

Workshop

Intelligence Analysis and Processing for Non-Physical Target Systems

IDGA Intelligence Analysis and Processing Conference 2007

Instructor: Ed Waltz
Ed.waltz@baesystems.com
703-622-4385

Distribution Statement A:
Approved for public release;
distribution is unlimited

Outline

- Introduction: Non-Physical Systems
- The Role of Fusion in Non-Physical Nets
- Level 1 Fusion of Non-Physical Entities
- Level 2-3 Fusion in Non-Physical Nets
- Tracking the Dynamics of Non-Physical Nets
- Current Challenges and Key Technologies

Introduction: Non-Physical Target Systems

Why are Non-Physical Systems Important?

Military Warfare is no longer waged and won by attrition of physical assets.

It hinges on the attrition of will and perception.

What makes this type of Warfare Different?

Outcomes involving Physical assets are more predictable; such assets lack free will, and conform to the laws of physics.

Non-Physical assets have free will, and exhibit unpredictable 'emergent behavior'.

Non-Physical Target Systems

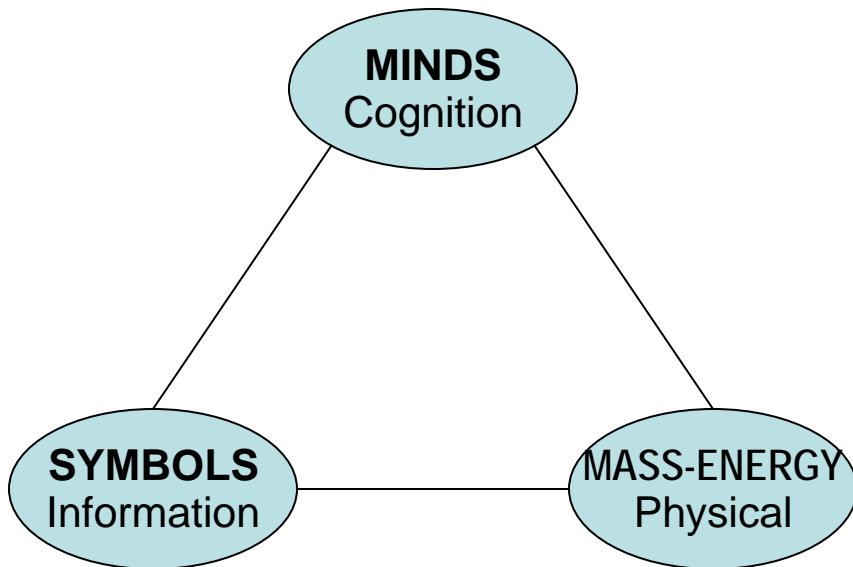
- Systems that are comprised of material and non-material elements, whose primary dynamic is based on the relationships and interactions of non-material elements (e.g. human cognition and emotion).
- Target systems primarily comprised of leaders, organizations, institutions and populations that include political, social, economic, military and information systems; these systems interact in cooperation, competition and conflict.

Terms

- **Physical System** – A system that has behavior that is predominantly driven by the laws of physics
- **Non-physical System** – A system that has behavior that is predominantly driven by human cognitive-emotive activity.
- **Human Factors** – The psychological, cultural, economic, and other human attributes that influence decision-making, the flow of information, and the interpretation of information by individuals and groups at any level in any state or organization.
- **Human Terrain** – The social, ethnographic, cultural, economic, and political elements of a population
- **PMESII** – Political, Military, Economic, Social, Infrastructure, Information System ontology

Physical and Non-Physical: Three Domains of Reality

- Intelligence targets exist as objects in three fundamental domains of reality – *and should be modeled as such:*



- Basic ontological distinctions made by Aristotle, Peirce, theologians (mind, word & deed) and semiotic philosophers

Categories of Objects (Entities)

DOMAIN

CHARACTERISTICS

COGNITIVE

- Human Cognition-Emotion
- Perception
- Conception
- Intention



SYMBOLIC

- Explicit information
- Communication and Media content



PHYSICAL

- Mass and energy
- Human bodies, Infrastructure, Weapon Systems



NON-PHYSICAL

PHYSICAL

Three Perspectives of Human Targeting

A. Human Factors and Decisionmaking

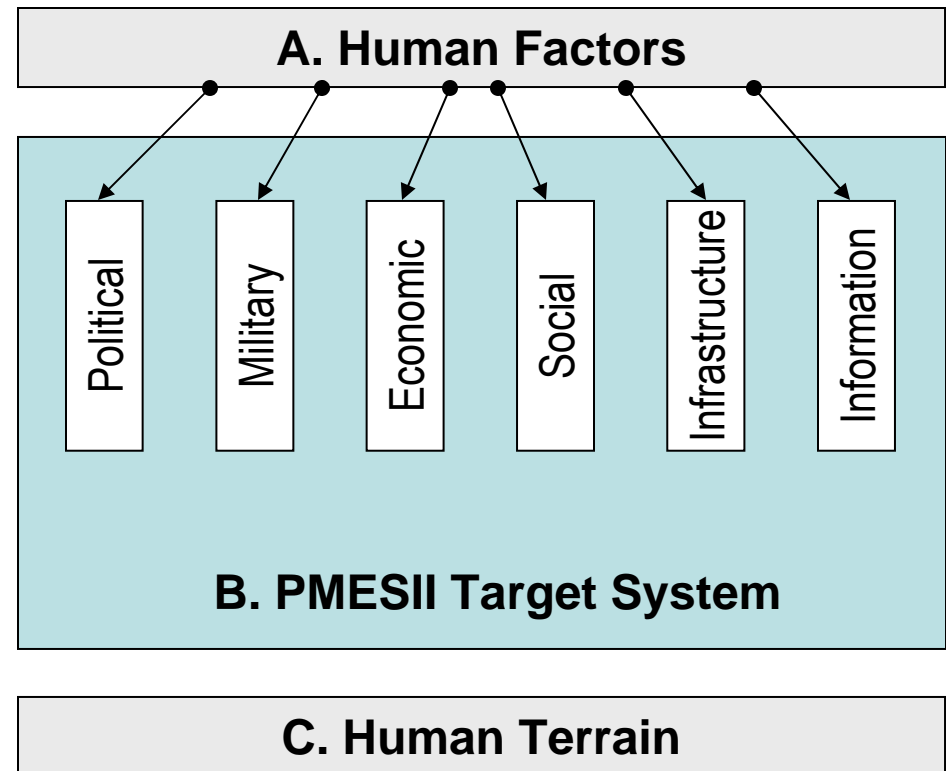
- DIA Human Factors Analysis Center (HFAC)
- JIOWS (San Antonio, TX)
- NASIC (Dayton, OH)

B. PMESII – System of Systems

- JFCOM Effects based approach
- JWAC

C. Human Terrain – Human Context

- JIOC (Joint Intelligence Operations Centers)
- Army, Interagency



A. Human Factors Analysis

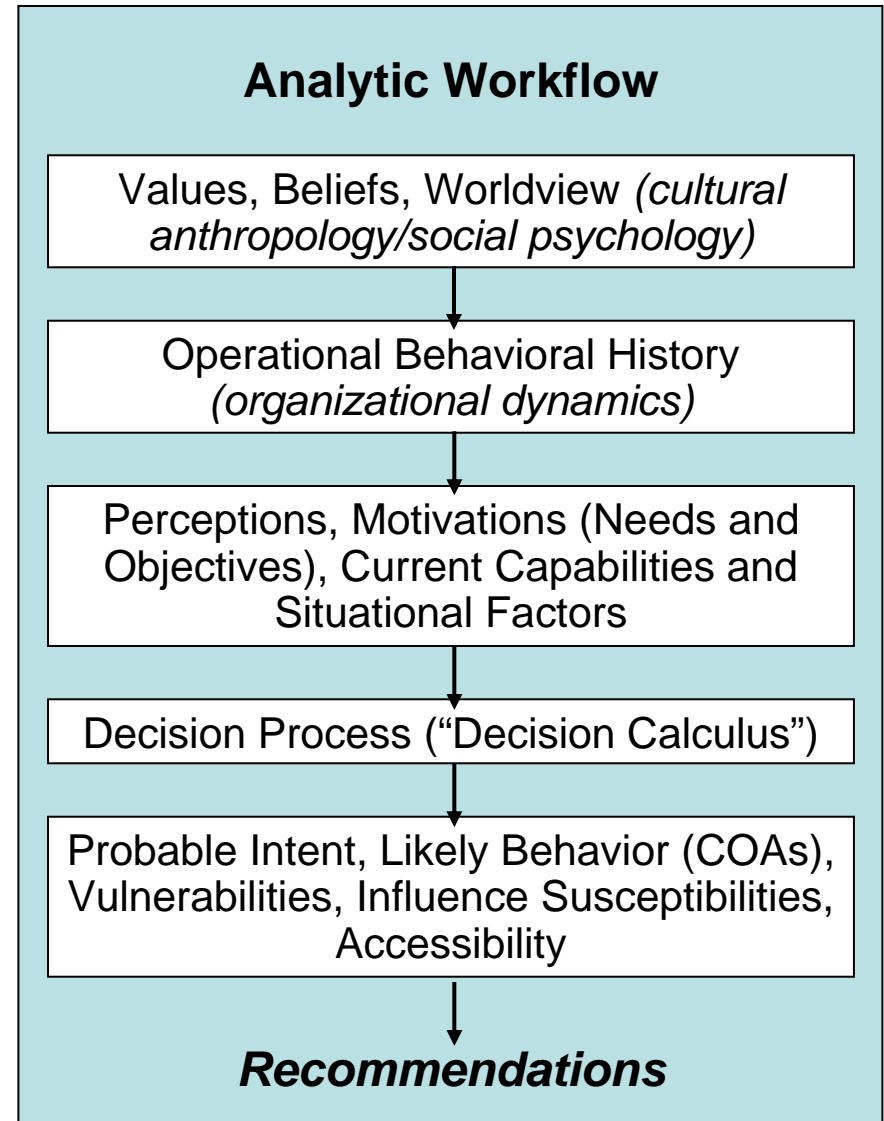
- **Human Factors:**

The psychological, cultural, behavioral, and other human attributes that influence decision-making, the flow of information, and the interpretation of information by individuals and groups at any level in any state or organization DCID 7/3 (5 June 2003)

