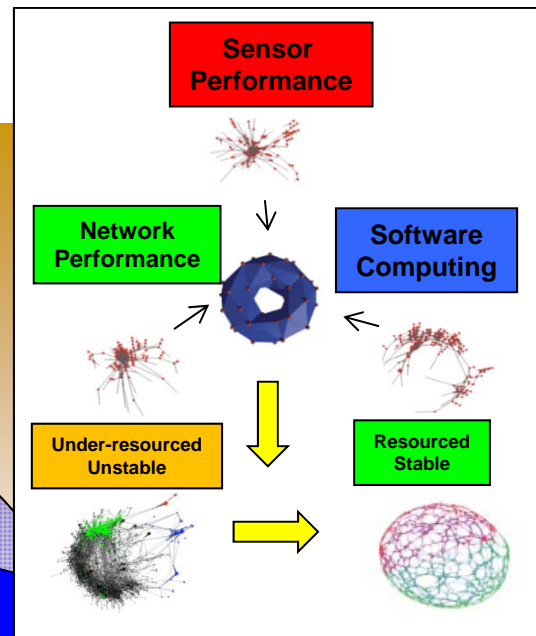
Unified Operation

Measure and verify information system properties among various system constraints

Measured Performance Regions

Heterogeneous Information

Network States packets, packet blocks packet groups)

Software States (variable, subroutine, program)

Hardware States (register, ram, virt. mem)

(timescale/level of abstraction)

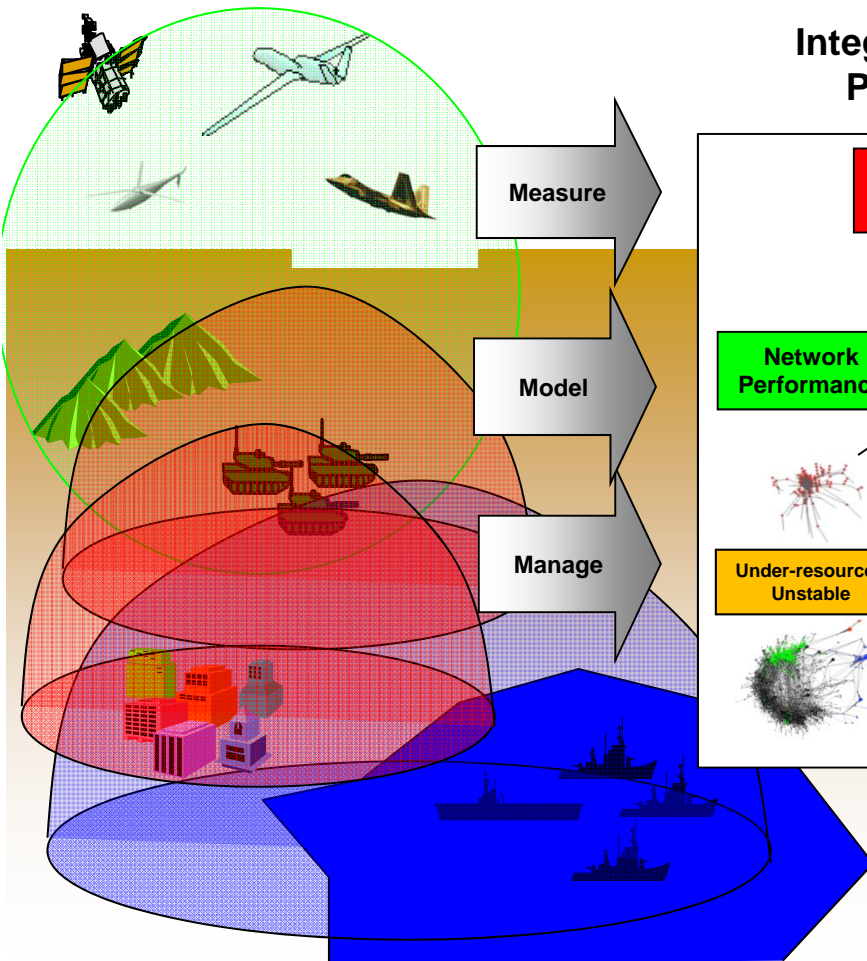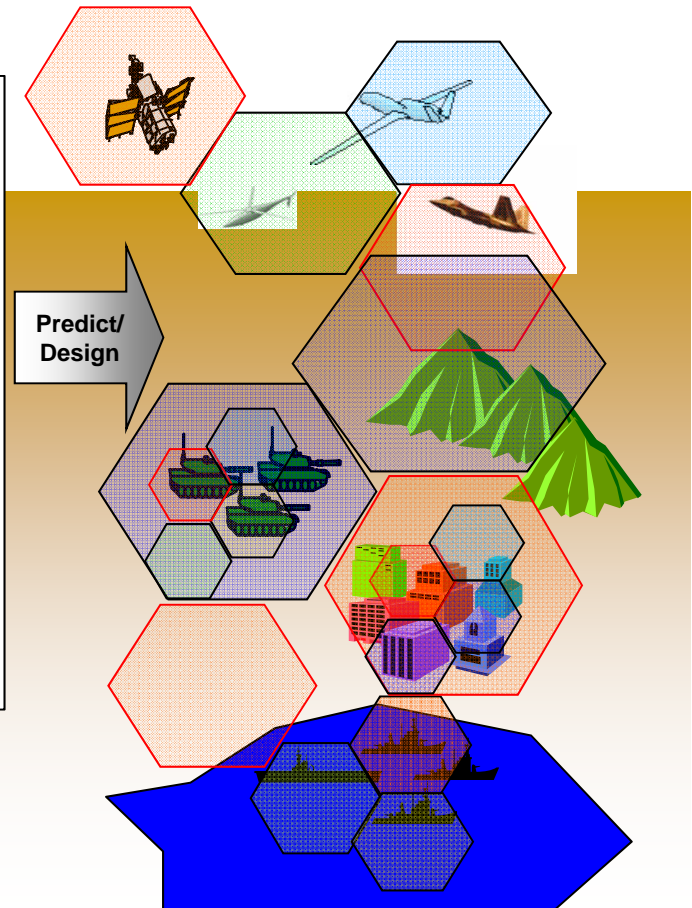**System Measurements**

