THE CENTER FOR ADVANCED STUDIES
IN SCIENCE AND TECHNOLOGY POLICY

Information Management and Policy In the Intelligence Enterprise

K. A. Taipale
Executive Director, Center for Advanced Studies
Senior Fellow, World Policy Institute
Adjunct Professor of Law, NYLS

Presented at:

McLean, VA • April 12, 2005

Presentation Overview

- Intelligence is a *knowledge creation* (not information sharing) enterprise
- *Converging missions* create *interstitial gaps* in policy and law
- Technology-enabled *networked organizational structures* challenge existing management processes and policies that are premised on earlier paradigms (hierarchies, industrial production, etc.)
- Requires *new intelligence production model* (KMIC enterprise model) serving dual obligations of *security & privacy* (business process needs)
  - production of actionable intelligence (~ security), and
  - protection of civil liberties (~ privacy)
  - *same technology and infrastructure* required to support both
- Policy Appliance Reference Model (both metaphor and framework for social construction of technology-enabled ISE)
- Discussion

The Center for Advanced Studies

- *Private*, *independent*, *non-partisan* research and advisory organization focused on *information*, *technology*, and *national security* policy
- Program on Law Enforcement and National Security in the Information Age (www.PLENSIA.org)
- Project relationships with World Policy Institute, Yale Law ISP, NYU Law CLS, Columbia CITI, NYLS, Markle Task Force, and others
- More info and papers: http://www.advancedstudies.org/

Related publications

- *Designing Technical Systems to Support Policy: Enterprise Architecture, Policy Appliances, and Civil Liberties*, in 21st Century Information Technologies and Enabling Policies for Counter-Terrorism Robert Popp and John Yen, *eds*. (IEEE Press, forthcoming 2005)

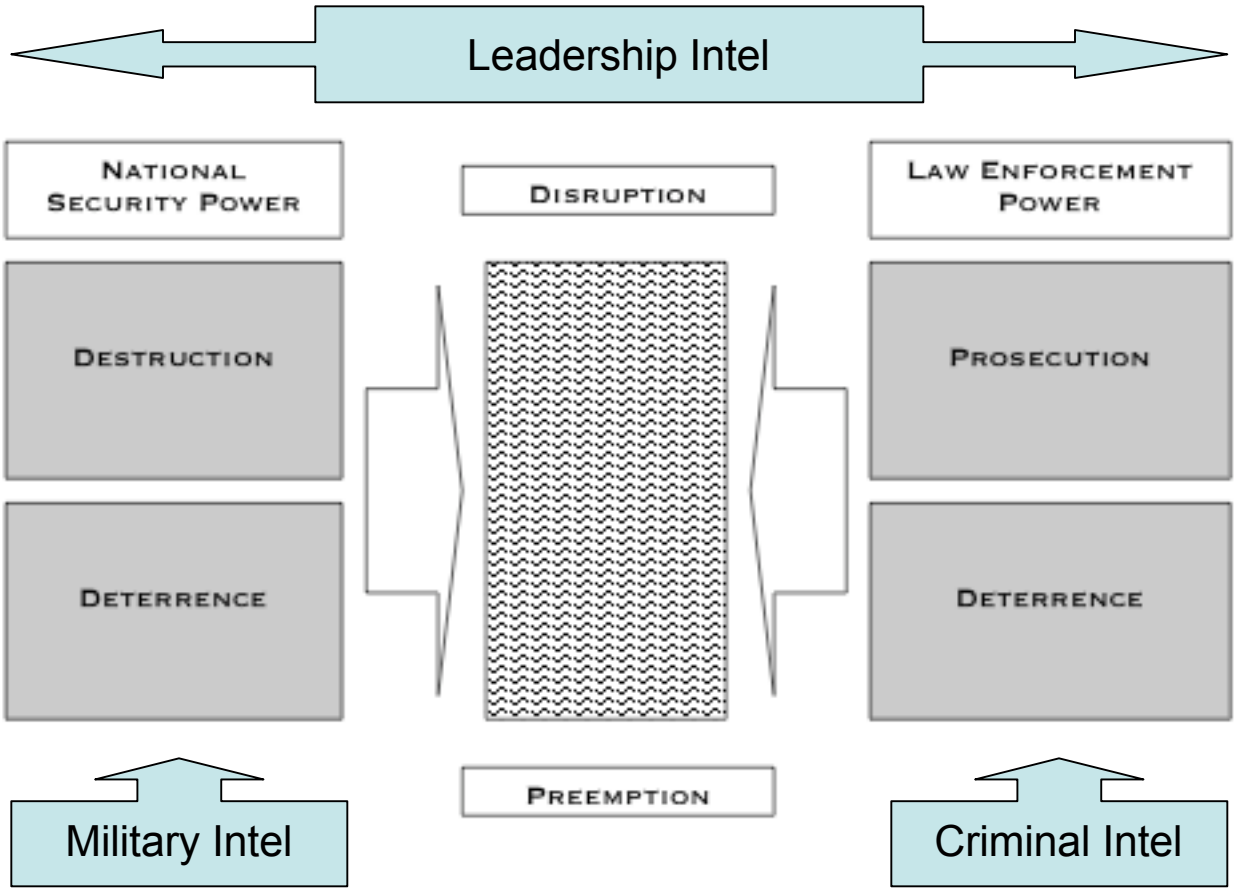
- *Technology, Security and Privacy: The Fear of Frankenstein, the Mythology of Privacy, and the Lessons of King Ludd,* 7 Yale J. L. & Tech. 123 (Dec. 2004)

- *Data Mining and Domestic Security: Connecting the Dots to Make Sense of Data,* 5 Columbia Sci. & Tech. L. Rev. 2 (Dec. 2003)

More caveats and context

- GWOT* as defining *threat metaphor* (but other emerging threats, cross border gangs, organized crime, rogue corporations, hostile state proxy networks, etc.)
- Scalarity, epistomology, “km”, description/prescription (analysis)
- HSPD-6 (TTIC), HSIS §892, 9/11 Report §13.3, EO 13354 (CTC), EO 13356 (“need to share”), IRTPA §1016 (ISE), etc. (“*information sharing*”)
- FEA, CIO Council, IC ISC 2004, IC CIOs, ICMWG, etc. (*but cf.* ICSIS/ICON 1999, ICDC 2001, etc.) (IC “*information systems*”)
- 9/11 Commission Report, Intelligence Commission Report, etc. (“*intelligence failures*”)

The challenge: converging missions and policy gaps



Why now? "It's the network, stupid"

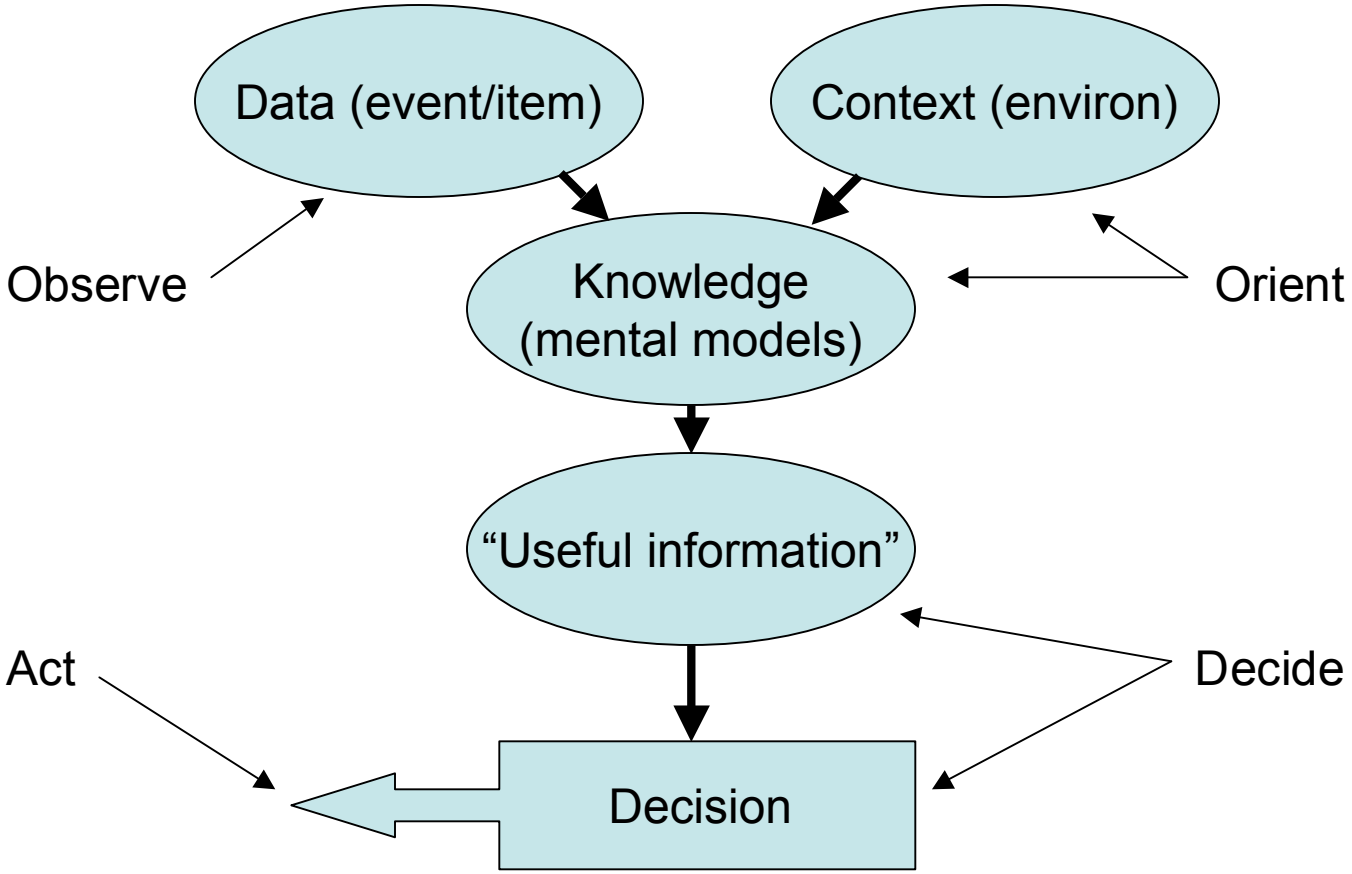
- ICT enables new management/social structures by facilitating asynchronous, non-bounded *convergence* among unrelated but like-minded individuals and groups (Novack & Johnson 2005)
- Clash between hierarchical power structures (nation-states) and emergent networked organizations (*seed value devolution*) GWOT*
- "*Networks and Netwars*" (Arquilla & Ronfeldt 2001 RAND)
  - *Networks have advantage over hierarchies*
    - Swarm resources
    - Adaptive/resilience
  - Power migrating to non-state actors (lower barriers to entry, lower transaction costs, and lower risk makes it "affordable") (Coase)
  - Not cyberwar (~military/HICs) but netwar (~LIC, OOTW, IC/LE)
  - "*Takes a network to defeat a network*" (coalitions, fusion centers)

Intelligence as Knowledge Enterprise

- Power of information is *contextual* (data + context = knowledge)
- Power derives from the *usefulness* (or uselessness) of knowledge for *decision-making*
- *Actionable intelligence* is contextually useful information or knowledge
- *Knowledge management* requires getting the right information to the right place at the right time for effective *strategic*, *tactical*, or *operational* decision-making (action)
- IC as *knowledge creation ~ learning community*

Distinguish km from KM, *e*-learning, knowledge sharing, project collaboration, business process management, CRM, e-commerce and publishing applications (not an integrated or holistic approach) and from enterprise architecture (focused on data and interoperability)

Data + context + knowledge = actionable information

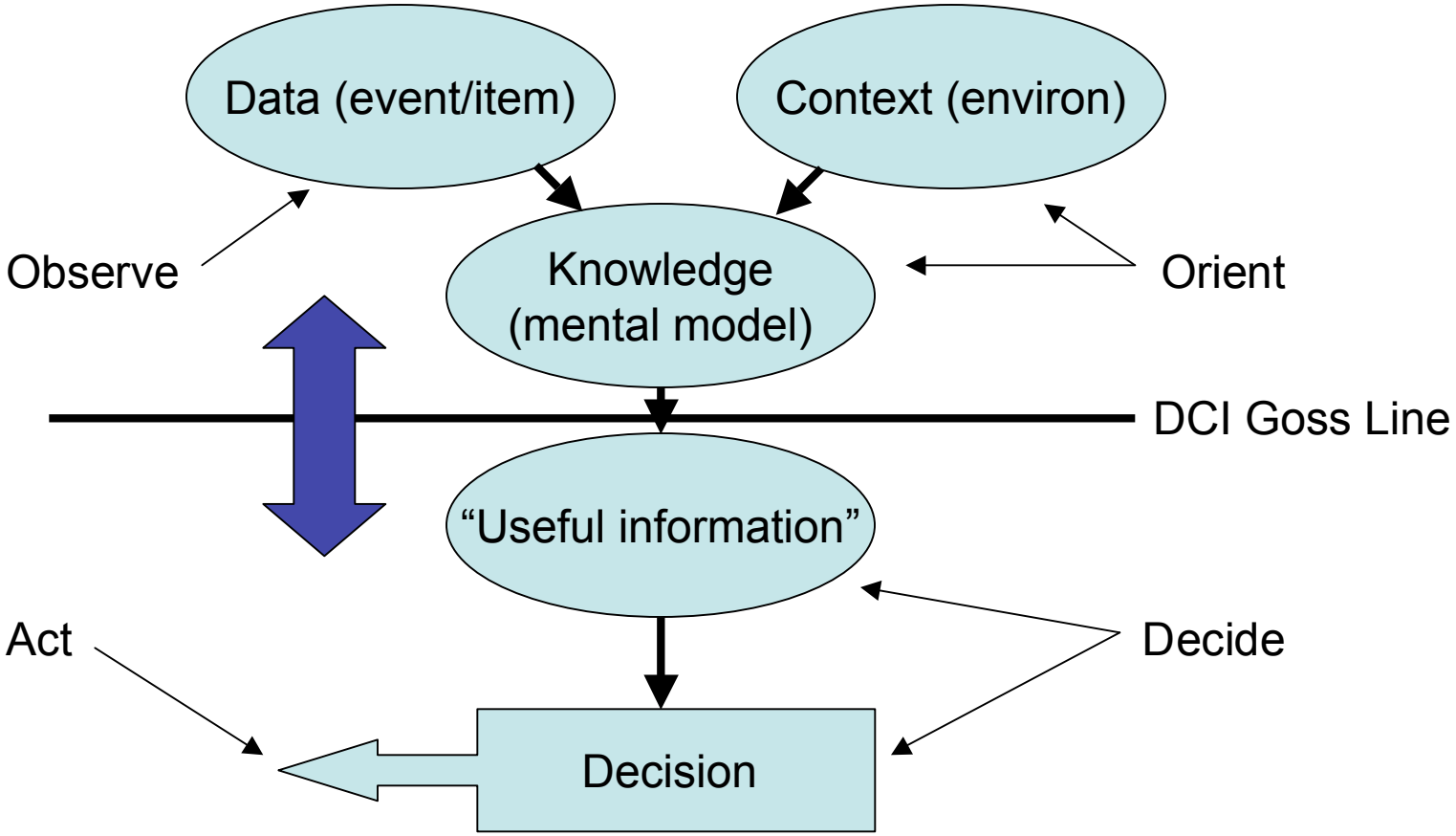


The intelligence "firm"