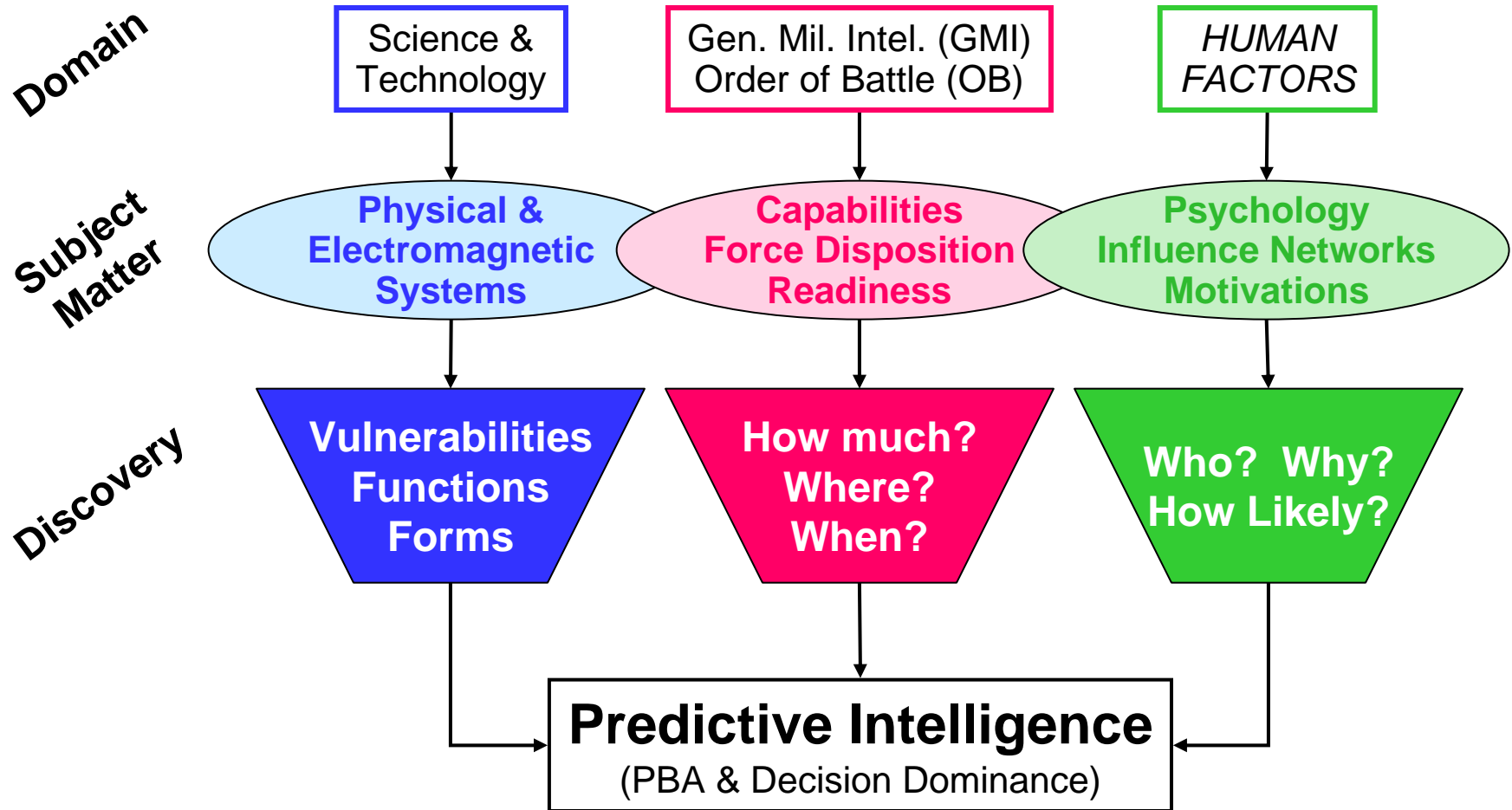
- **IO applications of HFA focus on influencing individuals by:**

- Computer Network Operations (CNO)
- Military Deception (MILDEC) Operations
- Psychological operations (PSYOP)

Source: Greg Jannarone, *Behavioral Influences Analysis Workflow*, USAF Behavioral Influences Analysis (BIA) Center, Maxwell AFB, Alabama



A: Integrating Human Factors Analysis



Source: Greg Jannorone, Behavioral Influences Analysis Workflow, USAF Behavioral Influences Analysis (BIA) Center Maxwell AFB, Alabama

B. Physical and Non-Physical Systems of Systems: PMESII

Political	Military	Economy	Social	Infrastructure	Information
Political Struggle; leadership influence	Armed forces Paramilitary Insurgency	System for producing, managing wealth	Human society, culture, characteristics	Facilities, services, installations to support society	System for producing, disseminating knowledge
<ul style="list-style-type: none"> National, provincial and local government Tribal leadership Competition for governing ideas Unarmed opposition groups Local and global ideological support, fundraising 	<ul style="list-style-type: none"> Organized government military organization, order of battle, readiness, force strength Organized paramilitary Armed insurgency 	<ul style="list-style-type: none"> GDP; GDP Growth; Inflation Trade balance (import, export, capital inflow) Construction; Public Finance; Debt Economic status of population elements, shortages, subsidies 	<ul style="list-style-type: none"> Security/Law & order (includes crime & criminal organizations) Public Health; mortality rates, disease rates Demographics; presence, distribution in city & environs 	<ul style="list-style-type: none"> Electrical Power production Water, sewer Transportation efficiency factors Manufacturing production Gas, Petroleum production, flow rates, efficiency Telecommunications bandwidth, coverage 	<ul style="list-style-type: none"> Broadcasting/publishing/website organizations Local, foreign (including US) media channels Media, content originators (political/social groups, writers, producers)
Human Behavior Analysis Focus			Human Terrain Focus		

B: Describing System of System Effects – Smith’s Categories

Area	Effects
Physical	<ul style="list-style-type: none"> • Destruction – Destroying an opponent’s physical capability to do something is clearly an effect. • Attrition – Incremental destruction of capability
Psychological	<ul style="list-style-type: none"> • Shock – Shock centers around a sudden collapse of the foe’s belief in his ability to produce an acceptable outcome in a given situation. It may be paralleled by a sense of despair and resignation to a fate that can no longer be avoided. • Attrition – Whereas shock is a sudden collapse of belief, psychological attrition is the product of gradual erosion of the will. • Foreclosure – Curtailing an opponent’s options or foreclosing potential courses of action. <i>Active foreclosure</i> is the use of military forces to block a course of action that a foe has already initiated. <i>Passive foreclosure</i> revolves around the fact that a sufficient array of capabilities exists or can be brought to bear in a given region to prevent a destabilizing course of action from being carried out.
Both Physical and Psychological	<ul style="list-style-type: none"> • Chaos/Entropy – Such confusion created in the minds of decision makers conveys that they are unable to command or fight their forces effectively. Psychological factors (disorientation, confusion, fear) yield a physical result, the inability to use otherwise applicable forces and capabilities; however the cause of this effect is inherently rooted in behavior, the inability of commanders and subordinates to handle the pace and scope of operations.

C. The Human Terrain

- **Human Terrain** – The social, ethnographic, cultural, economic, and political elements of the people among whom a force is operating [1]
 - Elements of the human terrain:
 - Entities (actors): People (Leaders, terrorists, agents), groups (families, clans, organizations) or populations
 - Relationships: Political, social, ideological, economic, armed influence
 - Context: Norms, metanorms (cultural, ideological, legal), economies (financial, information)
- **Civil Intelligence** – Intelligence derived from all sources regarding the social, political and economic aspects of governments & civil populations, their demographics, structures, capabilities, organizations, people, and events. [2]
 - Civil Considerations – The influence of manmade infrastructure, civilian institutions, and attitudes & activities of the civilian leaders, populations, and organizations within an AOR on the conduct of military operations (FM-06)

Sources:

[1] Montgomery McFate and Andrea Jackson, “An Organizational Solution for DoD’s Cultural Knowledge Needs”, *Military Review* (July-August 2005)

[2] Urban Sunrise – DARPA IXO; General Dynamics, Dec. 2003. © Ed Waltz 2007



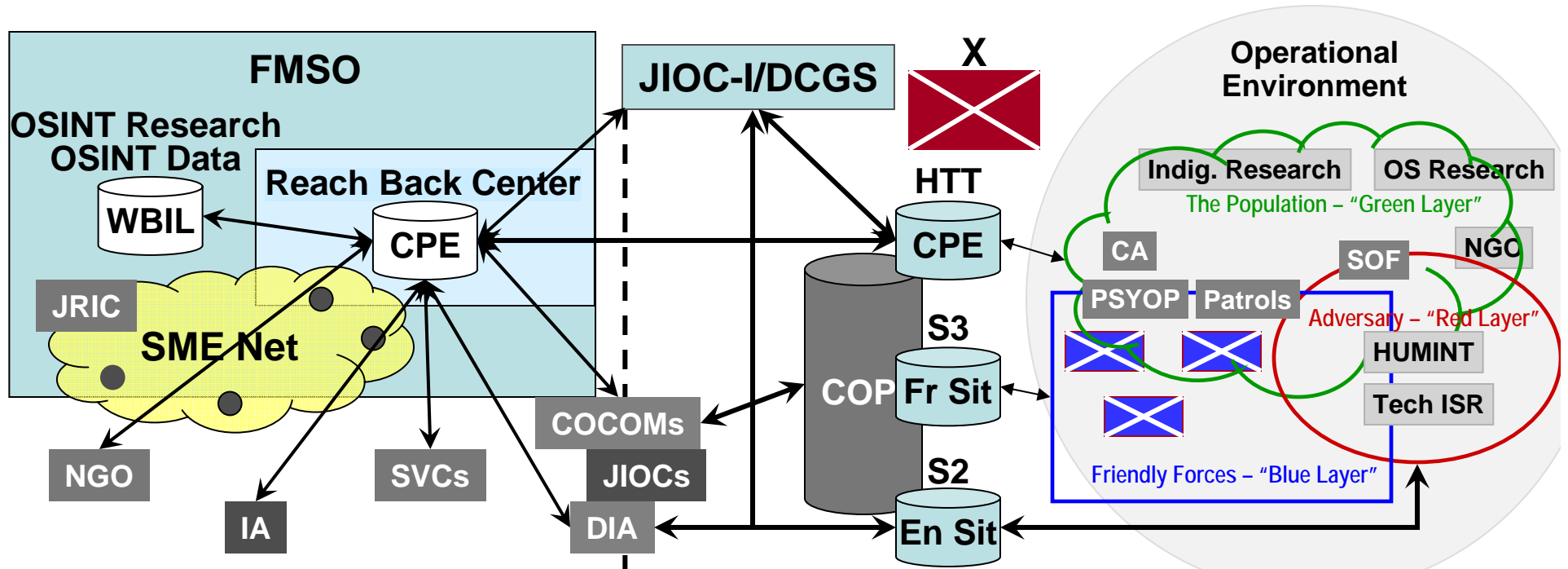
C. Human Terrain: Elements of Counterinsurgency – COIN*

Category	Functions-Factors Modeled	Factors included in Representation
Political Goals	Objectives and ideology base	Multiple Goal sets of actors as future states and desired levels of capability and capital (relative political, social, economic or armed power)
Contextual Factors	Environment – Geography	Geospatial area representation with key features, locations of fixed entities, lines of communication, infrastructure, etc.
	External influences (Regional, International)	External sources of ideological, financial, and material resource support.
	Environment – Rule of Law	Level of rule of legitimate authority; effectiveness of government services (Security, justice, health-sanitation, utilities, etc.) as a function of degree of corruption.
	Environment – Political Economic base	Stability of political power struggles; Strength of economic base in the region, state and local area modeled.
Insurgent Actors	Insurgent Leaders	Principal insurgent actors, their goals, relationships, leadership abilities, capital (relative economic, political, social and armed power) and resources.
	Combatants	Armed actors and capabilities of the military arm of the insurgency, their goals, resources and targets
	Political Cadre (party-militants)	Political actors of the insurgency, lines of influence and political capital
	Auxiliaries (active supporters)	Actors (groups) that provide sources of support
	Supporting population base	Civil population that support the insurgent cause
Insurgent Resources and Processes	Financial Sources (crime, external supporters, military acquisition, etc.)	Resource generation processes to create economic capital for recruitment, force containment, munitions and arms acquisition, etc.
	External support (foreign funding, ideology, personnel, information)	External supplies of economic and material support; external sources of information (intelligence) to increase effectiveness of activities.
	Recruitment	Rate of acquisition of combatants, political cadre and popular support.
	Training	Effectives of transition of recruits to effective combatants
Info Ops Processes	Media channels from sources to targets; effectiveness of info transfer and impact	Government and Insurgent processes (sources, messages, media channels, targets) of media to influence tone of news, belief in situation, and spread of positions or ideology.
Neutral Actors	Non-government organizations (NGOs), neutral populations, etc.	Goals, activities and influences of neutral organizations, institutions and populations.
Civil Population Groups	Mass (domestic) civil population for which government and insurgency seeks legitimacy and support	Population groups, goals, beliefs, and behaviors representative of demographics for the area modeled; relationships between groups represented according to influence, attraction, economics, etc.