Deterministic Content

Hybrid Content

Random Content

Deterministic Protocol

Hybrid Protocol

Random Protocol

Deterministic Architecture

Hybrid Architecture

Random Architecture

Less: Information Loss Under Disruption/Live
More: Latency, Resource Intensive/Safe

Best Integrated Performance Region

Less: Latency/Disruption Tolerant/Safe
More: Controllable/Live

Statistical Properties

Global Properties

Unstable/Un-resourced Insecure

Stable/Resourced Secure

Unified Operation

Units of information translate across heterogeneous domains and can be used to measure and quantify system performance

- Taking this approach can lead to a unified systems and security strategy

Basic Information Unit Scales

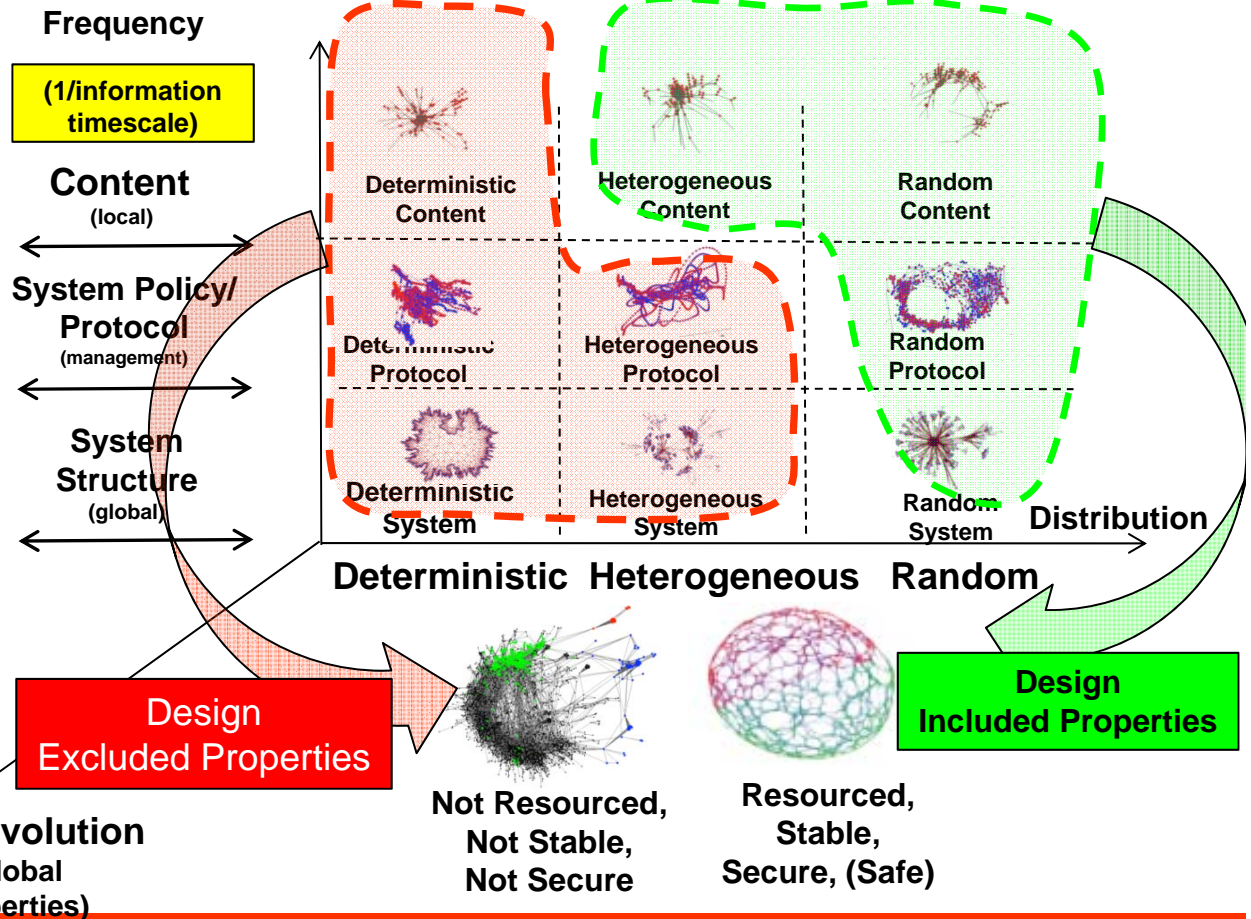| Data Network	Wireless Network	Hardware/ Software	Social	Biological
Packet	Modulation Unit	Register/ Variable	Words	DNA
Packet Groups	Waveform	Ram/ Subroutine	Phrases	Protein Synth.
Packet Blocks	Signal Array	Virtual Mem./ Program	News Reports/ Blogs	Cell Function