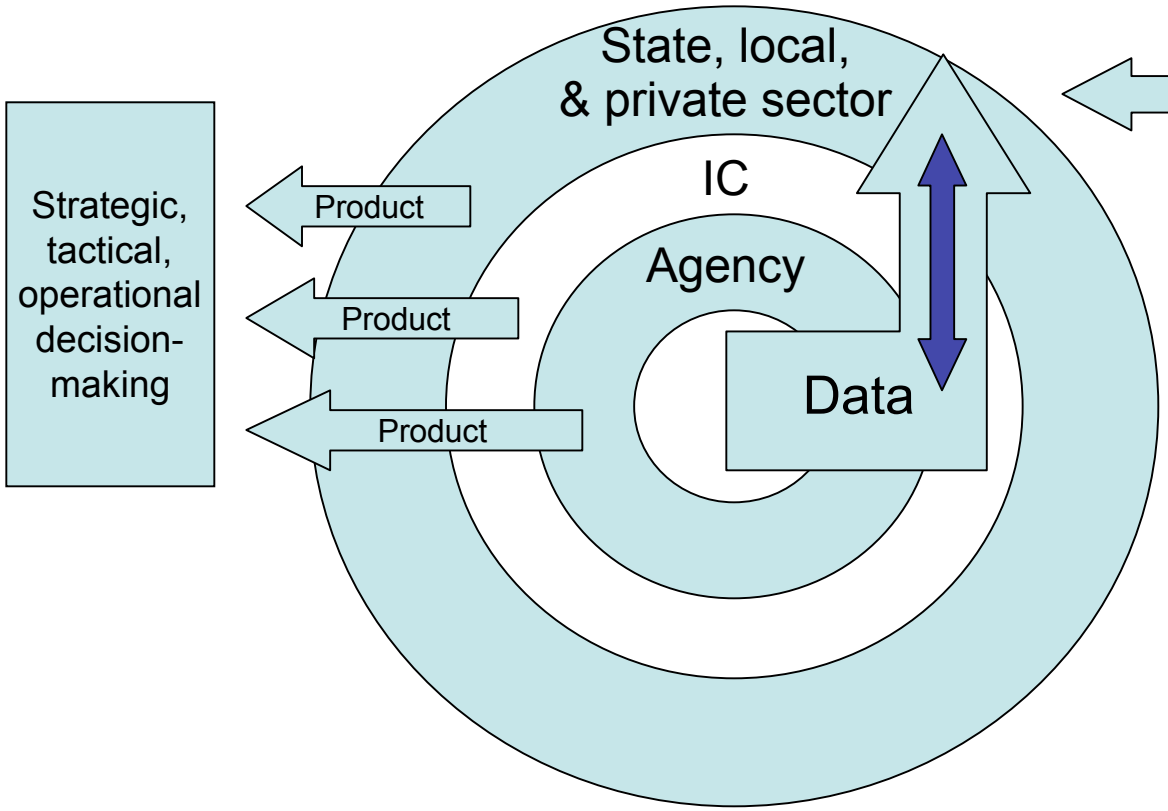
(specialization and lower transaction costs)

- Intelligence firm "value-added" functions:
  - Add value to information (collect, aggregate, summarize, abstract, model)
  - Develop knowledge and expertise (domain, process, methods, etc.)
  - Capture and share knowledge (transform)
  - *Facilitate application of knowledge* - service is only useful if it facilitates better decision-making (*cf. Intel Commission* conclusions re PDB - agenda setting rather than useful for decision-making)
- DCI Goss - CIA role to "collect and analyze" not to make policy
- *But cf*. "black box" advice (opaque) vs strategic advice (transparent)
  - E.g., consensus view vs. publishing debate/dissenting views
  - Situate information at the right level -- strategic, tactical, operational
- Value of information is its usefulness for decision-making; decision-making requires rich understanding of the mental models - *thinking behind analysis* and its applicability to local conditions (*local = up+down*)

Intelligence has little value if not useful for decision-making



Data has little value without use context



Make appropriate data available throughout the network so it can be contextualized.

And, activate network sensors.

Data may be meaningless to any particular “owner” and only becomes useful as people perceive it as useful within their local context or knowledge

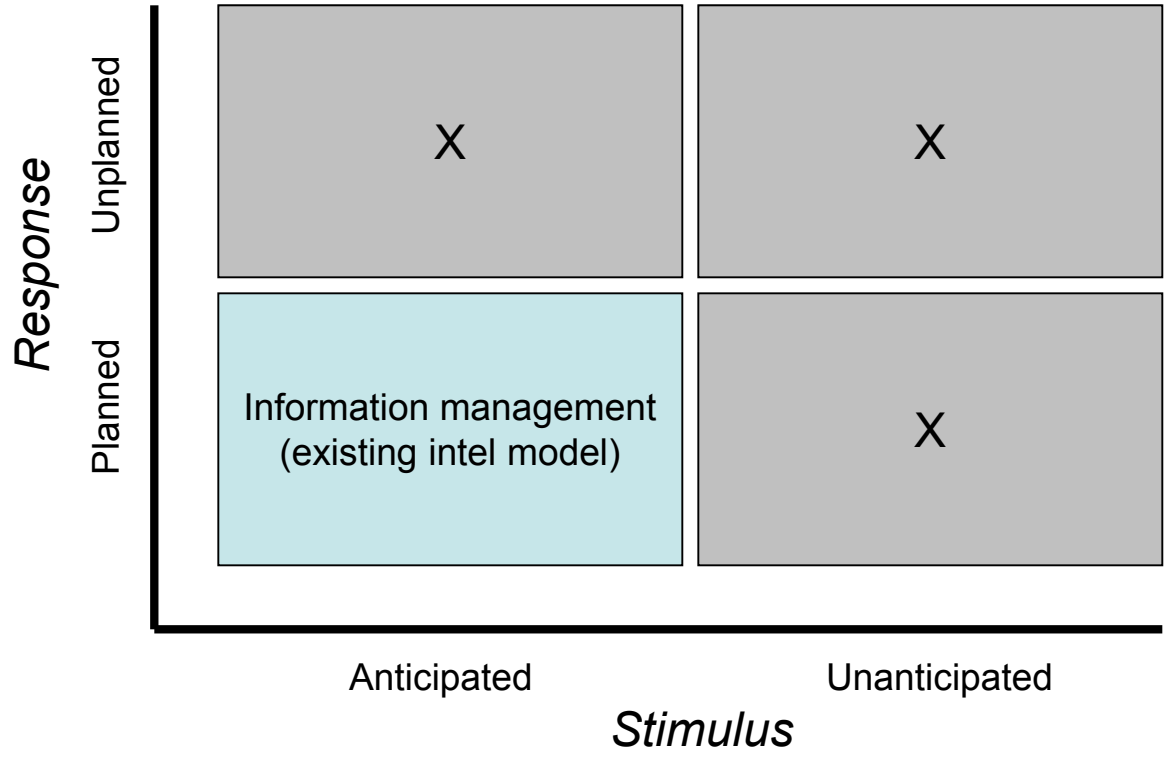
But data is only useful if the *capabilities* and *competencies* exist to produce knowledge

- Sharing data where it can't be processed results in *negative value* -- i.e., information overload
- *Capabilities* needed to process data include abilities to:
  - Summarize/condense/abstract (i.e., manage volume)
  - Contextualize (understand its relevance to local conditions)
  - Categorize (relate it to other information held locally)
  - Evaluate (process it within locally relevant mental models)
  - Verify (judge its reliability)
- *Competencies* required vary according to local decision-making requirements (strategic, tactical, operational, situational, etc.)
- Managing intelligence production requires evaluating capabilities and competencies, sharing appropriate information, and providing services/products to make up for processing deficits

Intelligence “failures”

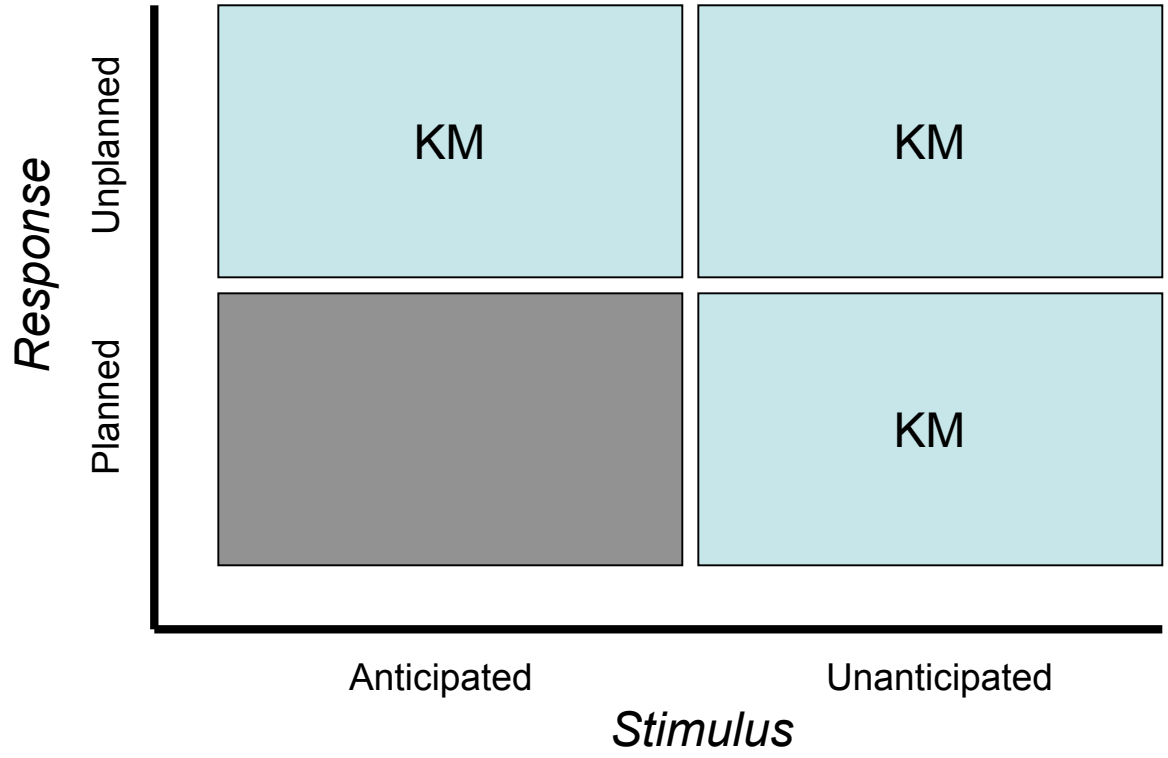
symptoms (stovepipes, turf) vs causes (production model)

Information management (need to know) paradigm:
good for preplanned responses to anticipated stimuli



Intelligence “futures”

Knowledge management (need to learn) paradigm:
to enable unplanned responses (innovation) to surprise stimuli



Knowledge management

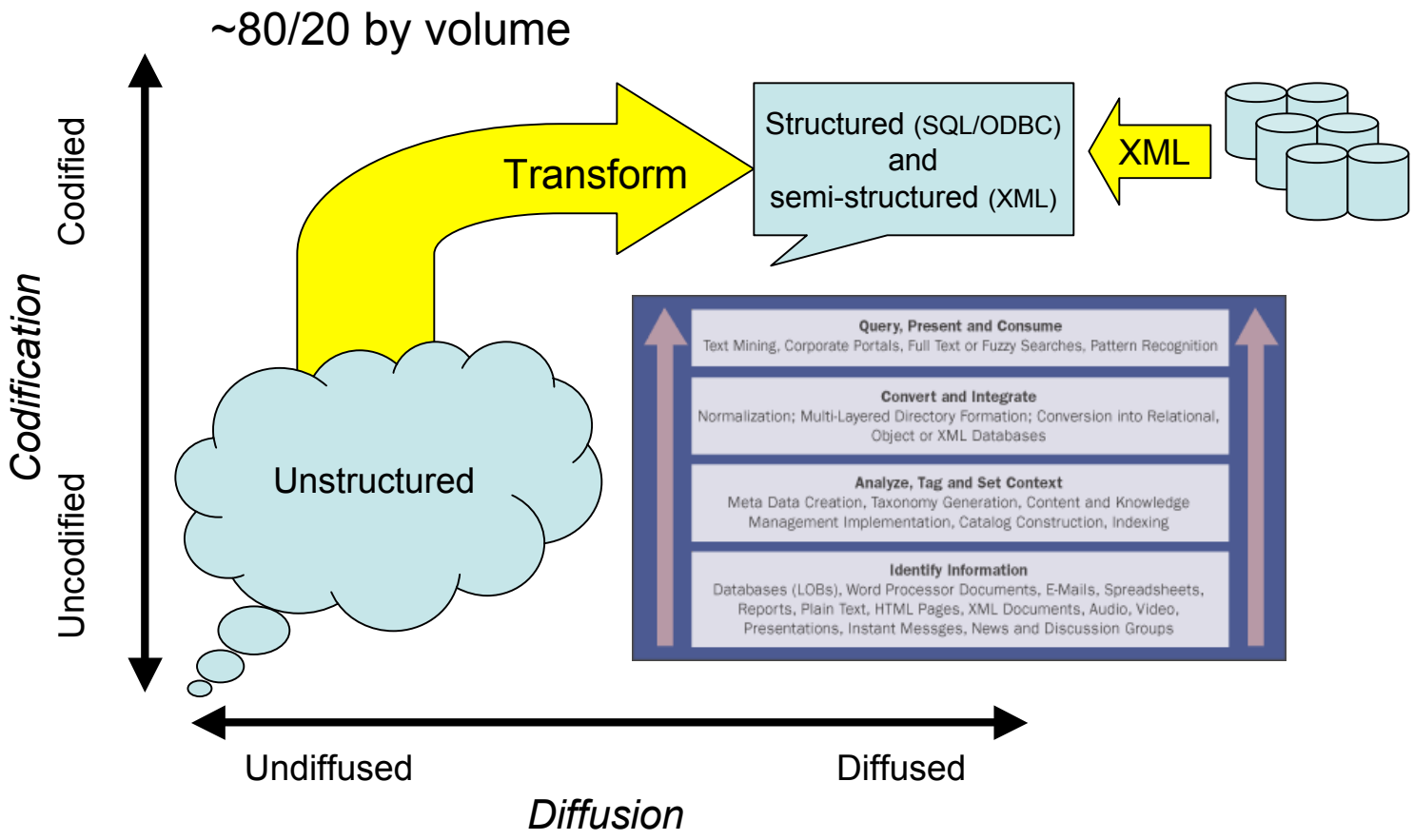
- KM is the leveraging of *collective wisdom* (experience + knowledge) to increase organizational responsiveness, innovation and resilience (ability to recognize and respond to novel threats, ~ InfoSec trends)
- KM supports *adaptive thinking* (think outside the box)
- Technology + reorganization + people = *social construction*
- Intelligence community ≈ *learning community*
- “Eyes and ears” vs *intelligence*

Components of a knowledge-centric system

- *Repositories* (hold explicated knowledge and associated rules)
  - Declaritive (concepts, categories, definitions, assumption)
  - Procedural (processes, sequence of events/activities, actions)
  - Casual (decision rationales, outcomes analyses, after-action)
  - Context (decision circumstances, assumptions, results of assumptions)
- *Collaborative platform* (support distributed work, info push/pull)
- *Network(s)* (data and communications connectivity)
- *Culture* (sharing and trusted community)