* Adapted from FM3-24

© Ed Waltz 2007

C. Human Terrain System Operational View



Cultural Knowledge Center

- Consolidates and archives cultural data, information and knowledge from across all disciplines and communities
- Provides SME access and assessments
- Coordinates cultural research efforts
- Leverages OSINT research and data
- Provides current cultural knowledge to planning, preparing, training, experimentation and decision-making activities

Organic Human Terrain Team (HTT)

- Captures, consolidates, tags and ingests cultural data
- Provides organic assessments
- Provides conduit to SMEs
- Assists commander and staff in understanding the "so what" of cultural information
- Conducts data-handoff to joining forces

Data Gathering in Theater

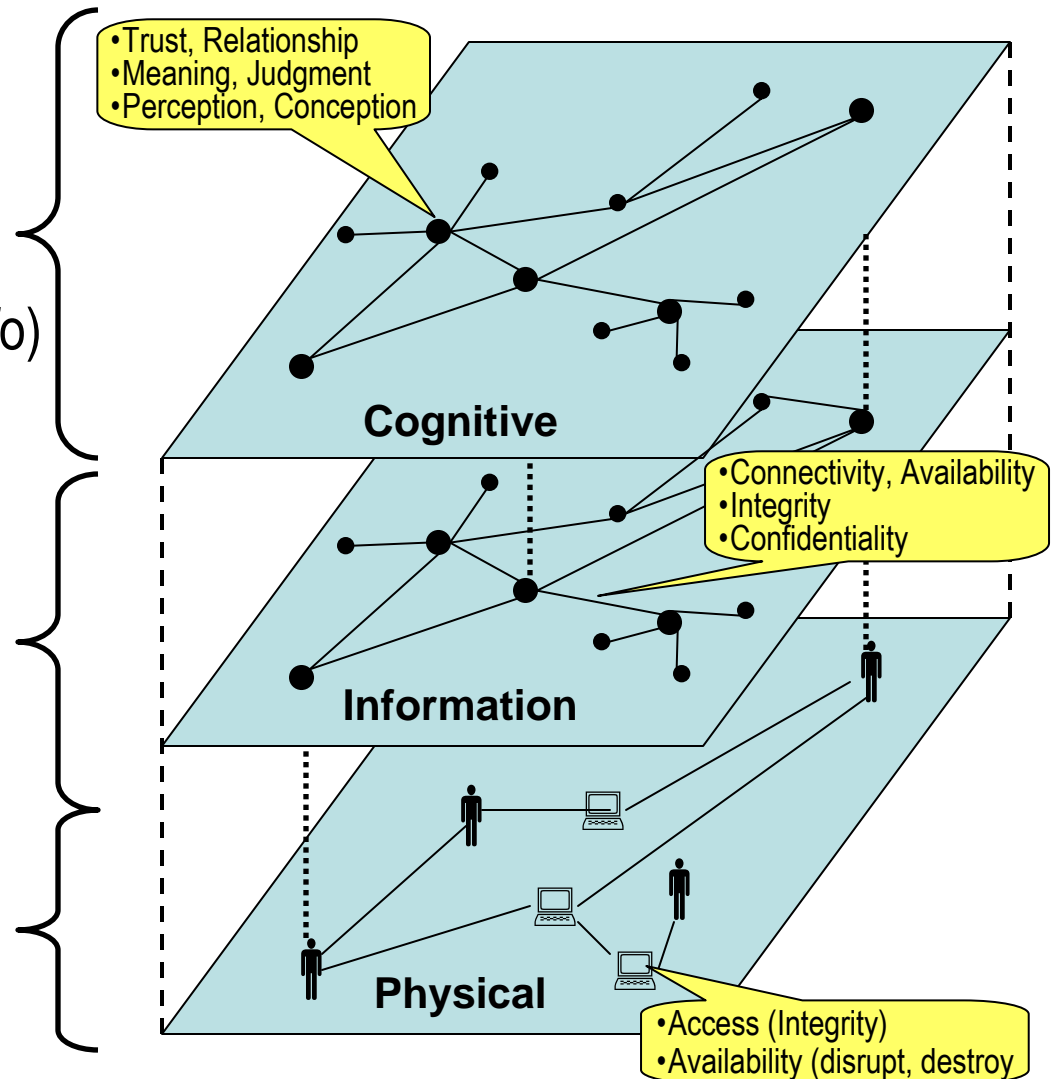
- Inputs from all sources
- Field Research Teams conduct focused cultural research
- Contact with non-BCT elements

Source: HTS Briefing, U.S. Army Training and Doctrine Command, Deputy Chief of Staff for Intelligence, August 2006

© Ed Waltz 2007

Information Operations: Networks Across Three Domains

- PSYOP, Deception
 - Cognitive effects on individuals
 - Perception, conception, belief
 - Meaning and values
 - Trust in the network (sources of info)
 - Tipping points of opinion
- Computer Network Operations
 - Effects on Information Network
 - Responsiveness, Latency
 - Accuracy, Integrity, Reliability
 - Information Flow
- Physical Attack
- Electronic Attack



Role of Fusion in Non-Physical Nets

Outline:

- Current manual methods of social network analysis and the tools
- Placing Net Analysis in the Data Fusion paradigm
- The role of fusion automation and current tools
- A workflow for non-physical fusion

Current Manual Methods of Analysis

- **Pattern Analysis Plot**
 - Records Data Time Group (DTG) of Events
 - Visualizes Repetitive Patterns over Time
 - Enables the analyst to visualize operating patterns
- **Association Matrix**
 - Association of individuals
 - Designate confirmed vs suspected associations
 - Can include deceased individuals
- **Event Matrix**
 - Relate people associated with events or activities
 - Horizontal associations (individual participation)
 - Vertical associations (roles of individuals)

Using automated tools, e.g. i2 Analyst Notebook®

- **Map entities and links from a database or manual entry**
 - Link diagrams
 - Timeline diagrams
 - Event diagrams
 - Activities matrix
- **Enables analyst to:**
 - *Accumulate* (marshal) evidence
 - *Visualize* evidence in temporal, relationship, event-activity views
 - *Maintain* and *Evaluate* alternative competing hypotheses

Applying the JDL Data Fusion Model to Non-Physical Systems

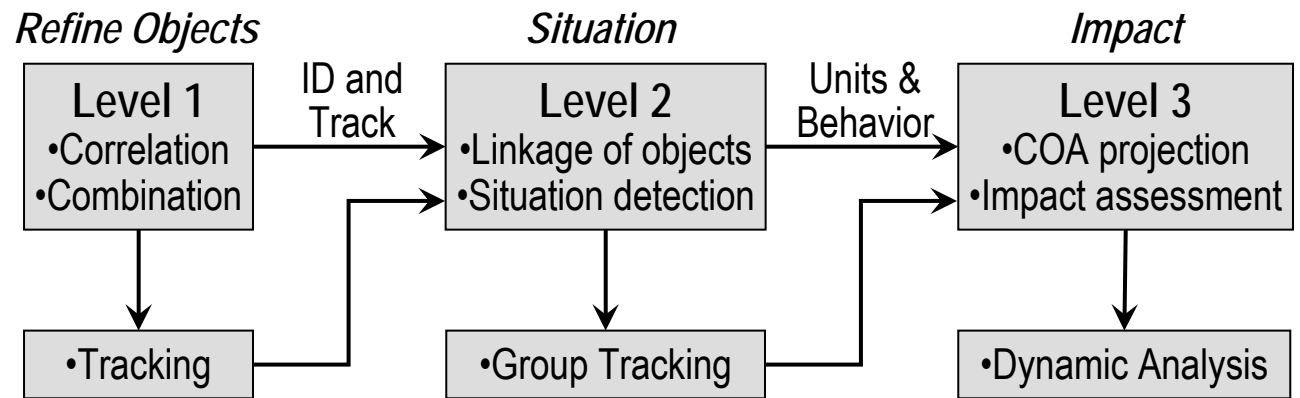
- Mapping traditional JDL fusion notions onto non-physical network fusion
- Terminology (JDL : NP Fusion)
 - Object : Entity
 - Group, Unit : Group, Network, System
 - Tracking : Dynamic model
 - Kinematics : Dynamics (social dynamics, network dynamics)

Level	Fusion Function	Non-Physical Net Fusion
Level 1 Object Refinement	Correlation	• Entity correlation
	Combination	• Entity identification; attribute linking from multiple sources and entity resolution
	Tracking	• Entity behavior over time; contact and network growth • Entity timeline
Level 2 Situation Refinement	Correlation, Association	• Group detection
	Combination Group Formation	• Group attribute combination and ID • Social Network Analysis (SNA) of static structures
	Group Tracking	• Group changes over time • Group action timeline
Level 3 Impact Refinement	Impact estimation	• Social system behavioral dynamics • Anticipate COA and Effects analysis • Effects-based approach to Planning

Comparison of Physical and Non-Physical Data Fusion

Physical Systems

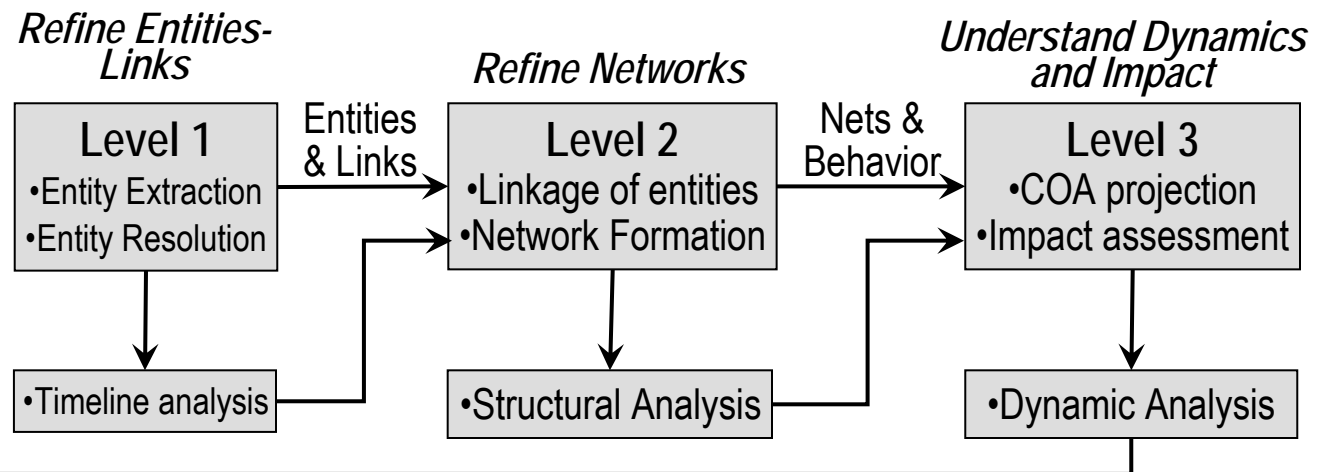
- Stochastic system
- Physical constraints
- Sensors
- Passive observation
- High-update rate based on sensor events