Digital Systems   General Systems

Frequency
(1/information timescale)

Content (local)

System Policy/ Protocol (management)

System Structure (global)

Time Evolution (Global Properties)

Measured System Properties

Deterministic Content | Heterogeneous Content | Random Content

Deterministic Protocol | Heterogeneous Protocol | Random Protocol

Deterministic System | Heterogeneous System | Random System

Deterministic   Heterogeneous   Random

Distribution

Design Excluded Properties

Design Included Properties

Not Resourced, Not Stable, Not Secure

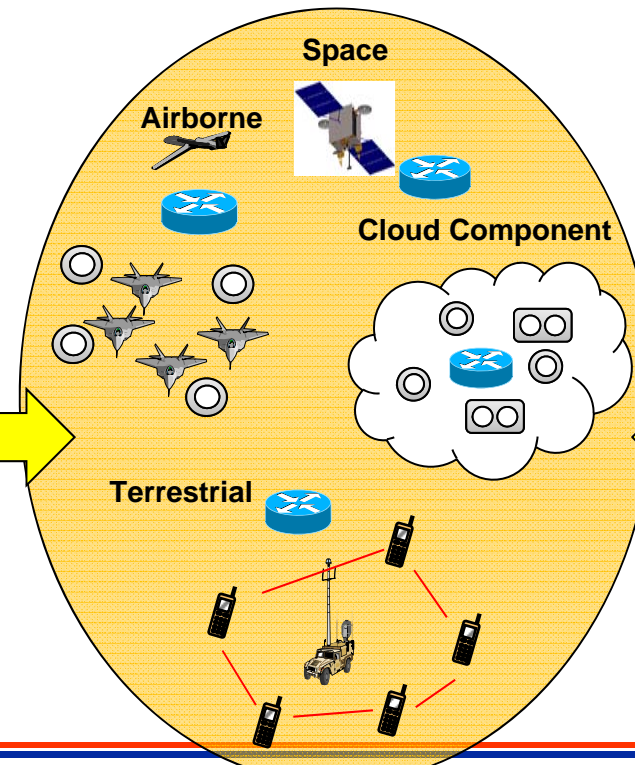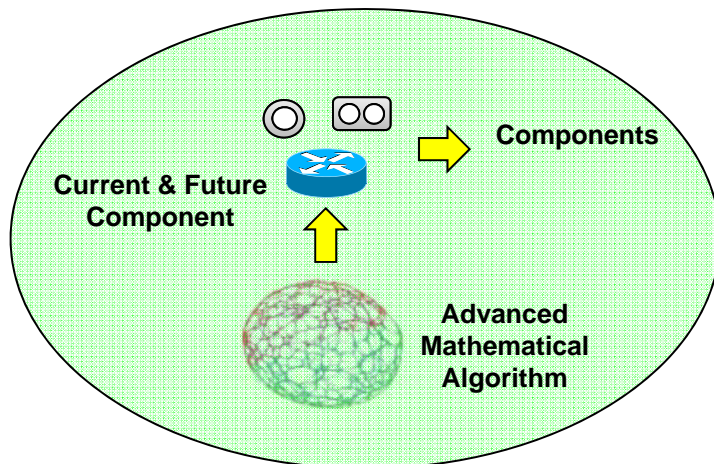**Resourced, Stable, Secure, (Safe)**

Current & Future DOD Architectures

An integrated framework to measure, model, and manage mission performance from the application to the physical asset enables to achieve mission performance guarantees in its future infrastructure

Introduce Into DOD Systems

Measure Model and Manage Using Advanced Data Methods

Components

Current & Future Component

Advanced Mathematical Algorithm

Space

Airborne

Cloud Component

Terrestrial

Enable Mission Performance Guarantees

Mathematical Systems Analysis