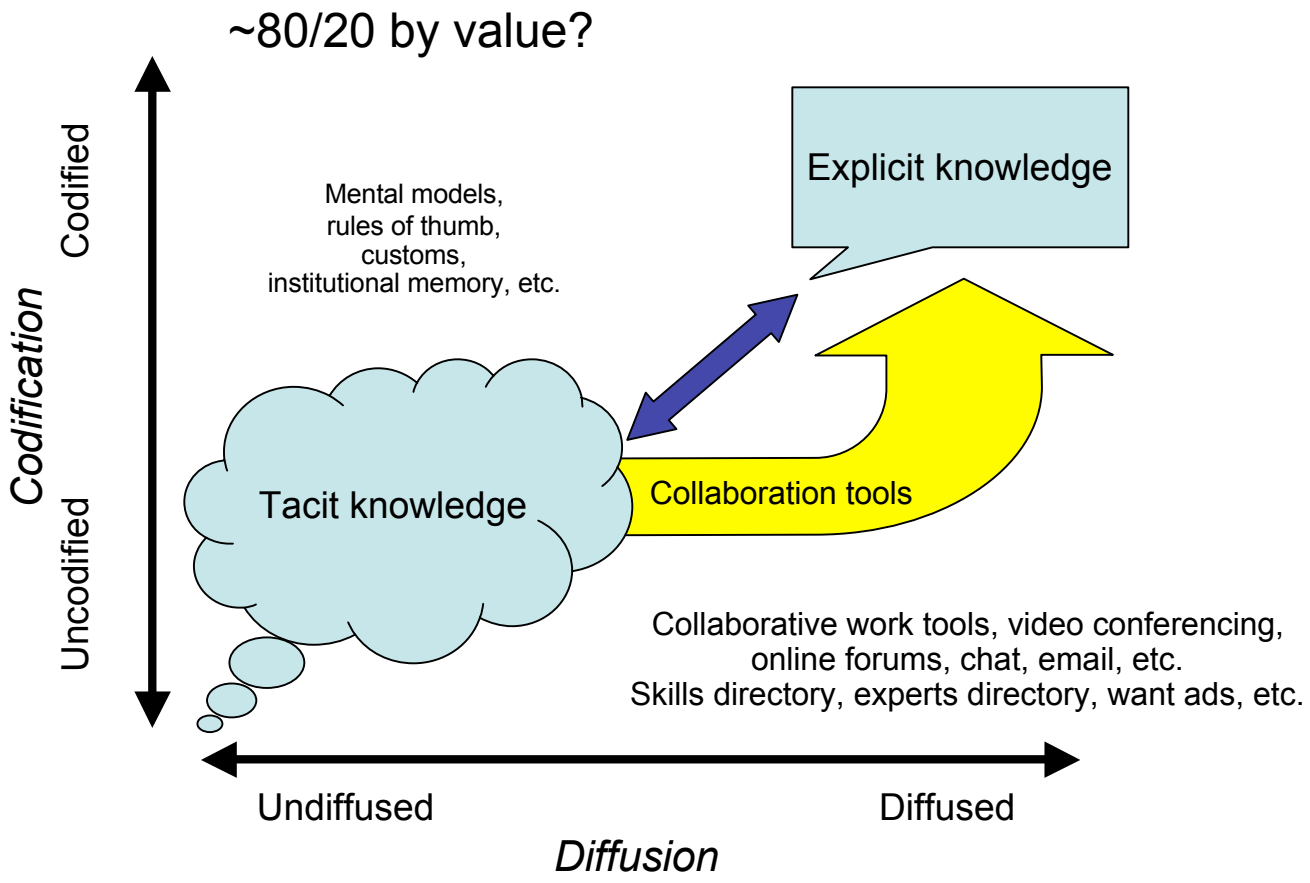
Information complexity - data-centric

(capture knowledge through structure)

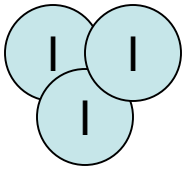
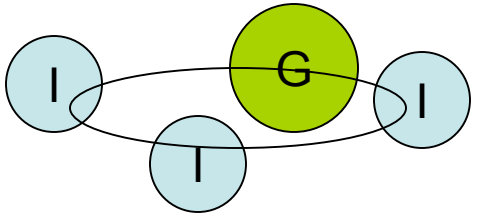
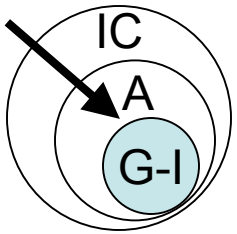
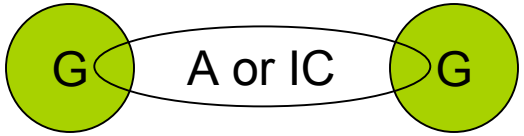


Knowledge complexity - user-centric

(capture knowledge through rich social interaction)



A taxonomy of knowledge development tools

Socialization (tacit-tacit) (F2F, Video, Email, chat, etc.) 	**Externalization (tacit-explicit)** (process capture, expert systems, etc.) 
Internalization (explicit-tacit) (knowledge networks and discovery, etc.) 	**Combination (explicit-explicit)** (collaborative work tools, intranets, forums, best practices database, etc.) 

Developing organizational information and knowledge capabilities - *technologies and skills*

	Individual Use (volume mgt)	Organizational flows (value mgt)
Technology	•Search engines •Filters and personalization •Intelligent agents •Analytic tools •Visualization tools, etc.	•Email, intranets & groupware •Skills/experts directory •Yellow pages/ •Electronic forums and chat •Videoconferencing, etc.
Skills and behaviors	•Time/attention management •Analysis and synthesis •Decision-making •Communication skills, etc.	•Organizational culture •"need to share" •Teamwork and collaboration •Fusion centers, etc.

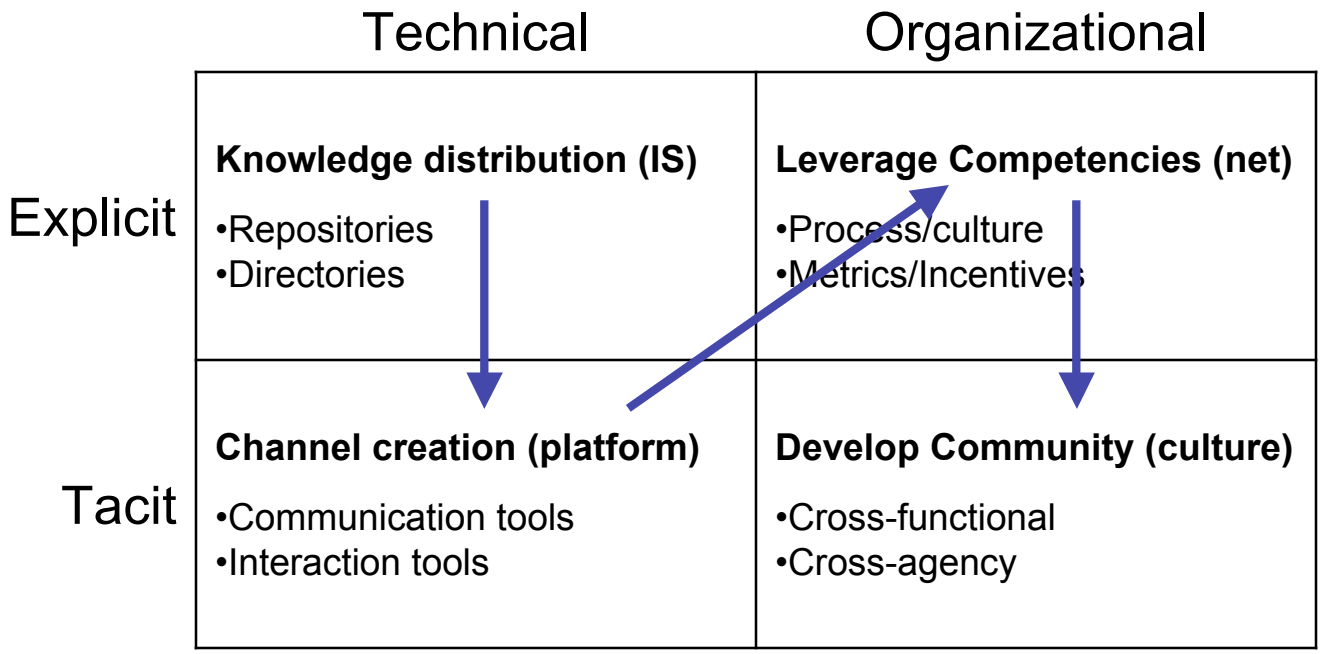
As interactive complexity (knowledge creation) rises, management system must shift to support more social, cognitive,and behavioral, and less technical, factors.

Developing organizational information and knowledge capabilities - *kinds of initiatives*

	Technical	Organizational
Explicit	**Knowledge distribution (IS)** •Repositories •Directories	**Leverage Competencies (net)** •Process/culture •Metrics/Incentives
Tacit	**Channel creation (platform)** •Communication tools •Interaction tools	**Develop Community (culture)** •Cross-functional •Cross-agency

Knowledge creation has to be enabled, not controlled.

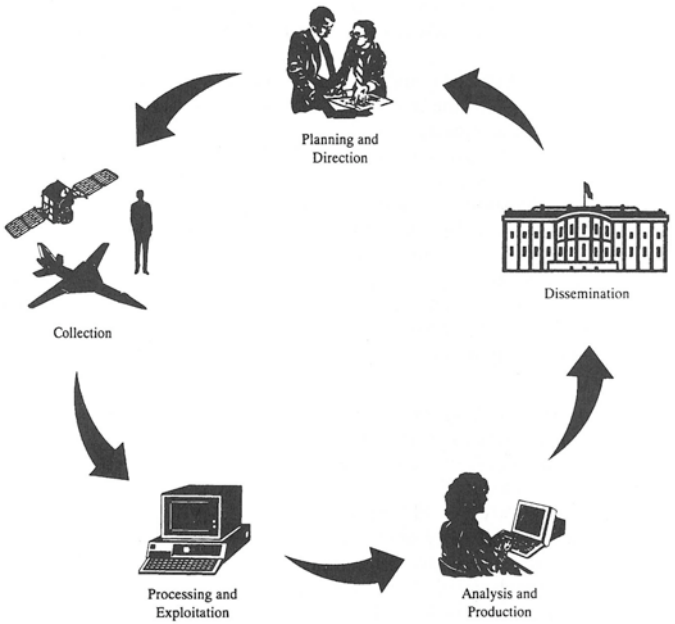
Systems development



Knowledge creation has to be enabled, not controlled.

Current production model: Has the "intelligence cycle" evolved?

1993



CIA *Consumer's Handbook* 1993

Today

CIA *Factbook on Intelligence* 2005

Intelligence Commission: is the problem rotational direction?

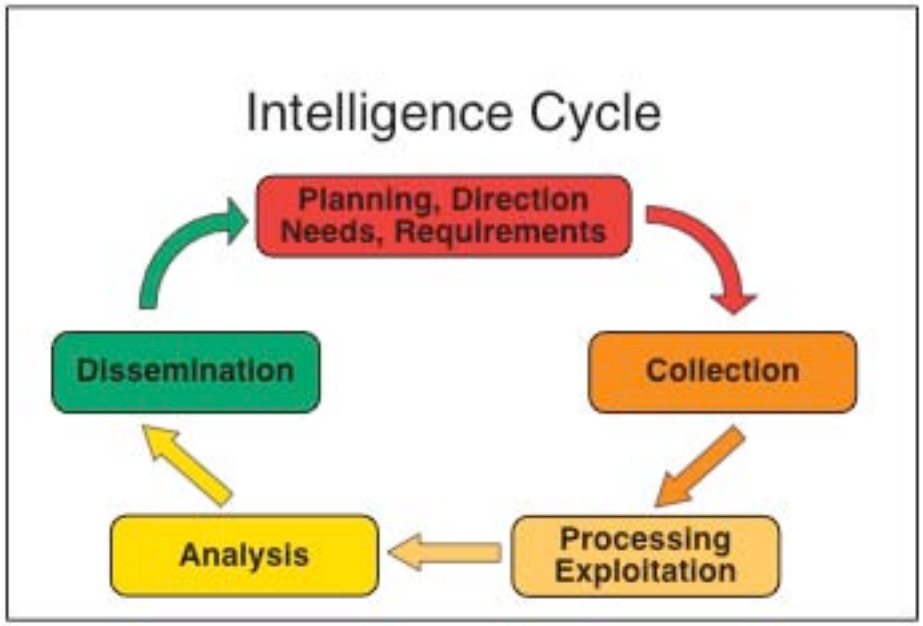
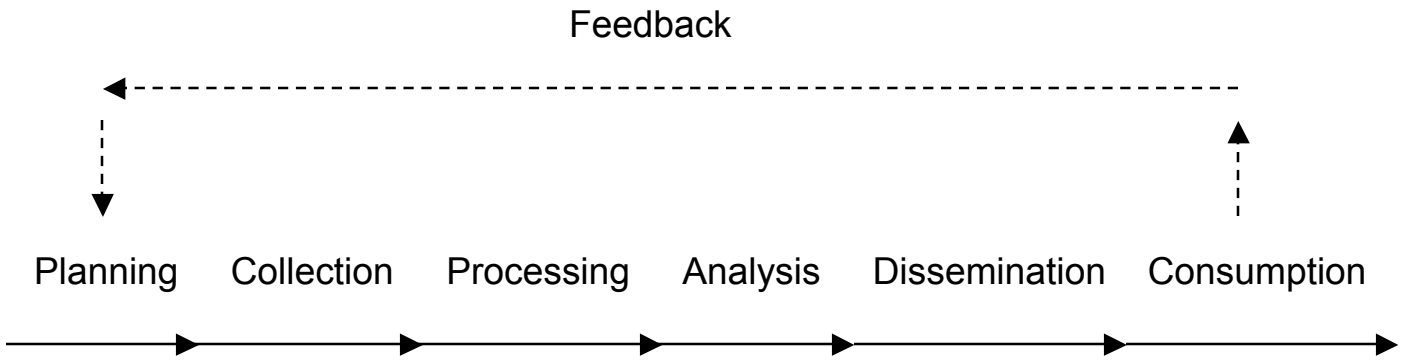
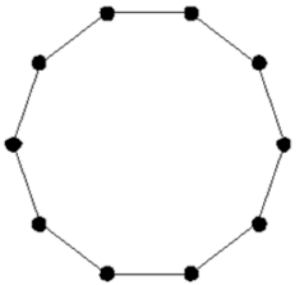


Figure 1. The Intelligence Cycle

Source: Commission on the Intelligence Capabilities of the United State Regarding Weapons of Mass Destruction Report Appendix C

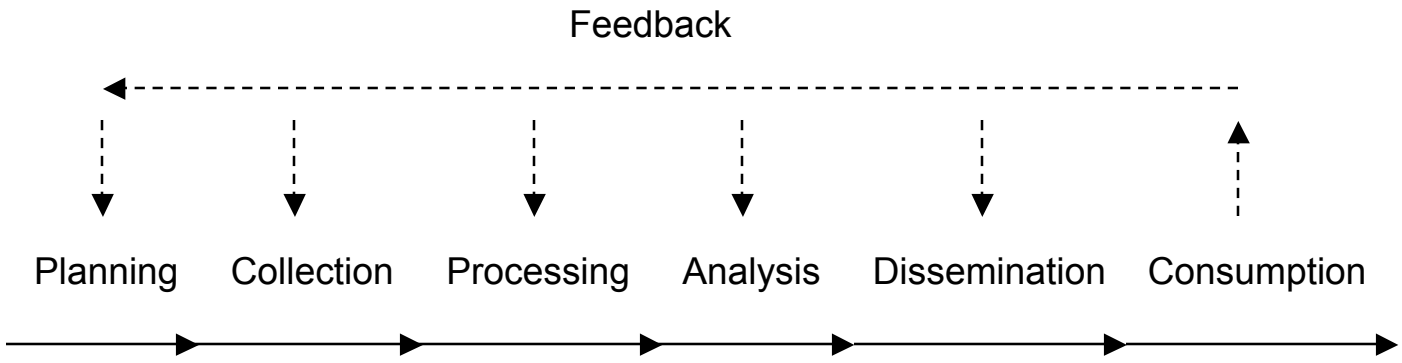
Real problem: the existing intelligence "cycle" is a linear network