Targets as predictable linear systems (Estimation & Prediction)

Non-Physical Systems

- Complex system
- Social constraints
- Sources
- Active probing
- Sporadic update based on target events and probes



Requested Probing Actions

Targets as non-predictable complex adaptive systems (Exploration & Probing)

Basic Approaches to Describing the Network

Approach Characteristics	1 Enumeration	2 Relation	3 Simulation
<i>Description</i>	<i>Enumerate</i> the entities within the organization; identify, characterize and index	<i>Model the structure</i> of relationships with the organization; identify and characterize formal roles, responsibilities as well as actual, informal roles and relationships.	<i>Simulate the behavior</i> of the organization and the dynamics of interaction among entities, as a function of perception of the environment, its resources, & goals.
<i>Analytic Focus</i>	SIZE: Organization scope, size, and scale of personnel, resources, locations, capabilities	STRUCTURE: Relations (roles, responsibilities, linkages) between organization entities	BEHAVIOR: Function and performance of individuals interacting; operational effectiveness of the functioning organization
<i>Modeling Tools</i>	<ul style="list-style-type: none"> Organized data sets in Databases 	<ul style="list-style-type: none"> Relational Databases Network graphs Statistical analysis 	<ul style="list-style-type: none"> Computational Social Science Social Simulations
<i>Typical Elements of Organization Model</i>	<ul style="list-style-type: none"> Lists of nodes (entities) Personnel profiles Psychological analyses Timelines of activities 	<ul style="list-style-type: none"> Social Network Graphs (nodes and links – relations) Organizational Net Analysis (roles) 	<ul style="list-style-type: none"> Actors and goals Environments Resources
<i>Examples</i>	<ul style="list-style-type: none"> Modernized Intelligence Database (MIDB) Operational Net Assessment (ONA) Database 	<ul style="list-style-type: none"> Link Analysis (e.g. Pathfinder; Analysts Notebook) Data Mining (automated induction of nets) e.g. Clementine 	<ul style="list-style-type: none"> Large-scale social simulations

Source: Waltz, Ed “Practical Approaches to Impact Adversary Decision-Making Processes” in “Organization Decision-Making and Information Warfare”, Artech 2007

© Ed Waltz 2007

Processing and Analysis Tools

Tool Category	Tool and Functions	Example Commercial and Government Tools
Enumeration	Field Collection – Collect narrative and structured data on individuals, relationships, events; for rapid codification and transfer to databases	CHIMS (CounterIntelligence-HUMINT Information Management System)
	Organization Intel Databases – Local data stores to ingest and index multiple sources of data	Heterogeneous Relational Databases, SQL, Oracle
	Processing Translation – Language translation	Apptek Machine Translation
	Processing Extraction – detect and extract named entities and relations from unstructured data; resolve named entities	NetOwl, ClearForest, Convera Retrievalware, AeroText
	Data Warehouse – to hold unstructured data and extracted index data	NCR Terradata
Relation	Link Discovery – Detect relevant network fragments and nets; screen and assemble groups and networks using knowledge in ontologies of the problem domain to determine relevance	Pathfinder; LGAT; Semio Taxonomy; Saffron Net; Inxight ThingFinder and Smart Discovery
	Social Network Analysis – Computation of network metrics for individual actors, groups and entire networks. Perform comparison of groups and nets by structural equivalence properties to detect groups that meet structural criteria	UCINET; InFlow; DyNet; VisualLyzer, MetSight;
	Network Visualization – Present entities and relationships in multiple viewing perspectives: Network node-link, Timeline and Causal views	i2 Analyst Notebook's Case and Link Notebook; Visual Analytics, StarLight, InXight StarTree; Orion Magic
	Geographic Visualization – Overlay network data on geospatial locations of entities and events	ESRI's ArcView, ArcIMS; Mapvision
	Relational Database – To store organization model hypotheses	Oracle 9i
	Organization Metrics – Assess organization performance, effectiveness, structural vulnerabilities against standard fitness criteria	OrgCon
Simulation	Organization State Models – Hidden Markov models track behavior changes of organizations that exhibit distinct states with observable transitions (e.g. C2 op sequences)	See models described in Chapter 2*
	Organization Behavioral Models – Dynamic models of actor behavior by representing transactions of social resources or influence, e.g. dynamic Bayes networks, agent based simulations, etc.	SAIP (SAIC Influence net); DyNet

Commercial Analysis Toolflow (Examples)

OPEN SOURCE

- InXight Crawler
- Convera Spider
- Virage Video Logger
- Copernic Pro

IMINT

- InterGraph Image Analyst
- ERAS Imagine

GEOINT

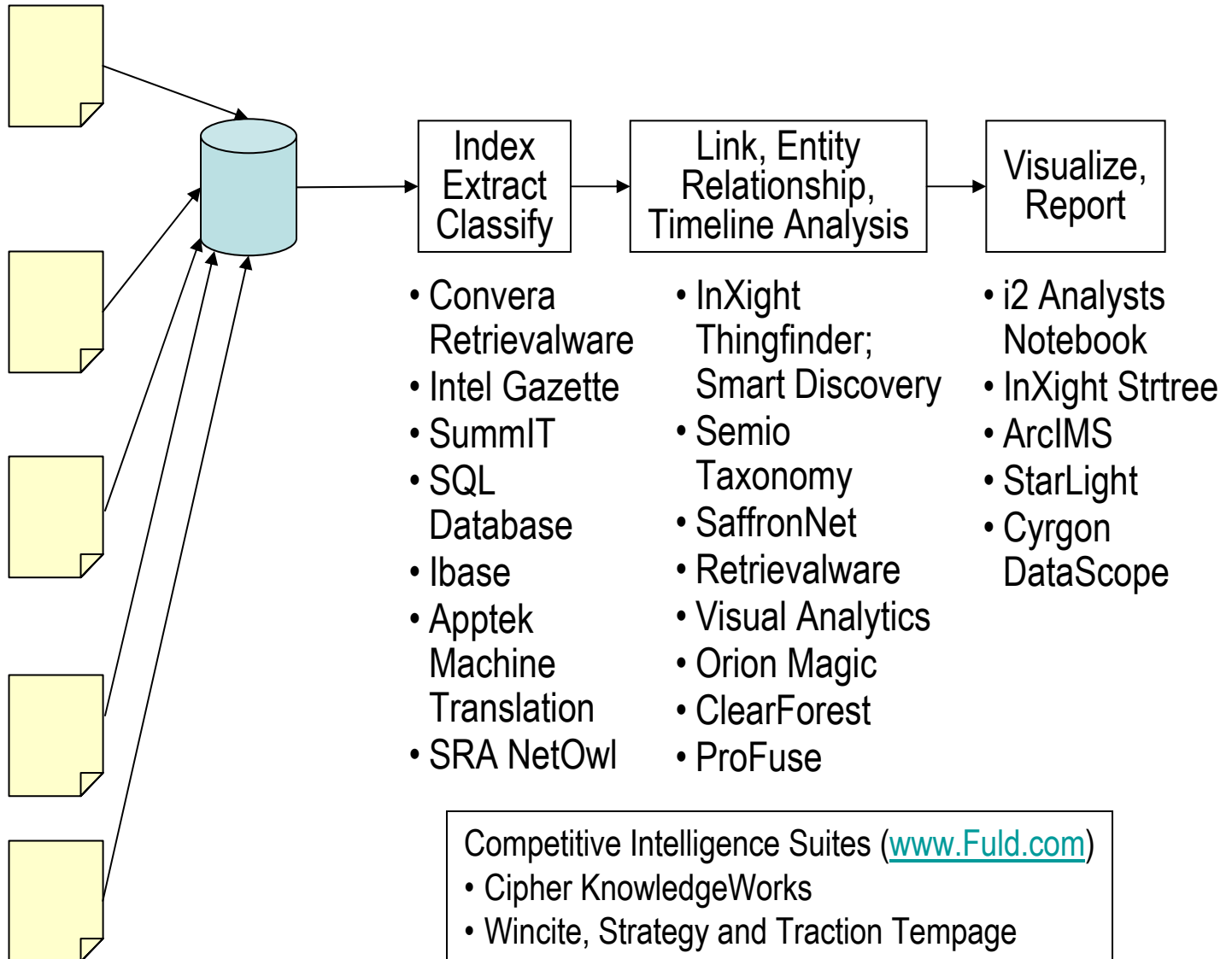
- Keyhole
- ArcGlobe, ArcView
- AutoDesk Map 3D

SIGINT, MASINT

- Clementine
- SPSS

Warning

- IntelliSeek Brandpulse
- Netrocity RealTime Event Radar



Level 1 Fusion of Non-Physical Entities

Outline:

- Level 1 Non-physical entity refinement (JP 3-13)
- Sources of data and methods of entity extraction
- Correlation and resolution

Sources of Information

Domain of Analysis	Required Information Elements	Typical Sources of Information
Cognitive Properties of the Organization	<ul style="list-style-type: none"> • Key Individuals – Identity and psychological profile of key leaders and influencers (e.g. advisors, key associates, and/or family members) affecting attitudes, perceptions and decision making environment. • Decision making calculus – How decision makers perceive situations, reason about them, plan actions, execute and then assess outcomes of their actions; the cultural basis of their perspective and doctrines applied to decision making • Organizational and Social Structure – Formal organizational structure (roles, responsibilities) and informal social relationships • Cultural and Societal Factors – Affecting attitudes and perceptions (e.g. language, education, history, religion, myths, personal experience, family structure) • Credibility Assessments – Of key individuals or groups; description of their sphere of influence. • Historical Factors – Key historical events that affect an individual or group's attitudes and perceptions of others. 	<ul style="list-style-type: none"> • Organization statements, reports, news releases, messages • Accumulated lists of members & associates, directories, attendance lists • HUMINT sources reporting on organization activities, plans, intentions • SIGINT and NETINT intercepted communications, message traffic, emails, courier messages • OSINT historical & cultural information, family & genealogical information
Information Properties of the Organization	<ul style="list-style-type: none"> • Information Infrastructure – Description of the capabilities of lines of communication, networks, nodes, their capacity, configuration, and the related computation. (This includes non-electronic networks such as rat-lines, dead-drops, etc., and other non-technical forms of communication.) • Technical design – Description of all technical elements of the information infrastructure, equipment, data models of all layers of the OSI stack. • Social and commercial networks that process & share information and influence (kinship & descent lings, formal & informal social contacts, licit & illicit commercial affiliations, and records of ownership & transactions, etc.) • Information – External traffic descriptions, internal content and context of information obtained from the information infrastructure. 	<ul style="list-style-type: none"> • SIGINT & NETINT intercepted communications, message traffic, emails, courier messages • HUMINT descriptions of technical components, networks and supporting infrastructure • OSINT commercial specifications, operating characteristics
Physical Properties of the Organization	<ul style="list-style-type: none"> • Geospatial Data – Describing location, geographic coordinates, and descriptions of organizations' facilities, infrastructure, physical lines of communication, etc. • Physical Site Data – Description of sites, facilities, power & communications equipment, critical links between physical locations 	<ul style="list-style-type: none"> • GEOINT imagery, terrain & supporting MASINT data about physical sites and facilities • HUMINT and OSINT descriptions of physical sites, construction, equipment.