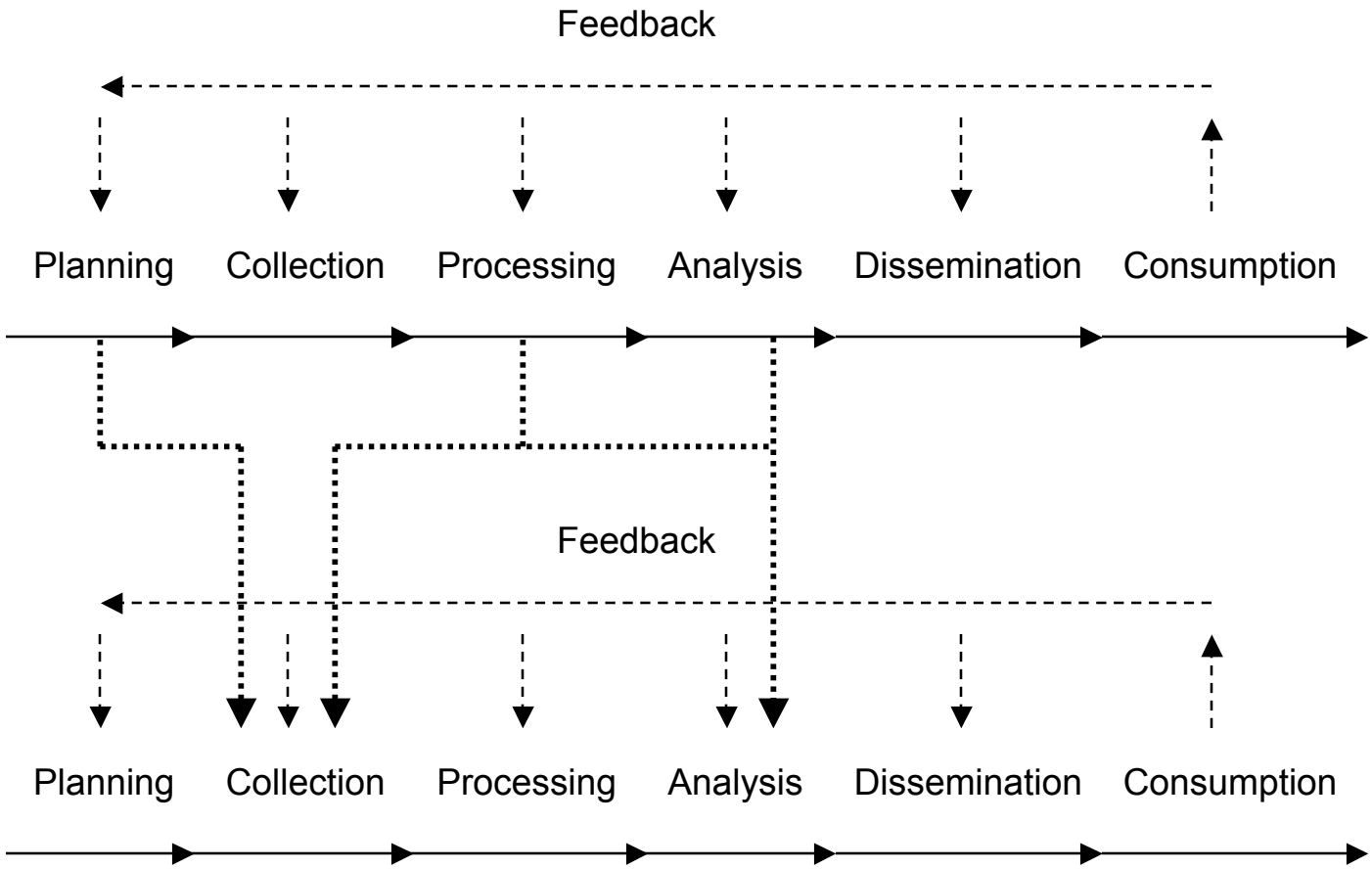


A linear network topology

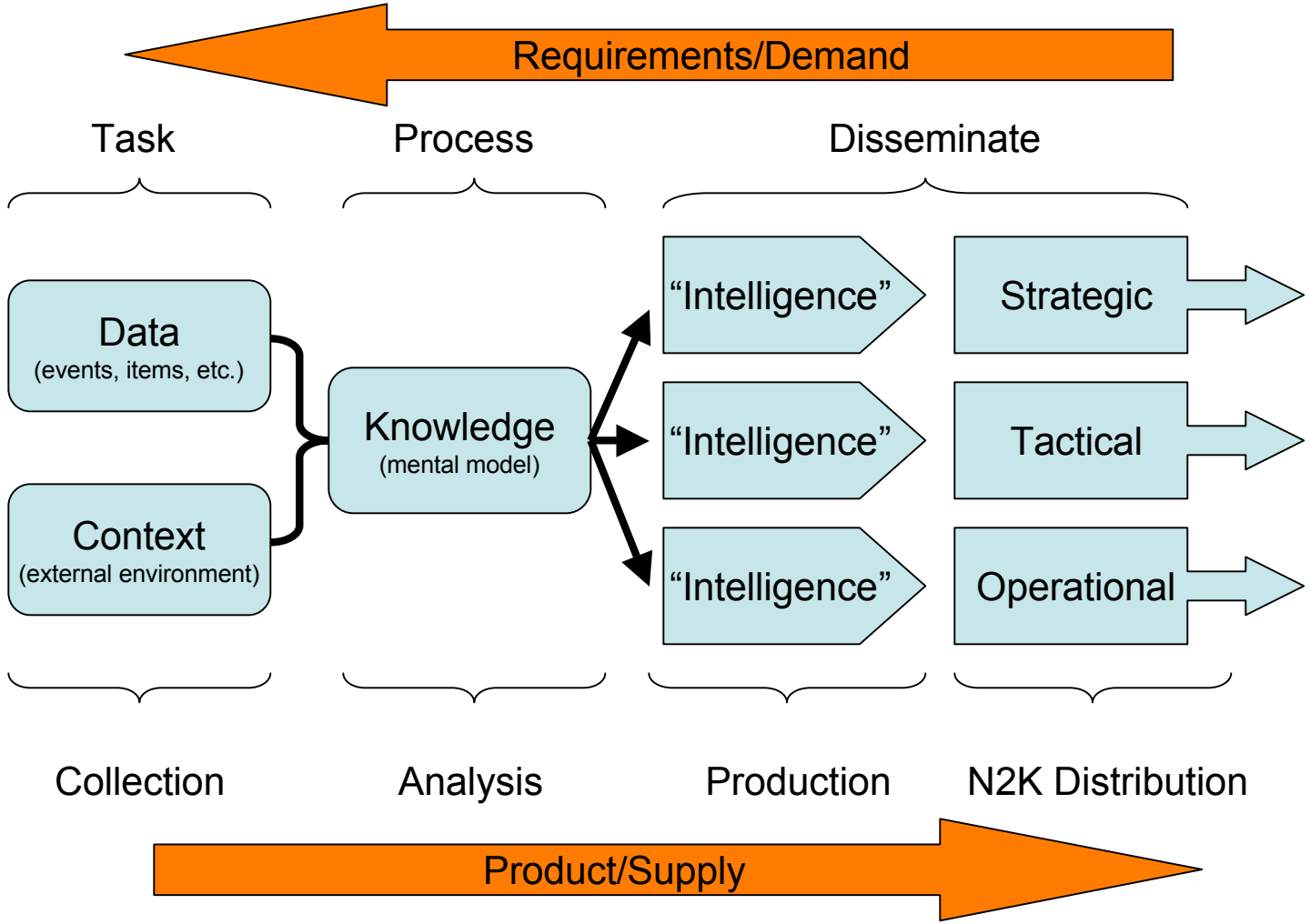
… is linear and multi-cyclical



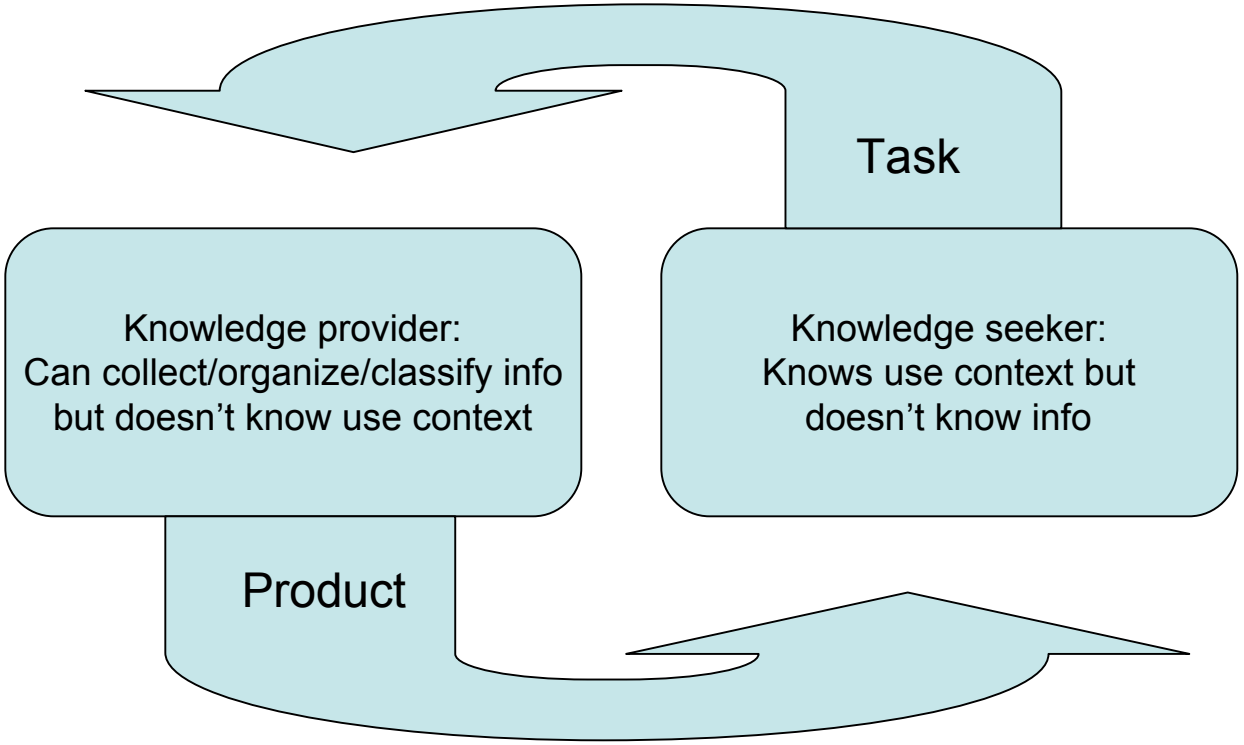
… is multi-linear multi-cyclical



But reflects *industrial model* of intel production



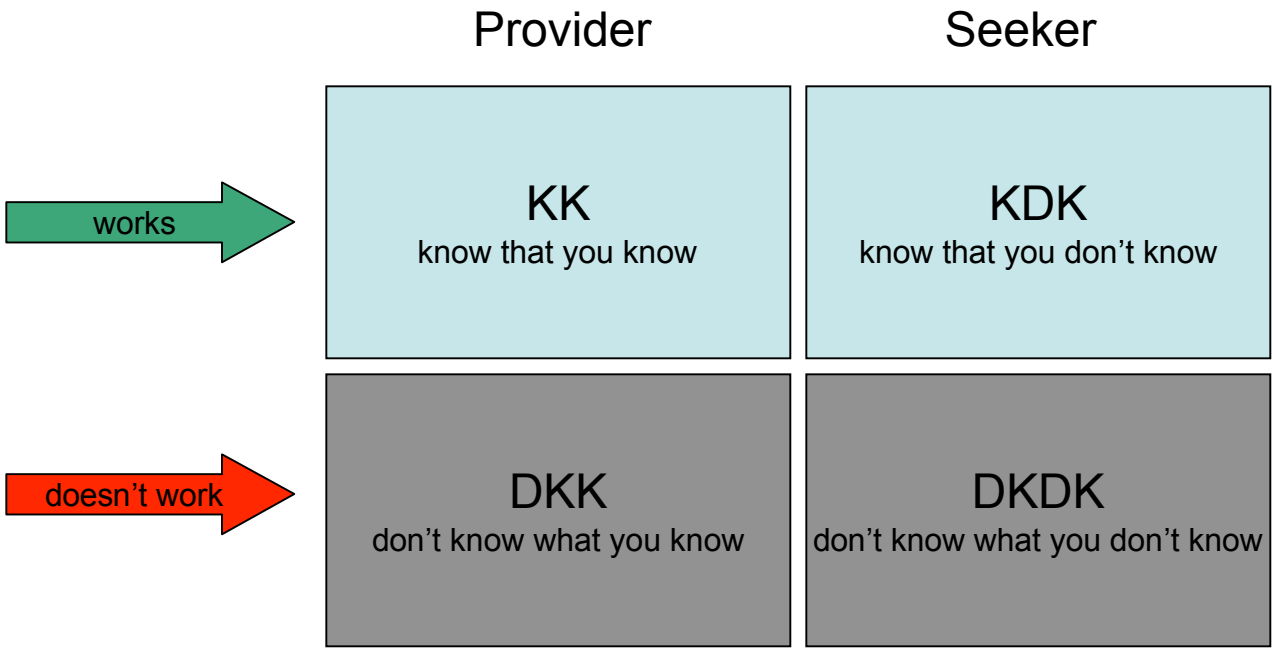
Problem with industrial model is the asymetric feedback loop and the deficiency of task/product integration



Cf. tech support (how do you do this vs here is what I want to do)

Model doesn't meet information needs

Existing intelligence process is based on
knowing what you know or knowing what you "need to know"!

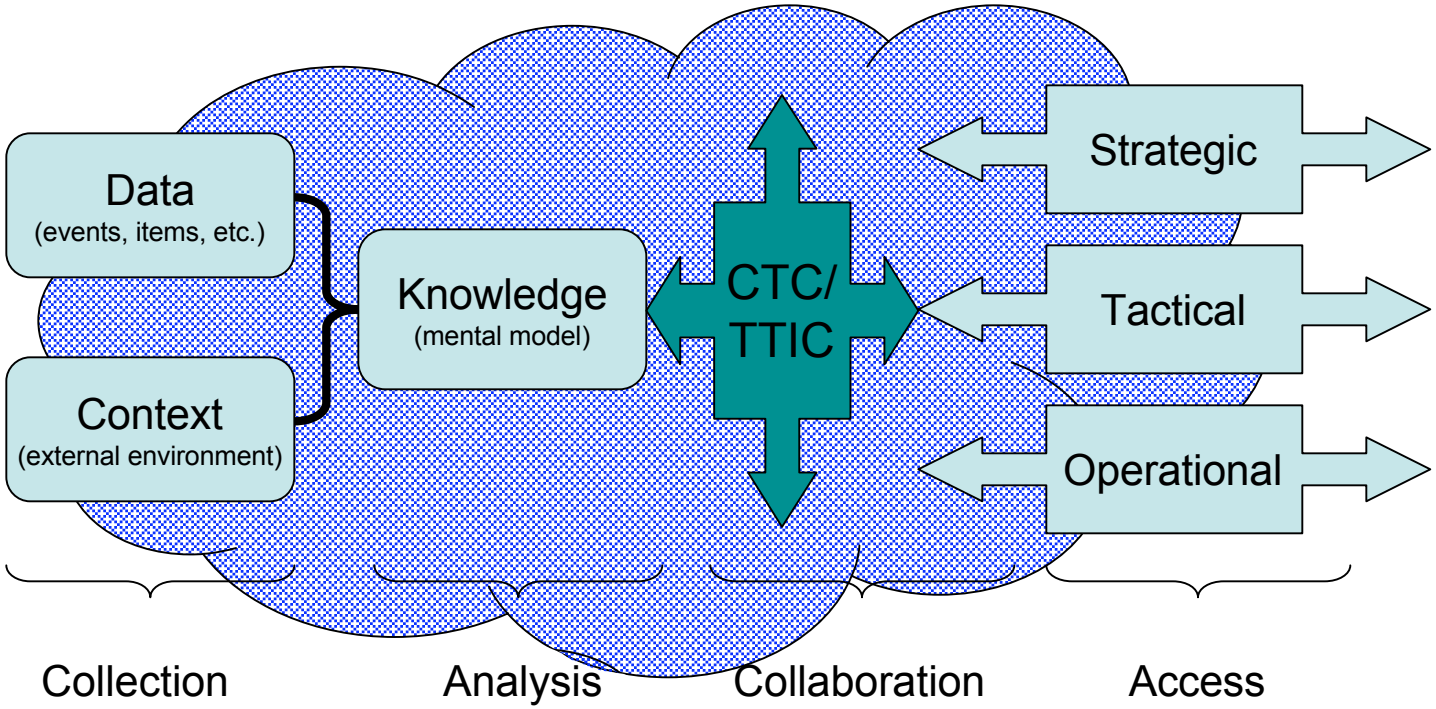


Footnote: same information needs exist in any information system

Technologies required to meet information retrieval needs:

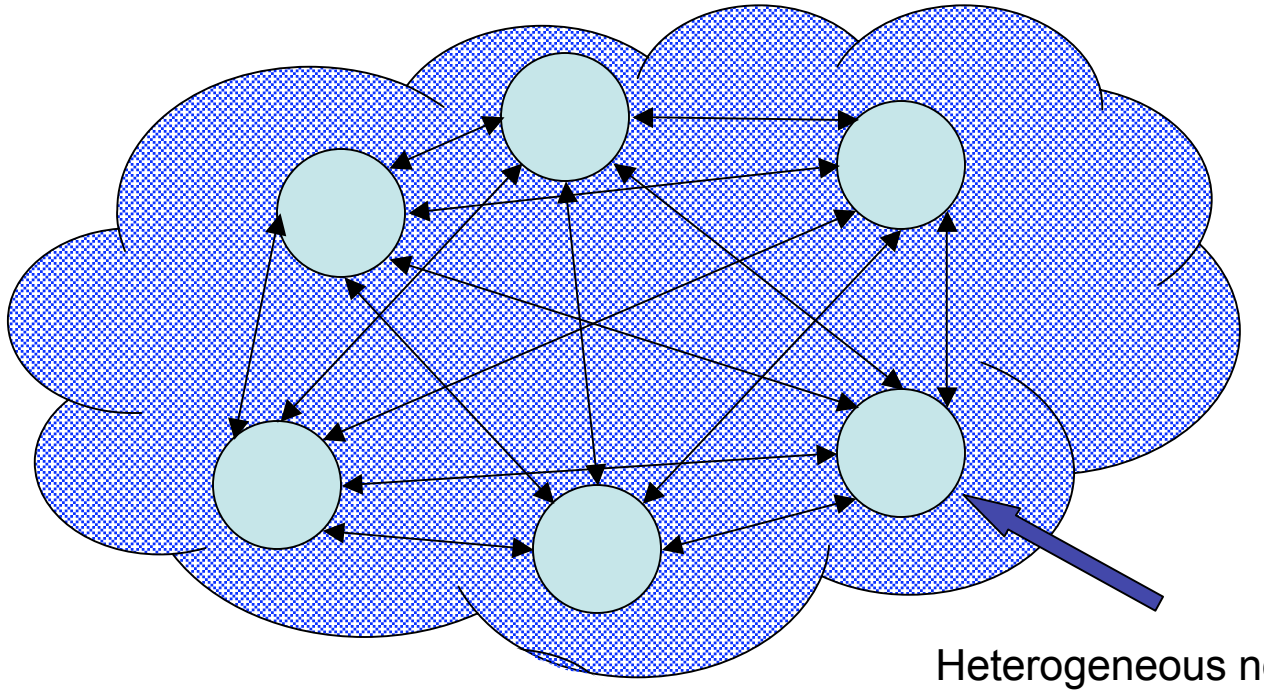
KK know that you know **SIMPLE SEARCH**	KDK know that you don't know **ACCESS/DIRECTORIES**
DKK don't know what you know **KNOWLEDGE DISCOVERY**	DKDK don't know what you don't know **INTELLIGENT AGENTS**

Is emerging "need to share" model (star or hub network) the answer?



All-source *fusion center approach* ("*we know*"). Better, but still info-centric and ignores complexity and novelty.

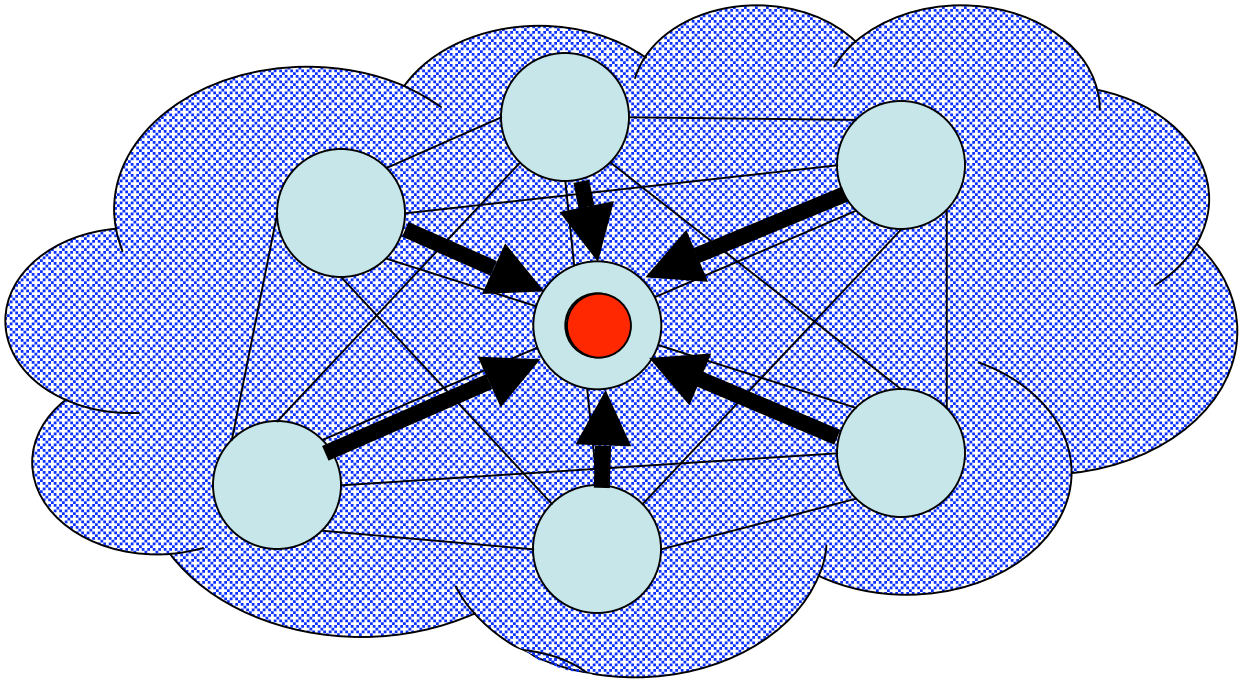
Better still (all-channel network): knowledge-centric learning network



Heterogeneous nodes w/ competing and complimentary capabilities and competencies

Why? To swarm resources

to enable unplanned responses (innovation) to surprise stimuli

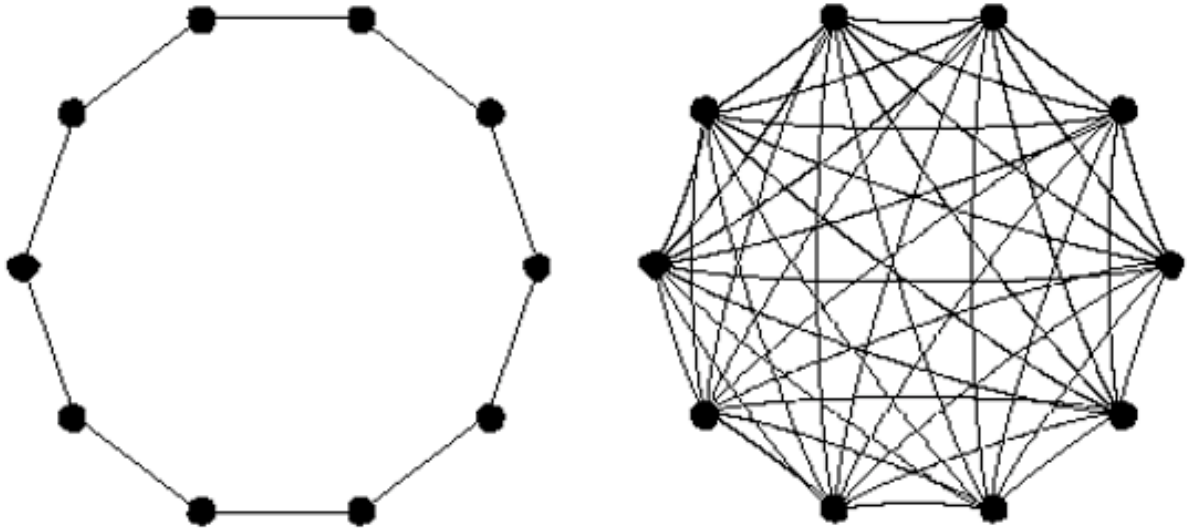


[*Cf.* *immunology*; and infosec *operational resilience*]

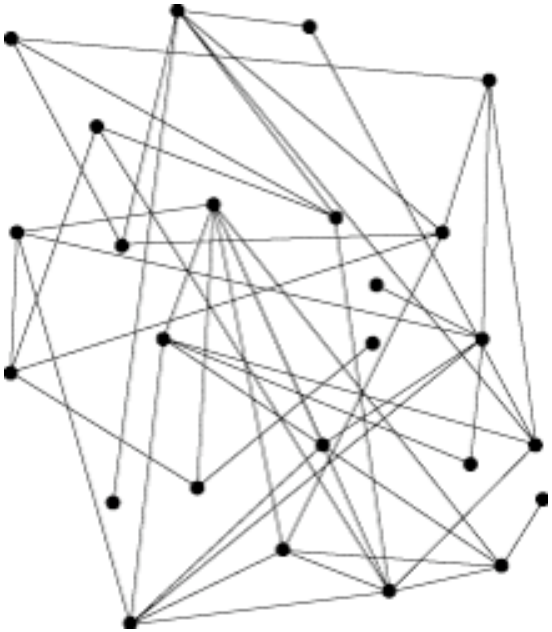
Intelligence and network theory

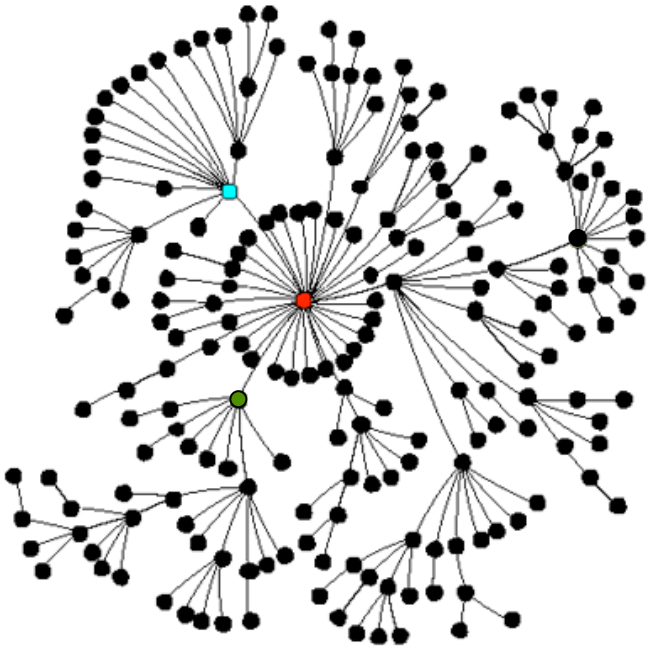
- Channel availability: linear vs all-channel
- Network organization: random vs scale-free
- Weak links vs strong links
  - Formal vs informal
  - Routine information transfers vs extraordinary
- Lessons for intelligence production (and privacy … later)

Communication channel availability: linear vs all-channel topology

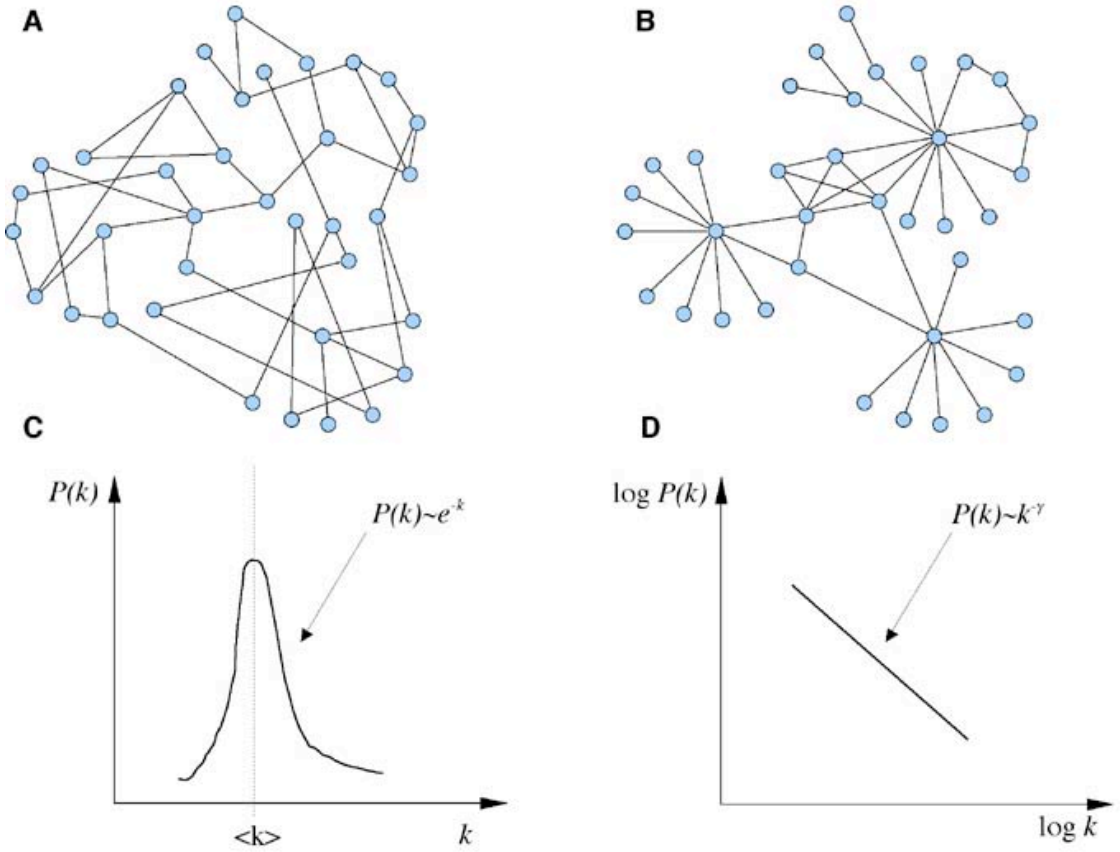


Network organization: random vs scale-free linking
(links, nodes, and super-nodes)