Approaches to Network Data Acquisition

- Distinguish a variety of types and means of collecting network data:
 - Quantitative – Explicit data (phone, credit card, travel transactions, passport, internet data, etc.)
 - Qualitative – Tacit (knowledge collection from SMEs, human sources, inferred from other information sources)
- Use Tacit collections to help navigate through Explicit

	Approach	Sources	Entities	~
EXPLICIT ↑	Extract entities and links from large volumes of structured data	<ul style="list-style-type: none"> • Phone records • Online • Email • Bank transactions 	<ul style="list-style-type: none"> • Accounts • IP, MAC addresses • Nodes 	>10¹⁰
	Extract named entities & links from large volume of unstructured data	<ul style="list-style-type: none"> • Intel reports • HUMINT • Open Source • Transcriptions • DOCEX 	<ul style="list-style-type: none"> • People • Places • Things 	>10¹⁰
	<i>Mixed Approaches</i>			
TACIT ↓	HUMINT	<ul style="list-style-type: none"> • Interviews 	<ul style="list-style-type: none"> • People 	>10³
	Subject Matter Experts (SMEs)	<ul style="list-style-type: none"> • Interviews • Workshops • Simulations 	<ul style="list-style-type: none"> • Actors • Institutions • Populations 	>10³

RDF Resource Descriptor Framework

- W3C Standard – Semantic Web basis for encoding knowledge more richly than tables (relational databases) or hierarchies (XML)
- RDF approach – Labeled, directed graph expression where each vertex represents an entity, each edge represents a fact, or a relation between two entities
- Basis of Expression
 1. Statement: A fact is expressed as a Subject-Predicate-Object triple; Example:
<Ankhira_hotel><is_located_in><Baghdad>
 2. Subjects, predicates, and objects are given as names for entities (aka resources or nodes).
 3. Entities represent *something*, a person, place, or more abstract states.
 4. Names are URIs that are global in scope, always referring to the same entity across all applications and databases to allow search and association.
 5. Entities can be given as text values, called **literal values**

An Ontology Example

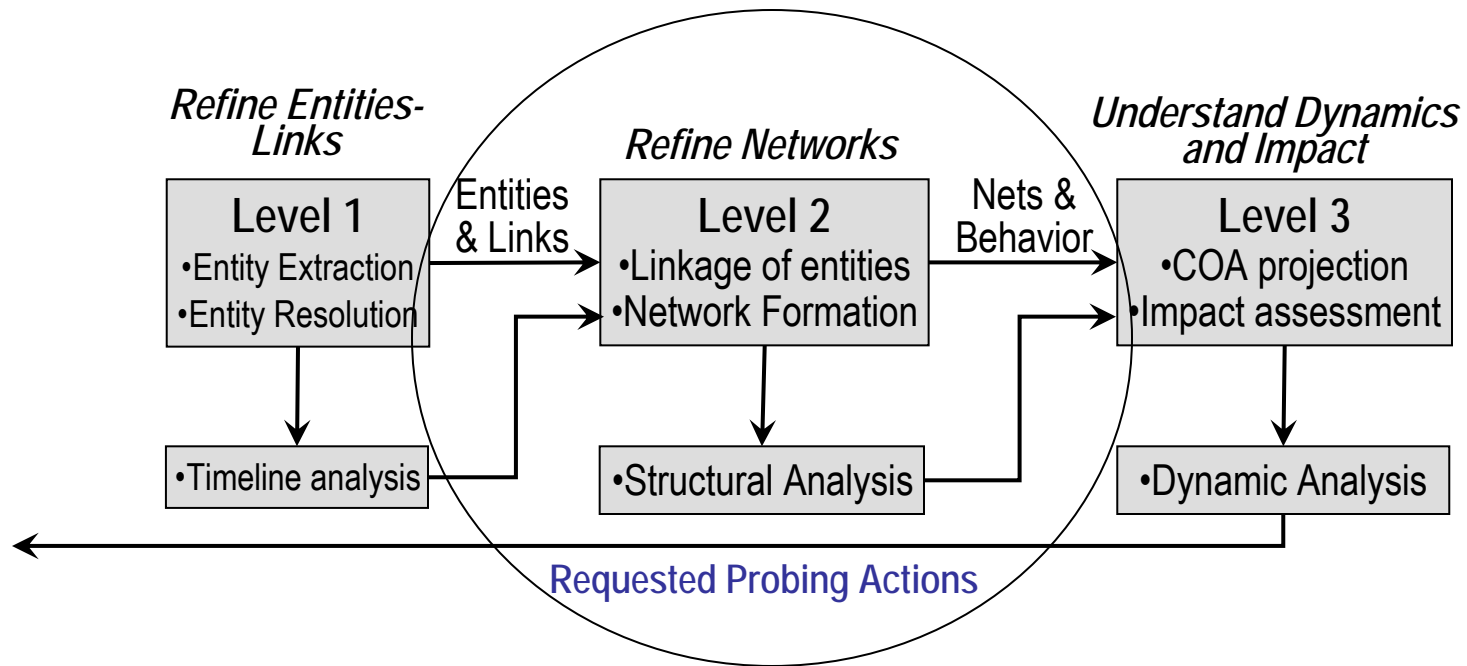
Ontology Concept	Examples from the general domain of Insurgency
Classes (general categories of things)	Leaders, combatants, supporters, financiers, organizations, border crossings, funding sources,
Instances (particular cases of entities)	Taliban, Mullah Omar, Mullah Dadulla Lang, Kulahn Pass
Relations: is_a, subclass_of, (kind_of), instance_of, part_of, etc	A leader is a kind of insurgent. Mullah Dadulla Lang was a Taliban commander. Kulahn Pass is a border crossing
Properties	Leader: Years' experience, training, sect, tribe, etc.
Values (for properties)	25 years, Mujhadin 1975-1982, Tajik, etc.
Rules (axioms) for reasoning	IF Mujhadin AND Tajik THEN accepted by Tajik population.

Level 2 Network Refinement in Non-Physical Nets

Outline:

- Level 2-3 Fusion in Non-Physical Nets
- Situation and Impact refinement applied to Organizations
- Basic Human Network Analysis (FM 3-24 approach)
- Social Network Analysis (SNA) to support Level 2 refinement
- Group detection and network assembly methods
- Visualization and analysis of organizational networks

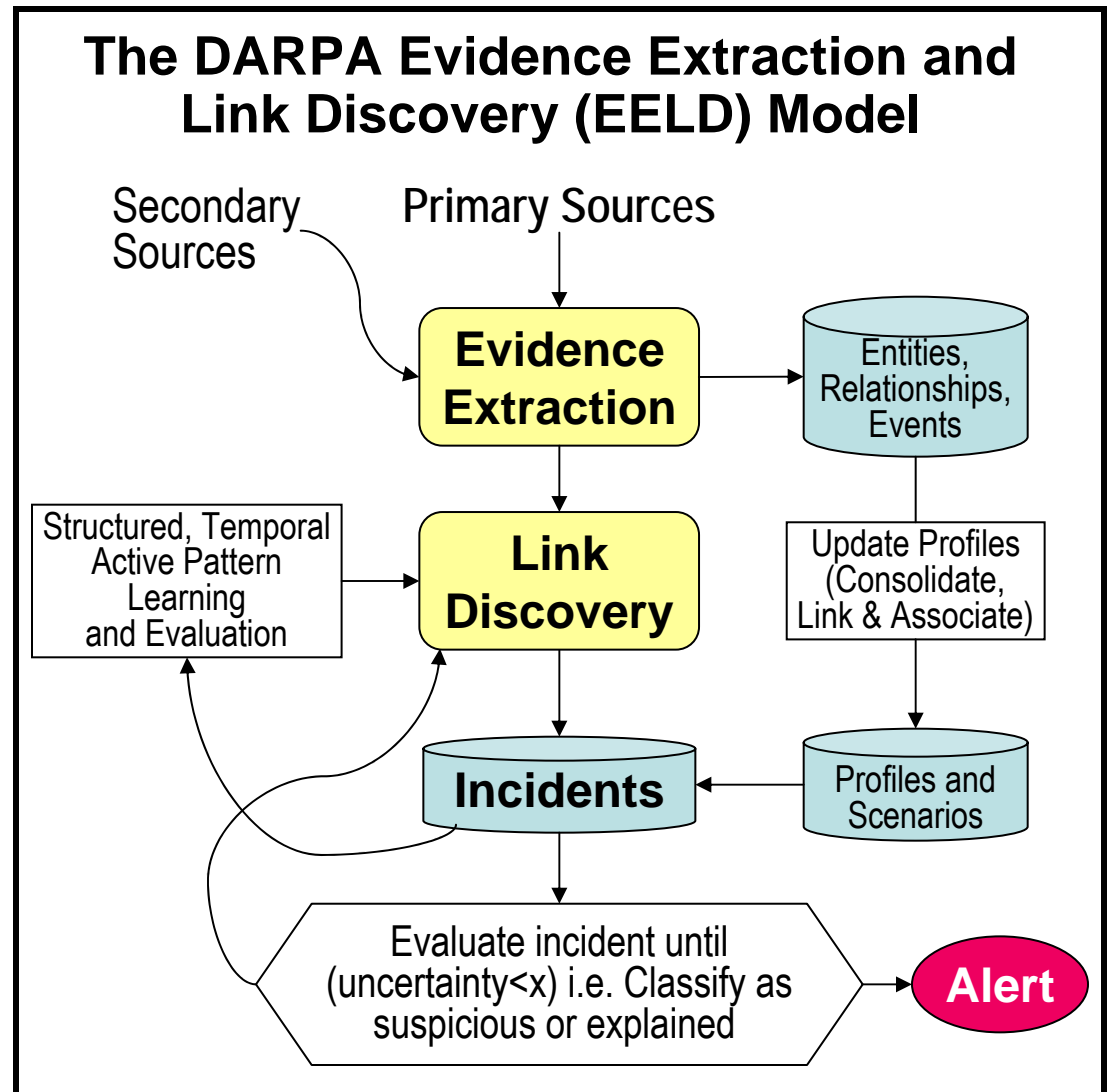
Situation and Impact Refinement Applied to Organizations



- Synthesize a model of the network
 - Accumulate and enumerate the entities
 - Identify the links
- Discover the structure of the network
- Quantify the metrics that characterize the network

Automated Link Mining and Discovery

- Approach: Project a relational data set onto a graph representation
- Apply graph theoretic methods to assess structural properties of the graph
 - Centrality – number of vertices between actors
 - Prestige – number of edges that “contribute to” an edge (directed graph)
 - Reach – number of vertices that a given vertex has edges to; number of edges to get to the target vertex
 - Sub groups – sub-networks that have structural equivalence (e.g. departments within an organization)



Link Mining Tasks

Link Mining Category	Task	Description-Application
Entity-Related Mining	Link-based entity classification	Label the entities within graph from a finite set of categories, based on links to each entity
	Entity clustering (Group Detection)	Cluster the entities into sub-graphs that share similar properties (discover the presence of groups)
	Entity Identification (Entity resolution)	Identify and label the complete set of unique entities within a graph
	Link-based entity ranking	Google Pagerank™ algorithm, HITS algorithm
Link-Related Mining	Link Prediction	Infer the existence of links between entities, based on the entity attributes (e.g. suspect a link where evidence does not yet exist)
Graph (network)-Related Mining	Subgroup Discovery	Discover “interesting” sub-graphs based on general criteria, or commonly occurring sub-graphs within a graph (e.g. discover patterns of terrorist cells with large graphs)
	Graph Classification	Classify an entire graph as a positive or negative instance (degree of correlation) of an example graph type.