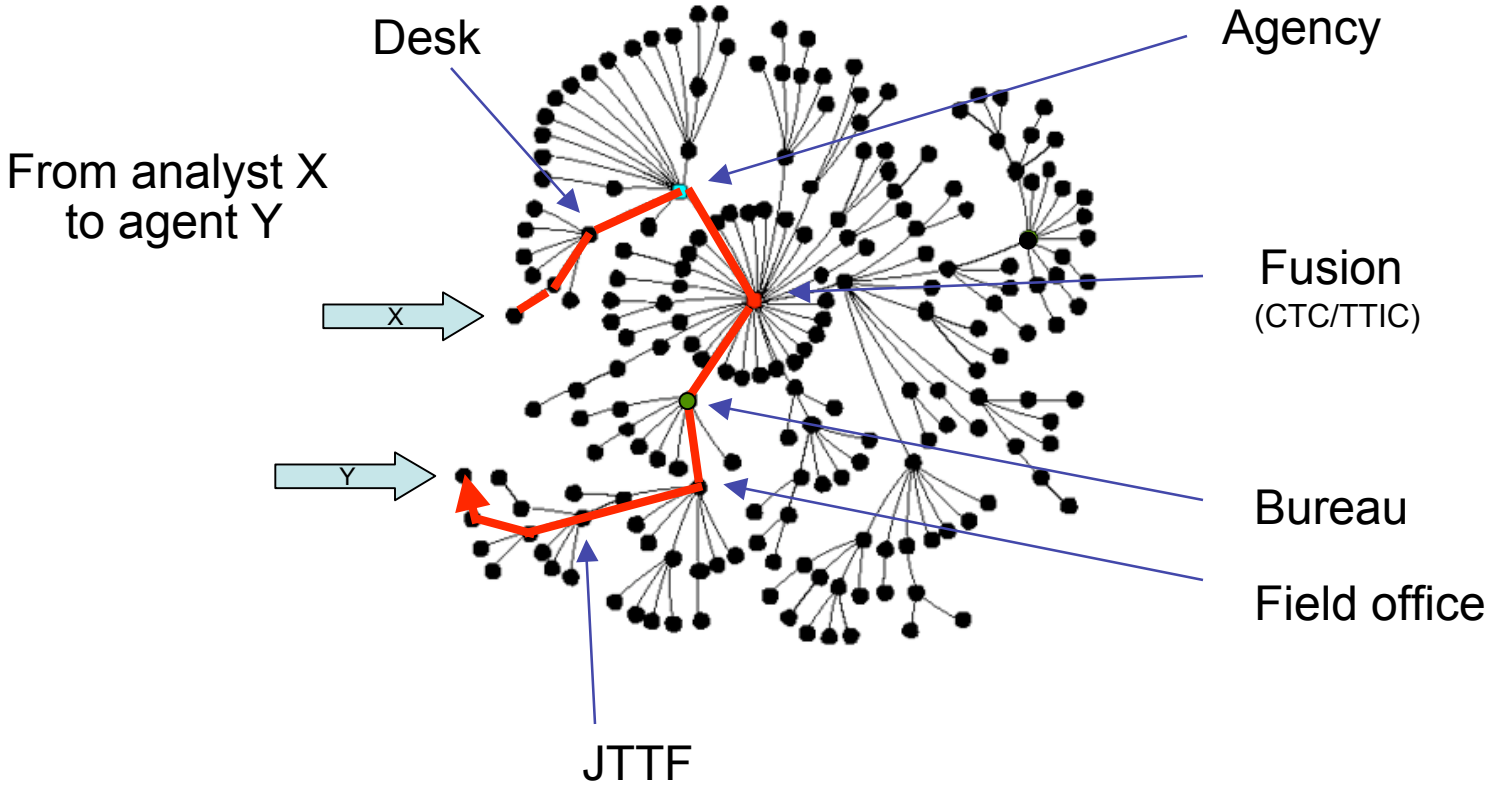




Random vs. scale-free networks



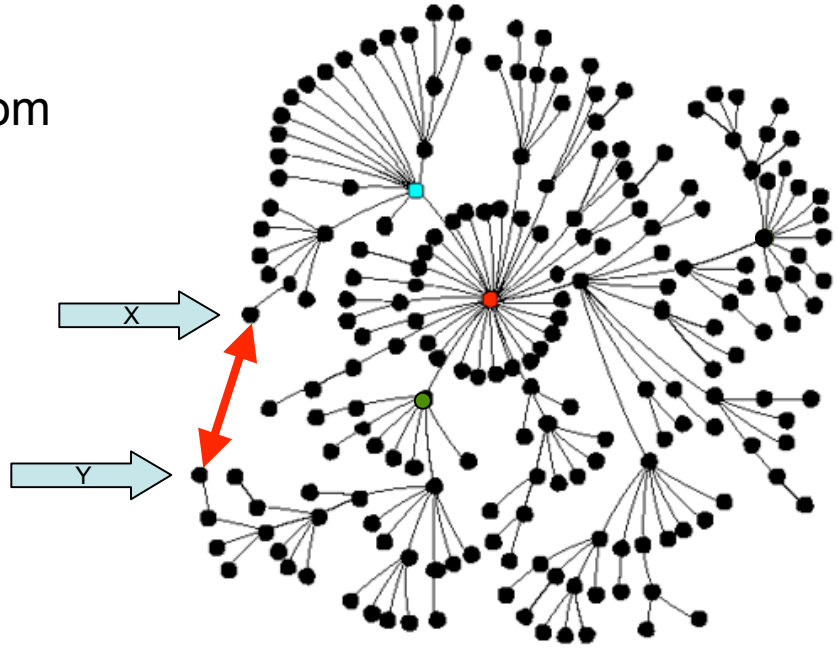
Implications for intelligence:
routine information sharing travels on formal/strong links thru supernodes (e.g., MOUs)



But, extraordinary information can travel on informal-weak links between clusters

(assuming channel is available)

Directly from X to Y



How do you build a community with viable weak link sharing of "routine" information that may only be extraordinary in shared context?

Stages of knowledge community evolution
(organizational change)

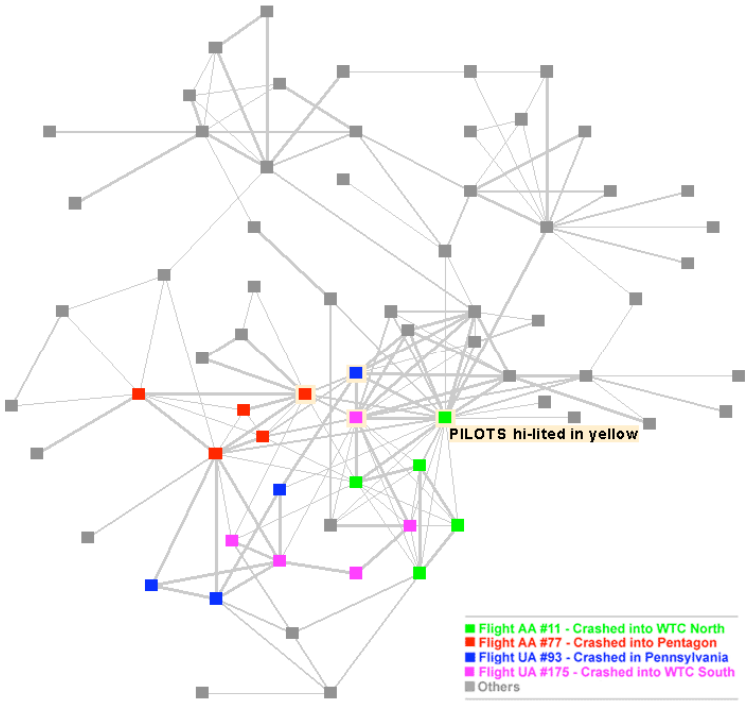
- *Individual fiefdoms* (CIA, NSA, NRO, FBI, State, DOD, etc.) (silos)
- *Process/product communities* (TTIC, CTC, etc.) (process hubs facilitate information flows)
- *Cross-functional communities* (JTTF) (heterogeneous hubs facilitate capabilities and competency sharing)
- *Enterprise best practices* (distributed knowledge useable throughout network)
- *Trusted community* (networked knowledge can be swarmed against threats or problems - *collective intelligence*) (ISE/Markle)

ISE ≈ scale-free all-channel network

with *complementary competencies* and *capabilities*
that encourages *entrepreneurial innovation*

- Some other considerations
  - *Creative abrasion* facilitates rich analysis (maintain competing intelligence centers with different POVs)
  - *Redundant* analytic competencies and capabilities insure depth
  - *Share inputs, compete on outputs* ("need to post")
  - Facilitate formal and informal coalitions, taskforces, joint projects and other *cross-functional or cross-organizational* interactions
  - Excess *formalization* prevents innovative behavior (cf., MOUs)
  - *Overcome* "not invented here" syndrome and project/taskforce *ownership* issues (separate analytic functions from operational, thus require cooperation) (e.g., Dalton)
  - Develop contribution based *incentives* (individual and institutional)

Footnote: modeling/destabilizing terrorist networks



See Disrupting Terrorist Networks 2005
http://counterterrorism.information-warfare.info/

New IC paradigm for knowledge production

- The intelligence community should model itself on the *immune system* (2nd Markle Report, Appendix C, 2003) (not sensory)
- The technical infrastructure to support an immunological approach requires developing the ability to recognize and identity novel threats throughout the system (i.e., *distributed capabilities and competencies and relevant data/information sharing*) and to swarm intelligence resources in response
- This requires managing information for *"knowledge creation" not* just production and consumption of intelligence *products or* undifferentiated information *sharing* (i.e., manage for knowledge creation at the right time and right place)
- Thus, more about *community building* than technical infrastructure building

Best practices for KMIC

- Idealized academic research model
- NGO and civil society model
- Government (WHO, CDC, EPA …)
- Open source movement, SETI, etc.
- Corporate model
  - Financial services (MS research, Buffett AR)
  - Pharmaceutical development (distributed research model)
  - R+D firms
- *Internal* -- look for entrepreneurial intelligence successes across IC and replicate conditions (*cf.* VC incubators)

Security and liberty

- Security and liberty are *not dichotomous rivals* to be traded one for another in a zero-sum game
- *Dual obligations* of a liberal democratic state, each to be maximized within the constraints of the other (wicked problem)
- There is no security without liberty; no liberty without security (John Locke, Founding Fathers, *et seq.*)
- *Cf.* liberty and *privacy*, and liberty and *license*

Risks to liberty

- Chilling effect (inhibit beneficial innocent behavior)
  - *Need for dissent* (Shannon, Bateson's rule, noise, new learning)
  - System stability (noise creates instability, but is required to meet novelty) (brittleness; catastrophic (USSR) vs elegant failure)
  - NB: USSR collapse was intel "failure" because no noise to analyze
- Slippery slope
  - *Normalization* of the extraordinary
    - Security paradigm (e.g., proliferation of ID systems, etc.)
  - *Mission creep* (proportionality)
- *Play [R]oom* (vs "innocent") and freedom
  - need to feel free to engage in social experimentation
  - *E.g.,"*cars, cigars, and bars" [see *Play Room in the Nat Sec State*]

"Security" as business process in CT

- Produce "actionable intelligence"
  - Potential catastrophic outcomes = political consensus for *preemption*
  - Preemption requires intelligence (information to anticipate future acts) (cf. *foreknowledge, prescience, precognition*)
  - Intelligence ~ knowledge creation
    - See the unknown: "connecting the dots", summarize, model, etc.
    - See the unknowable? Predictive modeling/statistical probabilities
  - "Actionable" = *Consequences* to civil liberties
    - Disruption/preemption (NatSec) vs. prosecution/reaction (LE)
    - *Cf*. presumptive heuristic - default state = innocent (Rosenzweig)
  - *Preemption, not sharing, is the challenge to civil liberty*

Privacy interests

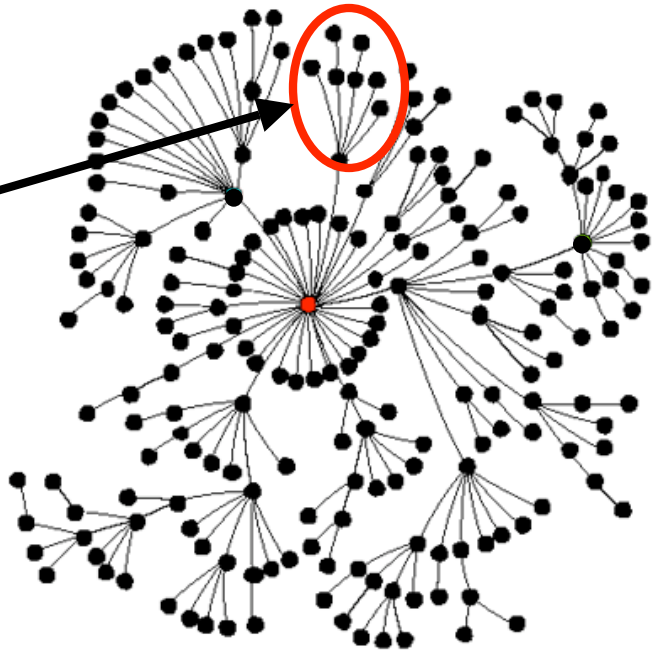
- Is privacy a first order or second order value?
  - Interests, rights, and legally protected rights (is it reasonable for society to recognize expectation of privacy)
  - (2º) *privacy protects civil liberties* (inefficiencies/practical obscurity)
  - But, is it 1º? Legal protection is contextual and privacy is alienable
  - Efficiency (technology) and *sharing challenges privacy*
- Parsing privacy interests (*Whalen* 1977 fn. 24) (Rosenzweig)
  - *Secrecy*
    - Keep things unknown
  - *Anonymity*
    - Keep things unattributed
  - *Autonomy*
    - Keep bad consequences from occurring from attribution

Parsed privacy values

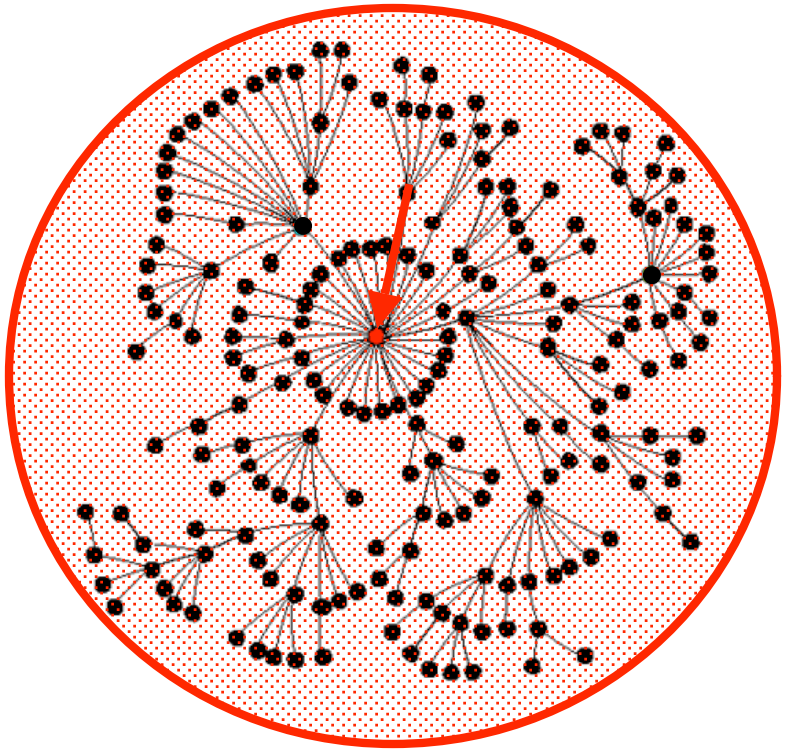
- Secrecy (protects not knowing "data") (2º?)
  - "get over it" (McNeely), no longer viable in dataverse/info society
  - Sup. Ct. 4th A analysis is premised on secrecy (*Miller*, *et cetera*)
  - "All or nothing" rules based on "acquisition/collection" are not viable
- Anonymity (protects individual) (1º?) (not attributing)
  - separate knowledge of behavior from knowledge of identity
- Autonomy (protects individual) (1º?) (proportionality)
  - due process (about fairness and power) (observation as predicate for consequence; error correction) (~ exclusion rule)
- Privacy right ≠ planning terrorist attacks in secret w/o discovery, thus, issue for CT can't be secrecy, rather it is *confidence interval* (limit collateral damage) and *due process* (correct errors)

Problem w/ all-or-nothing policy (and reuse): information shared locally for one purpose …

Can have legitimate expectation of privacy for locally shared information



… can flood throughout the network when it hits a super-node

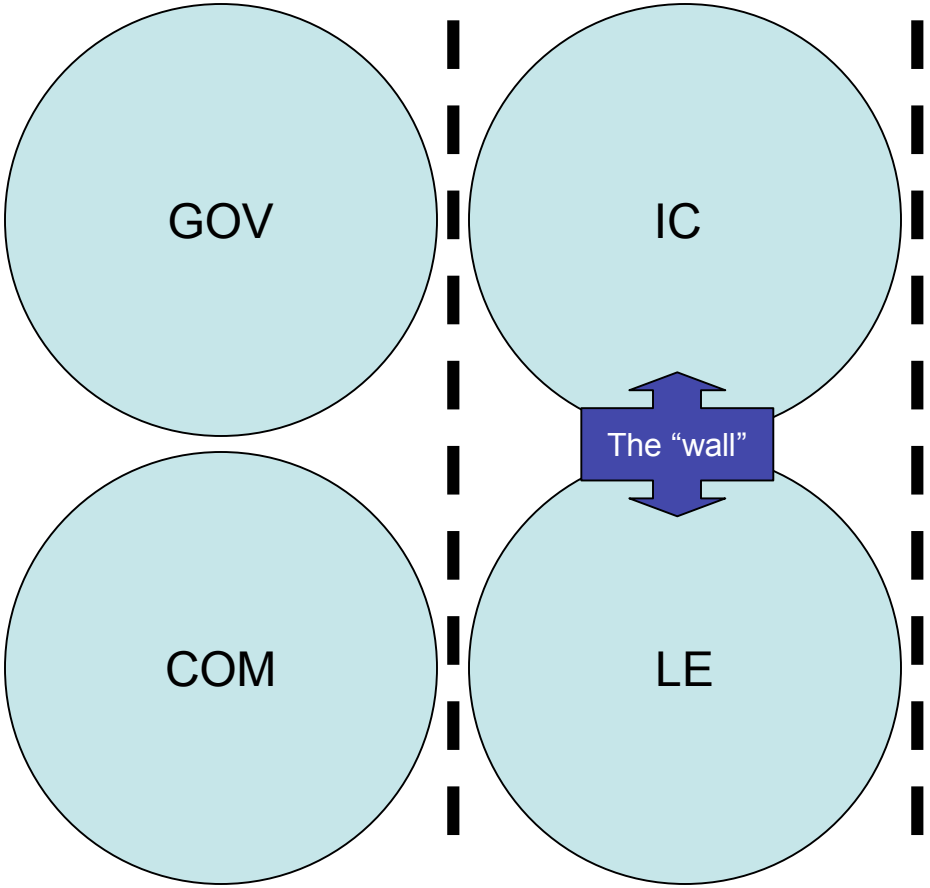


It may no longer be reasonable to say that once anyone in the network has information everyone should for any purpose

Inadequacy of statutory schema based on discrete domains and controlling collection …

Privacy Act
CMPPA
Census
Tax/IRS
DPPA,
etc.