Social Network Analysis (SNA)

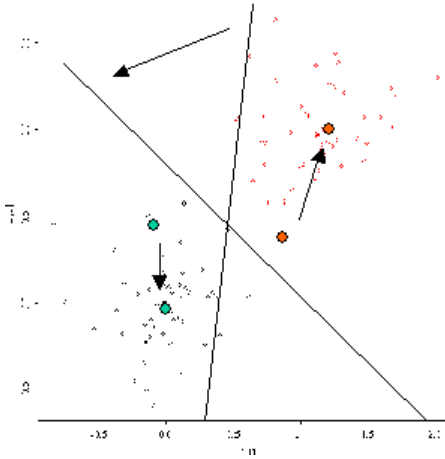
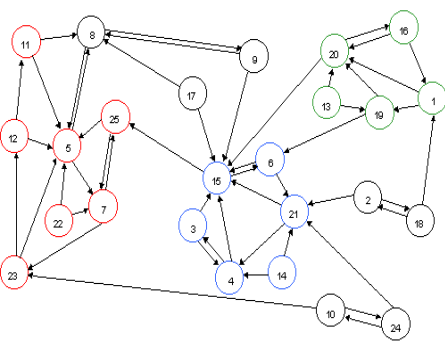
- **Social network analysis (SNA)** – Formal analytic method that quantifies the static, topological properties of the network as a whole (rather than any individual actor)
- **Premise of SNA** – Relationships are more important than individual actors; behavior of the network can best be described by the entire collection of interacting actors
- **Actors** – Interdependent representations of *entities*; in graph theory, an actor is a *vertex* of the graph; or *nodes* of the network.
- **Relationships** or *ties* describe the type and character of the relative roles, interactions, associations, transactions or other interdependency between actors. In graph theory, a relationship is an *edge* of the graph, of *links* between nodes of the network.

SNA Terminology

- *Degree* – The number of links to other actors in the network
- *Indegree* – The number of inbound links
- *Outdegree* – The number of outbound links in directional graphs
- *Clustering Coefficient* – The density of connections by quantifying the ratio of the total number of links connecting nearest neighbor actors, to the total possible number of links in the group.
- *Degree Distribution* – Distribution of probabilities that any actor in a group has k nearest neighbor actors.
- *Heterogeneity* – The degree to which actors' roles are similar to one another
- *Redundancy* – The degree of overlap in functionality such that the removal of a single actor or link does not prevent the organization from functioning
- *Latency* – The degree of potential delay in a network response to stimuli; ability to adapt to unforeseen events

SNA and Relational Data Mining

- Data Mining
 - Attribute based classification
 - Focus on attributes of entities – classify entities by similar properties (net without link consideration)
- SNA Relational Data Mining
 - Data is relational (emphasis on relationship links rather than attributes of entities)
 - Unit of analysis is the entire structure of social entities and links
- SNA software links
 - ORA <http://casos.cs.cmu.edu/projects/ora/software.html>
 - UCINET www.analytictech.com/ucinet_5_description.htm
 - Agna www.geocities.com/imbenta/agna/
 - NetMiner www.netminer.com/NetMiner/home_01.js
 - SociometryPro www.sociometry.ru/eng/index.php
 - Pajek <http://vlado.fmf.uni-lj.si/pub/networks/pajek/>
 - International Network for Social Network Analysis www.insna.org/INSNA/soft_inf.html

Data Mining	SNA Relational Data Mining
<ul style="list-style-type: none"> • Attribute-based • Classification of entities by attributes in a multidimensional space 	<ul style="list-style-type: none"> • Relationship-based • Set theory and multidimensional scaling to establish the strength and distance of connections
Unit of analysis: Entity (individuals)	Unit of analysis: Network Structure
<p>Attribute classification</p>  <p>Cluster Metrics</p>	<p>Sociogram</p>  <p>SNA Metrics</p>

Typical SNA Actor Index Metrics

Graph Type	Index Category	Actor Indices and Descriptions	Example Applications in Influence Analysis
Non-Directed Graphs	Centrality Measures	Degree Centrality (C_D) – Measure of the number of links an actor has to other actor.	Actors with High C_D and/or C_C have access to many other actors and may be candidates for recruitment or exploitation
		Closeness Centrality (C_C) – An aggregate distance measure (e.g. number of links) of how close the actor is to all other actors within the entire network; inverse of the sum of the shortest distances between each actor and every other actor in the network	
		Betweenness Centrality (C_B) – Measure of the extent to which an actor is located between other actors; the extent to which the actor is directly connected only to the other actors that are <i>not directly connected</i> to each other	Actors with high C_B may be critical liaisons or bridges; effective targets to reduce flows of information across the net.
Directed Graph (Digraphs)	Centrality Measures	Centrality Degree (C'_D) – Measure of the outward degree; the number of actors who are recipients of other actor outflows	Actors with high C'_C and C'_C may be critical decision makers; those who issue commands and policies.
		Centrality Closeness (C'_C) – An aggregate distance measure (e.g. number of <i>directed</i> inflow or outflow links) of how close the actor is to all other actors within the entire network	
	Prestige Measures	Degree Prestige (P_D) – Measure of number of links that transmit resources to the actor; a measure of how many actors provide to the prestigious actor	Actors with high prestige may be the focus of critical reporting, finances, or other resources; they may be key reservoirs or decision-makers.
Proximity Prestige (P'_D) – Measure of the number of actors adjacent to or from an actor			

Applied COIN Network Analysis (FM 3-24)

INDIVIDUAL-LEVEL ANALYSIS

- Characterize every member of the organization and identify key members, their location in the network.
- Relationship between the centralities of all nodes reveals:
 - Central Nodes dominate a very centralized network; removed or damaged, the network may quickly fragment into unconnected sub-nets.
 - Hubs are nodes with a high degree of centrality. A network centralized around a well connected hub can fail abruptly if that hub is disabled or removed.
 - Less Centralized Nets have no single points of failure; they are resilient to attack or random failure – many nodes or links can fail while allowing the remaining nodes to still reach each other over other, redundant network paths.
 - Degree Centrality describes how active an individual is in the network; degrees – the number of direct connections a node has. Nodes with the most direct links are the most active.
 - Between – A node with high “betweenness” has great influence over what flows in the network; betweenness centrality indicates whether an individual lies between other individuals in the net (intermediary, liaison, broker, bridge). A high betweenness centrality person can be a single point of failure.
 - Periphery nodes – Receive very low centrality scores; they can often be connected to nets that are not currently mapped; they may also be resource gatherers or actors that have their own network outside of their insurgent group.

ORGANIZATION LEVEL ANALYSIS

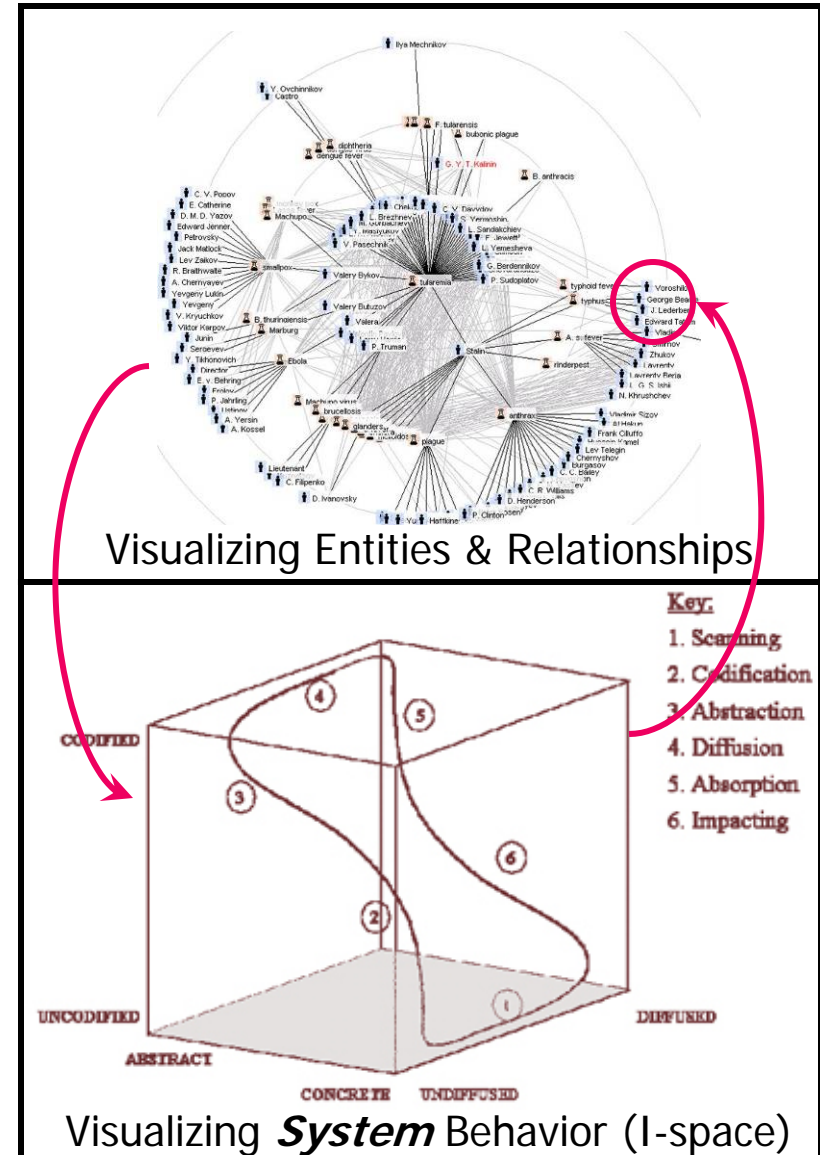
- Insight about the organization’s form, efficiency, and cohesion.
- A regional insurgency may consist of large numbers of disconnected sub-insurgencies. As a result, each group should be analyzed based on its capacities as compared to the other groups.
- Organizational-level capacities described by the social network measures:
 - Network density – General indicator of how connected people are in the network; global-level density is the proportion of ties in a network relative to the total number possible.
 - Cohesion, Efficiency, Core-periphery actors
- Using someone else-s data to come to conclusions that determine military operations is inadvisable; commander should take ownership of the collection process, decide which people are tracked as nodes in the data, and what constitutes a relationship between nodes.
- How a commander collects and organizes data changes which social network measures are most relevant to the situation.

Approaches to Model Iranian Nation-State Decision Making

	MODEL	Nation-State Behavior	Alternative Sub-Models
First Generation Graham Allison, Essence of Decision	I (Classical) Rational Choice Model	Basis: Nation-state as <i>ACTOR</i> (or team of actors): <ul style="list-style-type: none"> • Rational actor: Decision-making is a <u>choice</u> by optimal selection from among a set of alternatives • Basis of decision-making is overall utility 	<ul style="list-style-type: none"> • <i>Unitary</i> – single rational actor • <i>Factional</i> – Internal factions in cooperative-competitive decision-making • <i>Strategic</i> – Triangular power game among international players
	II Institutional Model	Basis: Nation-state as <i>MACHINE</i> <ul style="list-style-type: none"> • Government machine of process model: Decision-making is an <u>output</u> or <u>result</u> of integration, or bargaining/compromises among institutions • Basis of decision-making is integrated process result. 	• <i>Organizational Process</i> – Result of coordinated, but independent institutions
	III Bureaucratic Government Politics		• <i>Bureaucratic Politics</i> – Result of bargaining and compromises among competitive institutions of government
2nd Gen Janis, ...	Interacting Multi-Actor Models	Basis: Nation-state as <i>ORGANISM</i> <ul style="list-style-type: none"> • Mixed-actor models: Decision-making is a group process • Basis of decision-making is emergent group behavior, dynamics, psychology 	• Many complexity-based multi-actor model structures

Representing the *Dynamics* of Networks

- Current / traditional analysis in Physical Domains “count molecules”, not the pressure in the bottle
 - Thus we don’t know if *conditions* are going to change
 - Nor how to effect a change in conditions
- How to represent and view the dynamics of non-physical networks?
 - A holistic system of systems
 - Understand behavior and drill down from behavior to critical entities
 - In the information and cognitive domains
- This is the most abstract of all knowledge.



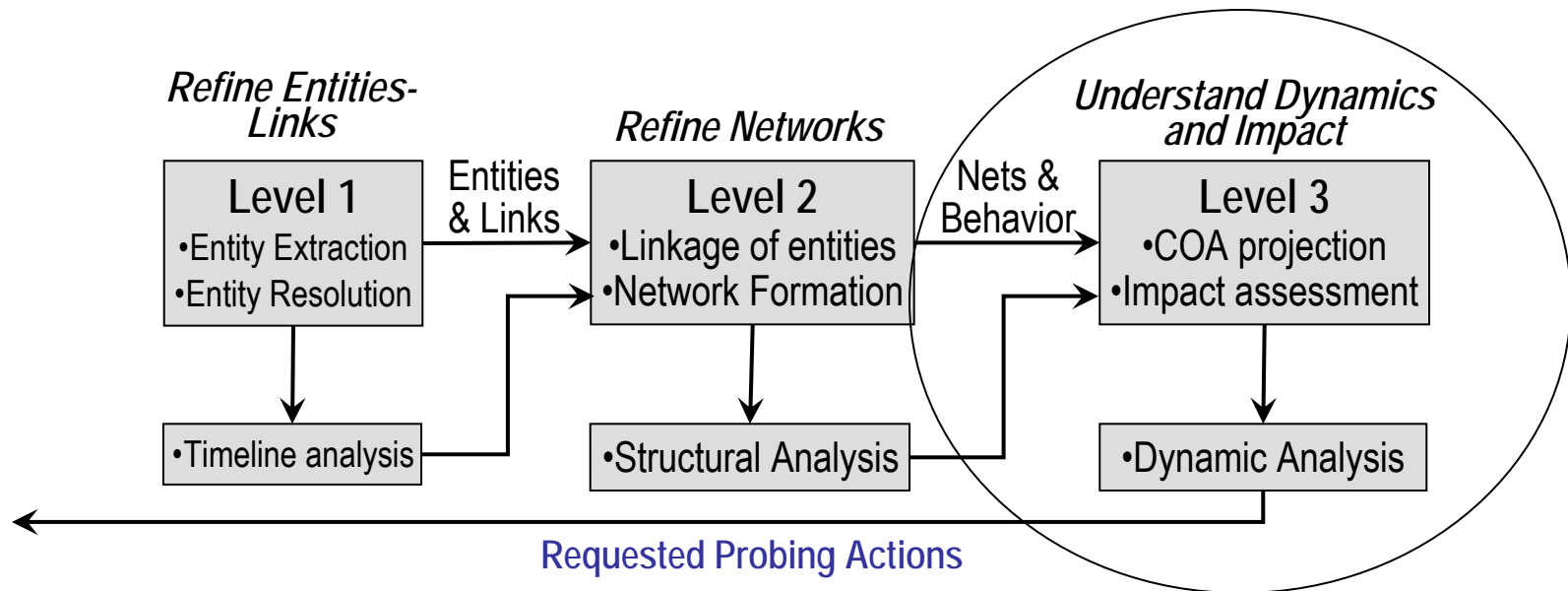
Source: Boisot, M. *Knowledge Assets: Securing Competitive Advantage in the Information Economy*.

Level 3 Impact Refinement: Anticipating the Dynamics of Networks

Outline:

- Tracking the Dynamics of Non-Physical Nets
- The complexity of Social Networks – Computational Social Science
- Tracking complex Social Dynamics – Elements of State, and what about Kalman filters?
- Exploratory Dynamics Analysis and Tracking
- A survey of Dynamic Modeling Methods and their applications

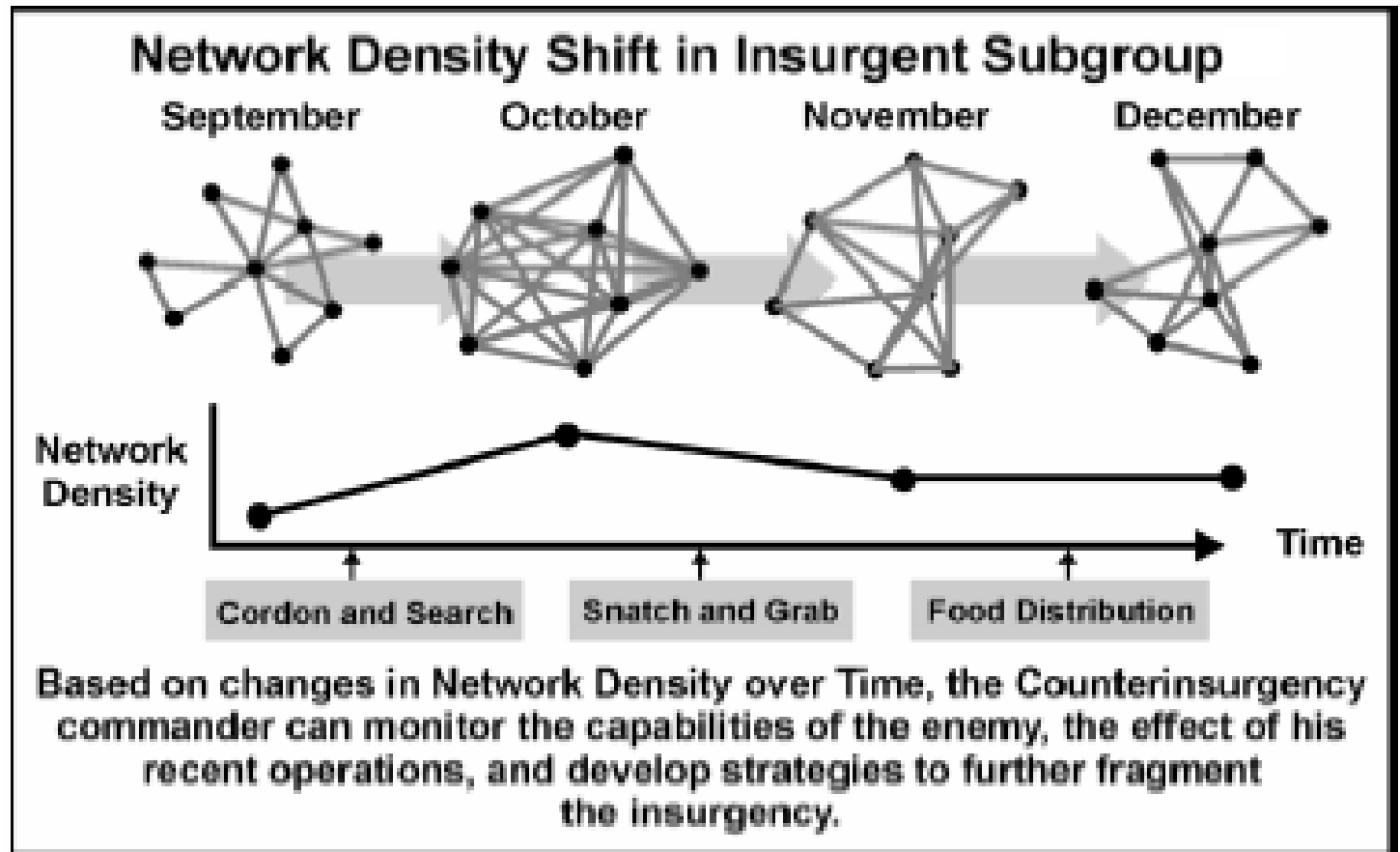
Situation and Impact Refinement Applied to Organizations



- Analyze the dynamic behavior of the entire network
 - Stability, equilibrium, growth, vulnerabilities
- Project possible course of action (COAs for the network)
- Assess the sensitivities of the net to external influences
- Explore the effects of alternative policies (actions)

Operational Need: Tracking Network Dynamics

- Network Dynamics
 - Entities
 - Structure
- Effects of Operations
- Emergence of Threats



Source: FM 3-24 Counterinsurgency, Appendix E, Social Network Analysis

Social Network Characteristics

- **Small World Phenomena**
 - Stanley Milgram experiment (1967)
 - Average path length ~6 (“six degrees of separation”)
- **The Reason: Funneling Concept**
 - Connectors (hubs) in real world networks make them uniquely different from random networks; Malcolm Gladwell Tipping Point (2000)
- **Resulting in Scale Free Networks**
 - Structure and dynamics independent of the system’s scale (N, # of nodes)
 - Degree distribution follows a power law:
P(k) is proportional: $P(k) \sim k^{-\gamma}$
 - $P(k)$ = probability that a node connects with k other nodes
 - Coefficient γ is generally between 2 to 3
- **These networks exhibit the Tipping Point Dynamic (Phase Shifts)**
 - Morton Grodzins, Gladwell
 - Rare phenomenon becoming rapidly & dramatically more common (epidemic)

Characteristics of Scale Free Nets

1. Hubs (with degree well above the average) distributed throughout
2. Hierarchy of hubs – large, medium, small provide funneling and low degree of average separation
3. Tolerant to random link failures
4. Vulnerable to hub failures; relatively small percentage of hub failures can bring systemic collapse
5. Can be subject to rapid collapse (tipping point) when critical number of hubs have concurrently failed.