FCRA
VPPA
CCPA
Telecom
GLBA
HIPAA,
etc.

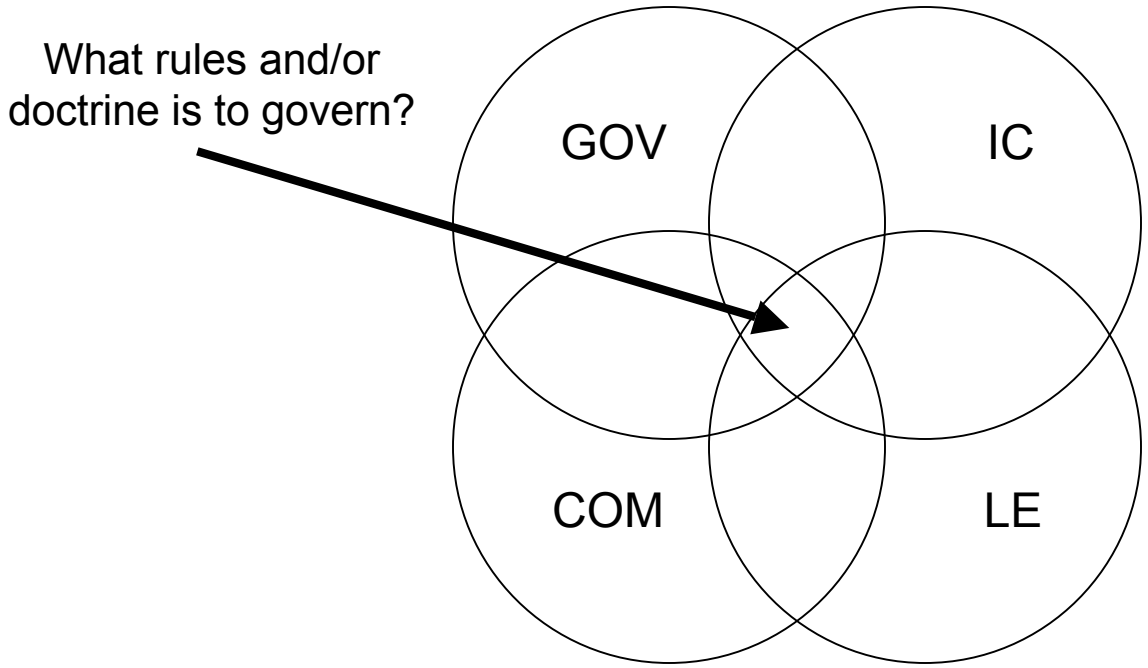


IOA
EO

FISA

Title III
ECPA

… is inherent in the dataverse and network:
where *use* (and *reuse*) not *collection* is the issue



4 domains
= at least
13 use cases

data largely "exists" (new data economics) and question is under what circumstances can it be *used* for any particular purpose (i.e. *accessed*, *analyzed* and *attributed*) and what *consequences* result from its use

Maintaining anonymity

- Separate knowledge of behavior from knowledge of identity (control *data attribution* see *Frankenstein* 2004)
- Anonymization of data for information sharing, matching, and analysis technologies
  - One-way hashing for watch lists
  - Searching on encrypted data for analysis
- Pseudonymization of identity for identification systems and collection technologies
  - No government monopoly on ID, encourage multiple issuers and use of multiple nyms (see http://releases.usnewswire.com/GetRelease.asp?id=39202)
  - Use one-to-one strategies, not one-to-many, etc.

Maintaining autonomy

- Due process (fundamental "fairness") principles
  - Investigative tools, not evidentiary (*cf*. *watch lists*) (*see Secure Flight*)
  - Openness, notice, transparency, options/choice
  - Accuracy and error correction (*Kafka* v. *Big Brother*)
- Political control
  - Authorization
  - Oversight
  - Review
  - Checks and balances
  - Transparency

Confidence interval

- Limit collateral damage
- False positives (bias and decision heuristics)
- False negatives (bias and decision heuristics)
- *Variable to threat environment*
- *Cf*. “perfect machine” (Rosenzweig)

Due Process

- Fundamental fairness (Dworkin *Law's Empire* 1986)
- Due process analysis factors:
  - Predicate (general vs specific, *Martinez-Fuerte* 1976) (probabilistic vs probative, *Cortez* 1981)
  - Alternatives (least intrusive alternative, proportional)
  - Consequences (minimization -- *cf*. USSID-18 minimization of intrusion, selective attention standards based on CI)
  - Error correction (authorization, review, remedy, remediation)

In any case, security = privacy

- Both security (intelligence production) and privacy *require the same information management and infrastructure*:
  - *Distributed access and user accountability* (authentication, logging and audit)
  - *Information assurance and data quality* (protection, error discovery and correction)
  - *Rules-based processing* (policy enforcement appliances)
  - *Selective revelation* (policy enabling appliances)
- Enterprise architecture (data and interoperability)
- Social construction (architecture and organizational change)

Rules-based processing

- *Limit the scope* of inquiry or use based on:
  code plus policy rules (*policy appliance*)
- Two aspects (move from client/server to agent/computing)
  - "*intelligent agent*" - credentialed query accessing distributed DB (proof carrying code, analytic filtering) (~ bots)
    - *Mobile agents* reduce network load, operate in real time, encapsulate protocols to move seamlessly between heterogeneous systems, perform asynchronous and autonomous execution, and are fault tolerant
  - "*smart-data*" - meta-data labels (or wrappers) specifying information about data and how it can be used (ICMWG - IC XML Schema)
- ~ web services provider (controlled processing)
- ~ shrinking perimeter of defense (systems, application, data)

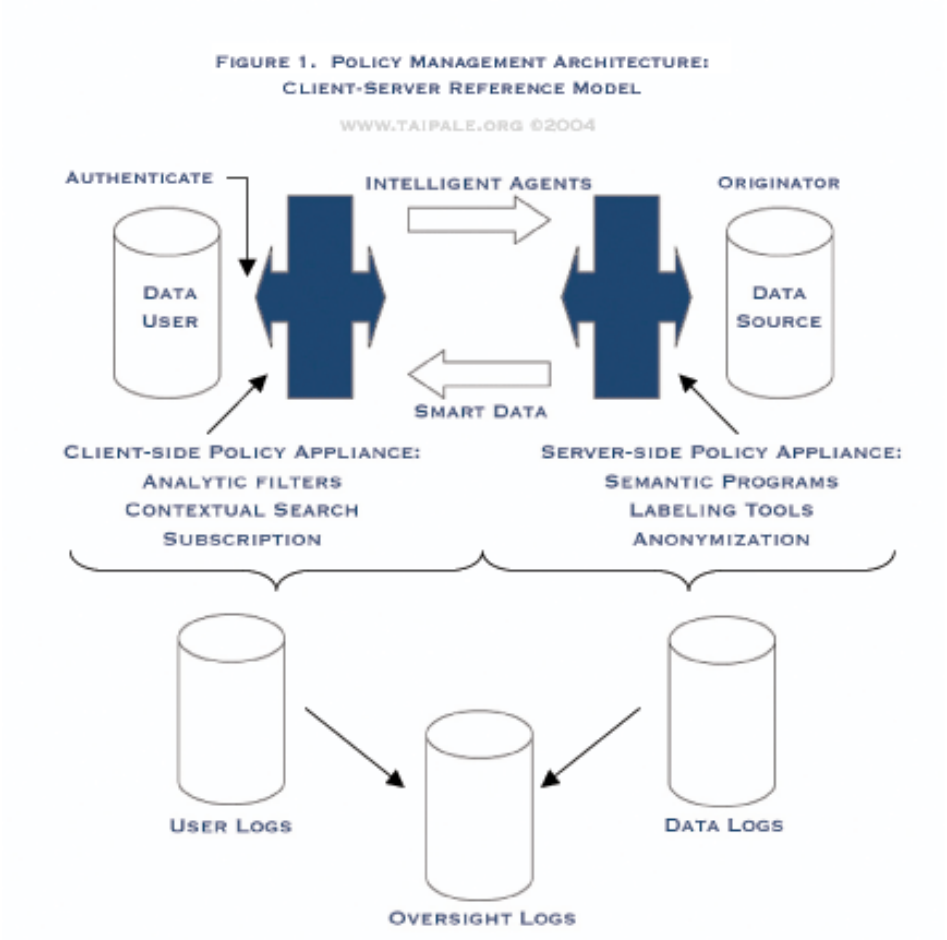
Selective revelation subject to due process

- Iterative, incremental revelation of data
  - Initial revelation by statistics or categorical analysis
  - Subsequent revelation(s) justified on prior results
  - Need to develop predicates and authority/review mechanisms
- Allows for data analysis without revealing personally identifying data (or other attributes, e.g., tearline for sources and methods)
- Anonymized/de-identified data
  - One-way hash v. cryptographic approach
  - Staying one step ahead of data analysis identification
- *Impose appropriate legal or administrative procedure before each revelation or attribution*

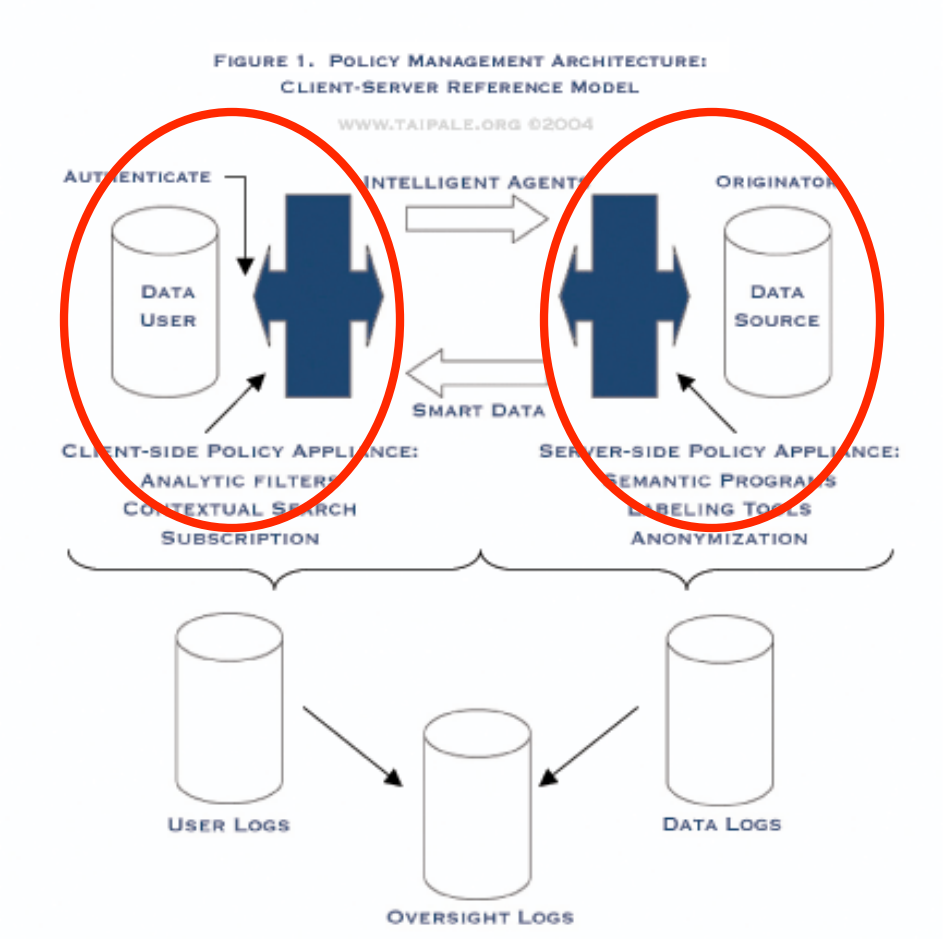
Strong credential and audit

- *"Watch the watchers"* (democracy demands accountability)
  - Technology creates potential for abuse but also allows for "perfect" oversight and control (good test for trust, "eat your own dogfood")
  - Immutable audit trails for users and data
    - Distributed
    - Cross-organizational
    - Cross-validation
  - But, *safe harbors, incentives, and work environment* needs
- *who controls the logs*? (CIO, CTO, CSO, GC, IG, CPO, external?) (also, are logs themselves subject to PA/FOIA, etc.)
- Query: can certainty of *post hoc review* help make up for eliminating or reducing ex ante *predicate*?

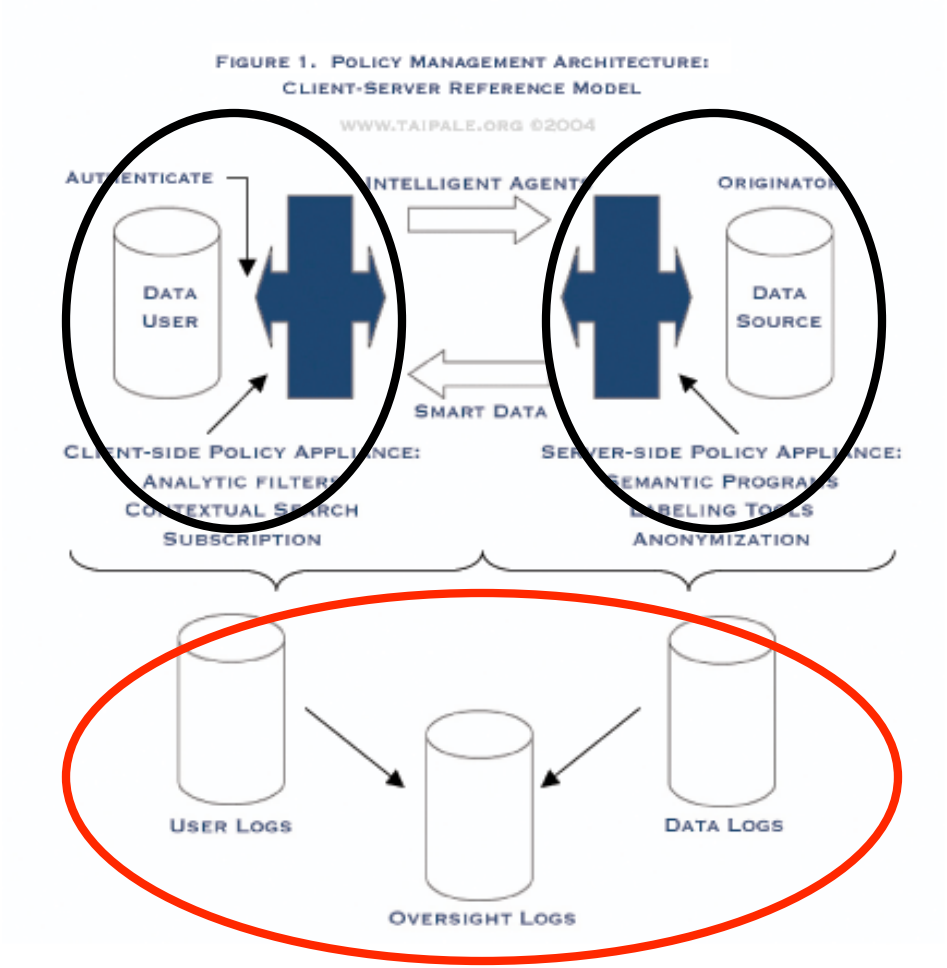
Policy Management Reference Model
(from client/server to agent/computing)