Cautionary Notes

- These structural properties presume a static network;
- Self-healing networks (e.g. small hubs become large hubs) can keep the network scale-free in the presence of link failures (small degree of separation, and responsive)

Causality and Complexity

Causality – The relating of causes to the effects they produce.

- The belief in and description of causation
- Related to determinism – The naturalist teaching that every event in the universe is caused and controlled by natural law; that there is no free will in humans and that all events are merely the result of natural and physical laws.
- Believing and knowing (*two different things!*)

Complexity – Measured by the scale of its independent actors and the scale & effects of Interactions between those actors.

Characteristics of Complex Adaptive Systems

- **Unstable** – Subject to large effects being produced by small inputs or changes;
- **Appear random** (*they are not*) – Chaotic and relatively simple deterministic process rules that produce *apparently* random behavior
- **Irreducible** into their independent components and relationships between components (defy decomposition)
- **Hyper-sensitive** to initial conditions; the same elements together as a system may exhibit drastically different behavior depending on starting conditions
- **Paradoxical** – Can have multiple independent solutions or explanations for behavior
 - Root cause(s) rarely discernible
- **Emergent Behavior** arises from interaction of all the independent actors in the system
 - Unpredictable by analyzing system's component parts
 - Driven by *interaction* of its parts

Source: Casti, John L. *Complexification: Explaining a Paradoxical World Through the Science of Surprise*, NY; Perennial, 1995

Emergence in Complex Systems

- **Emergence** – The development of novel and coherent structures, patterns and properties that are not predictable from the individual elements of the network; the result of self-organization
 - **Weak Emergence** – Behavioral properties arising in systems as a result of the interactions at an elemental level
 - **Strong Emergence** – Behavioral properties not directly traceable to the system's components, but rather to how those components interact
- Emergence as a property of “Chaos” (Complex) Systems –
 - Not social disorder, anarchy or general confusion –
 - A class of dynamical behavior that occurs in deterministic nonlinear systems
 - Occurs in feedback systems in which yesterday's events impact today's events and today's events will affect tomorrow's.
 - A semi-orderly system in a state, which despite an apparent disorderly appearance, has some predictability and order in its structure and process.
- Although Chaos systems may appear random in their behavior, they gravitate toward phase spaces of stability over time:
 - *Attractors* are phase spaces that can be predicted but only over the short term, and only if the system is well understood enough.
 - *Bifurcation* is an abrupt change in the system's state due to a changed parameter(s)
 - When a system bifurcates, it may do so into disorder, or simply move to another settled state.

Network Defeat Strategies

- Waltz's 3 domains of defeat [1] (*right*)→
- Glenn's 4 categories of system targeting modes: [2]
 1. Decrease variety
 2. Decrease interaction
 3. Decrease energy (entropy increasing)
 4. Alter Feedback control

[1]Waltz, Ed "Practical Approaches to Impact Adversary Decision-Making Processes" in "Organization Decision Making and Information Warfare", Artech 2007

[2]Glenn, Kevin B. (Maj, USAF), *Complex Targeting: A Complexity-Based Theory of Targeting and its Application to Radical Islamic Terrorism*, Maxwell AFB:Air University, 2002; Ch.4

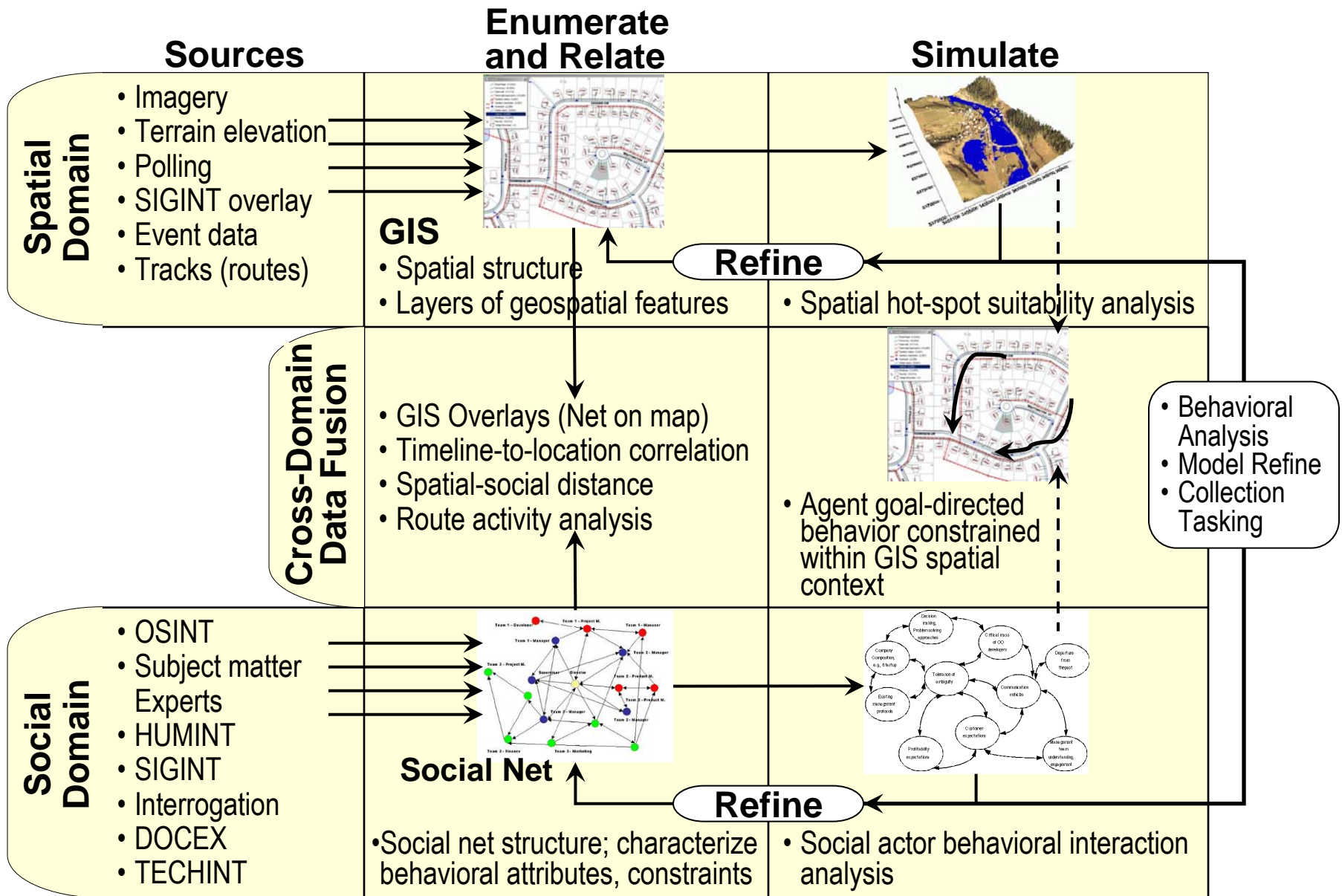
Domain of Attack	Defeat Mechanisms	Typical Military Effects Sought
Physical Defeat	Attrition of adversary's physical resources that are sources of power; Maneuver of attacker's physical resources to exceed the rate of target response	<ul style="list-style-type: none"> • Annihilation of physical resources; inability to track and respond to attacker's physical assets
Information Defeat	Attrition of adversary's information resources that are sources of power; Maneuver of attacker's information-based attacks to exceed the target's ability to recover and restore networks	<ul style="list-style-type: none"> • Degradation and loss of data sources, communication links, and ability to process information • Degradation of networks to reduce information-sharing, shared situation awareness and ability to self-synchronize
Cognitive - Emotive Defeat	Attrition or Exhaustion of the mind and will Dislocation of state of mind of leadership Disintegration of the state of mind of combatants and their cohesion	<ul style="list-style-type: none"> • Uncertainty, distracted, misdirected, and deceived state of mind • Loss of situation awareness • Indecision, Delayed and erroneous decisions • Loss of morale and will to fight

Current Challenges and Key Technologies

Representative Operational Needs

DoD	Intelligence Community
<ul style="list-style-type: none"> • Army <ul style="list-style-type: none"> – Human Terrain (ARL, INSCOM IDC) • Air Force <ul style="list-style-type: none"> – Behavioral Analysis (NASIC; AFRL HE) • Navy-Marine Corps <ul style="list-style-type: none"> – Cultural Intelligence (Quantico) • JFCOM <ul style="list-style-type: none"> – J9 Effects based approach to planning • JWAC <ul style="list-style-type: none"> – Joint Warfighting Analysis Center • JIOWC <ul style="list-style-type: none"> – IO Intelligence preparation of the environment (IO IPE) – Precision influence target 	<ul style="list-style-type: none"> • NSA <ul style="list-style-type: none"> – Network analysis (extraction and link from structured/unstructured) • NGA <ul style="list-style-type: none"> – Geospatial-social data fusion • CIA <ul style="list-style-type: none"> – Unstructured data analysis • DIA <ul style="list-style-type: none"> – Human Factors Analysis Center (HFAC) • IARPA-DTO <ul style="list-style-type: none"> – CASE, ARIVA, PAINT

Geospatial-Social Terrain Data Fusion



Socio-Cultural Modeling R&D Programs

Research

- AFOSR – Socio-Cultural modeling
- AFRL – HEC; IFR
- DTO – PAINT
- NRL – Willis, Goolsby
- NPGS – MOVES Institute
- DHS – Center for Behavioral and Social Aspects of Terrorism
- DARPA – PCAS
- NDU – PCMT
- AFRL – CPE Vision

Applications Development

- Complex Operations Modeling, Planning and Outcome Exploration COMPOEX (DARPA-JFCOM)
- SEAS (JFCOM)

Operations

- DCGS-A
- CHAMPION
CI-HUMINT Advanced Modernization Program / Intelligence Operations Now JCTD
- Human Terrain System (TRADOC/FMSO)
- Marine Corps Intelligence Activity Cultural Awareness program

Network Attrition and Attraction

Two complementary and interdependent approaches

	Attrition (Annihilation)	Attraction and Exhaustion
Function	Locating and capturing individuals hidden within the environment	Transforming the social environment in which the adversary is embedded; secure support, apply pressure, affect and degrade operations
Basis	Attrition based warfare Open military confrontation Local scale (military will to fight)	Information and 4G warfare Political-social influence Global scale (public will of opinion)
Military Role	<ul style="list-style-type: none"> • Broad area search; Cordon search • Surveil, Tentative Detect, Identify • Tag, Track and Locate (TTL) • Entice, Trap, Capture, Kill • Exploit, forensics, analysis 	<ul style="list-style-type: none"> • Cultural Intelligence • Social Surveillance, • Social Network Analysis • Broad area PSYOP; focused PSYOP • Process Tips, Cues, Suspects
Focus	<ul style="list-style-type: none"> • Point targets – direct effects • Military action 	<ul style="list-style-type: none"> • Population targets – indirect effects • Strategic influence