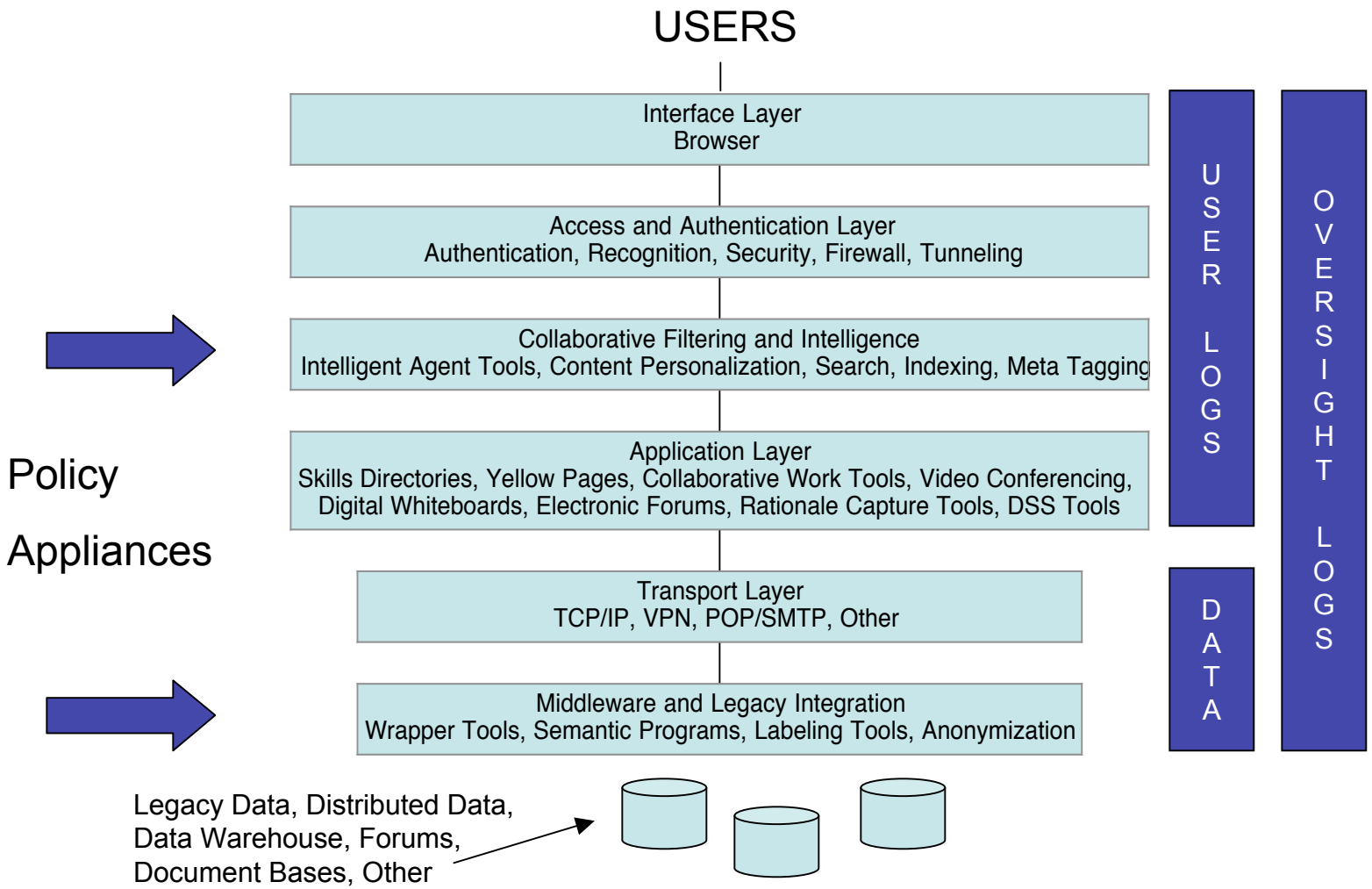
Distributed rules-based processing



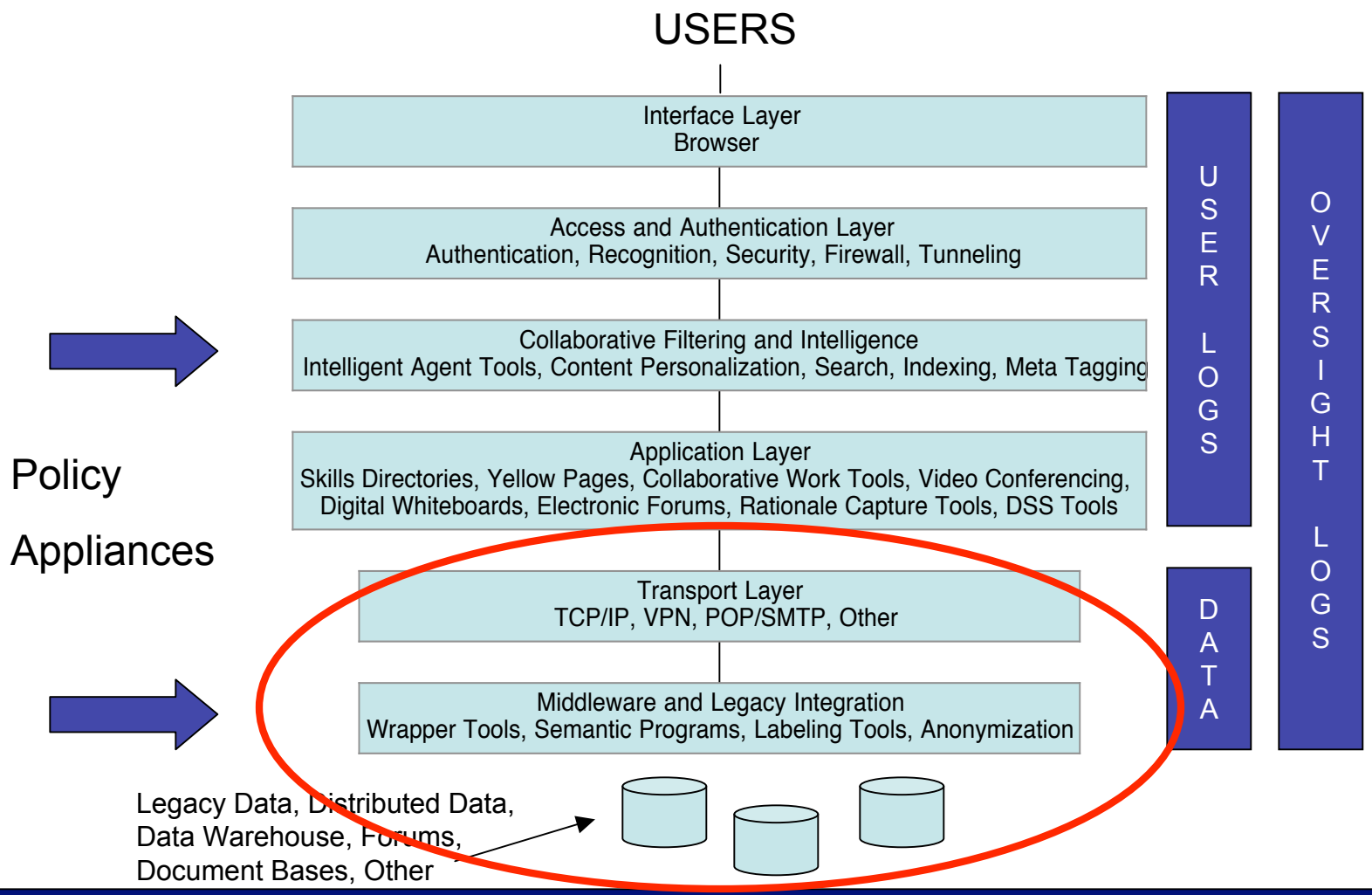
Distributed rules-based processing w/ accountability



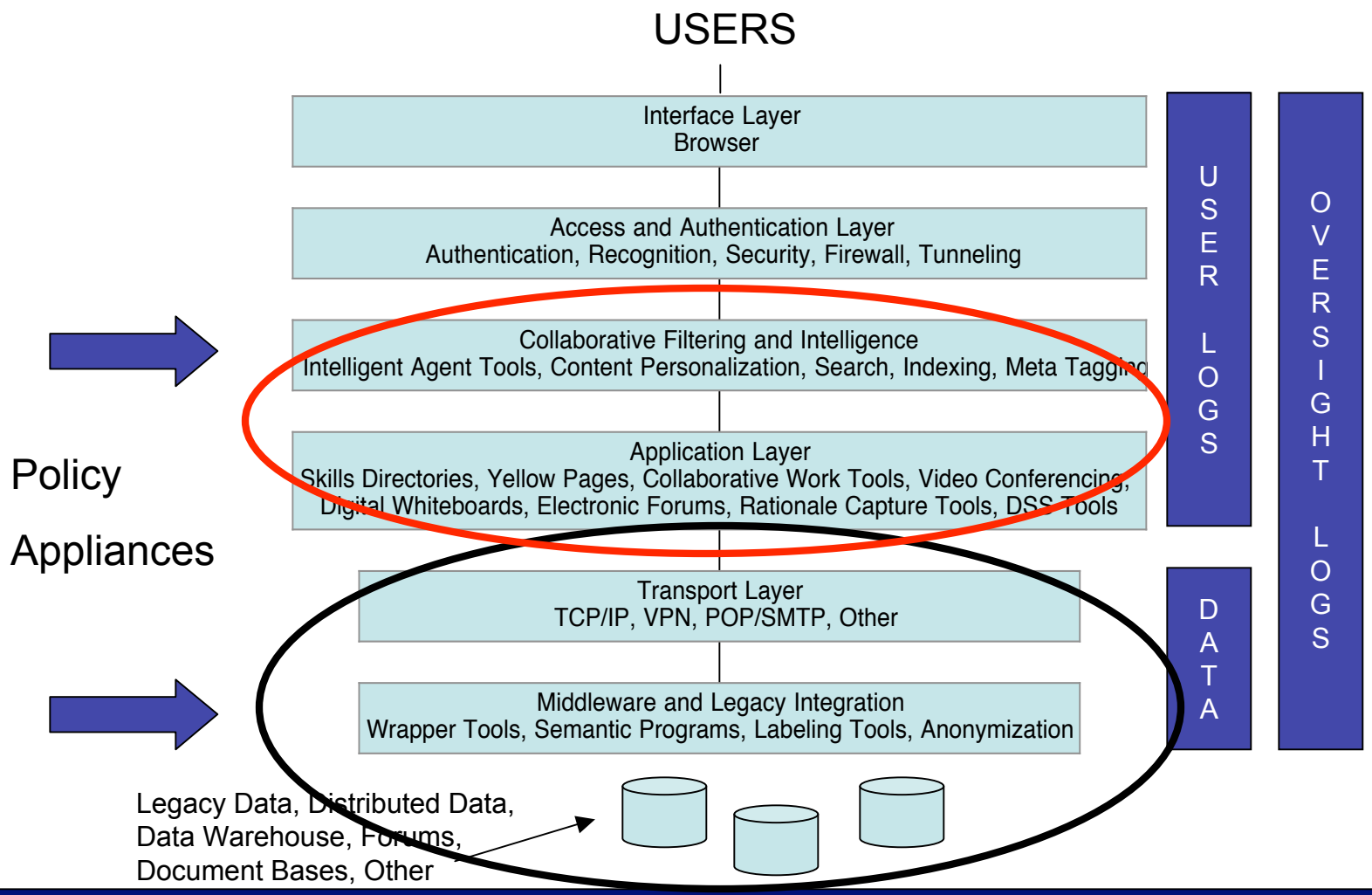
Network Stack Reference Model (7 layer arch.)



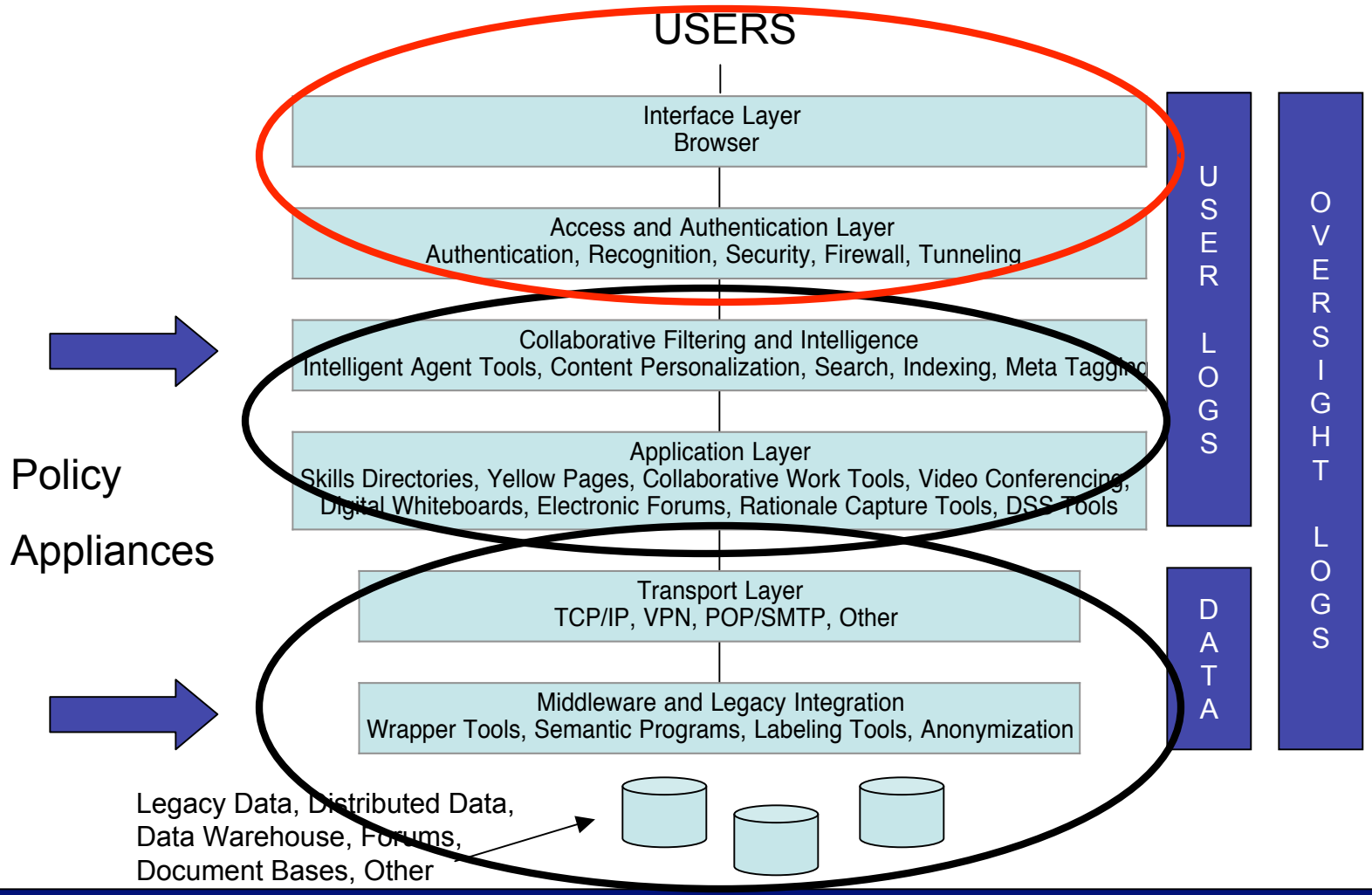
Data and interoperability



Collaboration



Enabled user community



Other misc. issues

- Is *metadata* subject to same rules as underlying data?
  - contexts of the subject
  - the service desired and its security
  - rules for information access and exchange; and
  - descriptions of data elements and resources
- Is the *result of analysis* (output) subject to the same rules as the underlying data (input) (is there a distinction between human and machine product?)
  - related tuples
  - linkages
  - patterns
- Is there a useful distinction for data analysis analogous to the traffic/content distinction in communications surveillance

Open policy issues

- Who authorizes and controls “policy appliances”?
- Who controls the logs? How and by whom are they audited?
- Is the data subject involvement in notice, challenge, and error correction? Applicability of PA, FOIA and FIPs.
- How to protect sources and method (e.g., using classified/sensitive evidence in court)
- Developing incentives, safe harbors, and accountability
- Setting predicate standards for particular access or action
- Relying on open or closed model for operational security (risk management vs “security thru secrecy”)

Conclusion

Goal: technology enabled knowledge management with security *and* privacy.

<meta author="taipale"></>
<meta date="041205"></>
<base href="http://intelligence.information-sharing.info/"></>