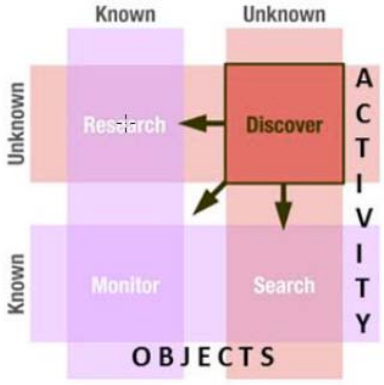
Semantic eScience Comments 25 September 2015
R. Zager

1. "Only Semantic eScience of Security provides the foundation required for capturing both cybersecurity context and experience in order to enable the tradecraft to be captured and utilized to automate and formalize that knowledge. This allows the technology to automatically "connect the dots" between the new data and information entering the system and the knowledge already processed. A key enabler to predictability is understanding what happened before." (p. 36)

Regrettably, understanding what happened before is not generally predictive of the future. Karl Popper, in works such as *The Poverty of Historicism*, established that, except in limited cases, the past is not predictive of the future. Popper observed that the predictive power of incremental information was applicable only in systems which are well-isolated, stationary, and recurrent (like our solar system). The predictive power of Semantic eScience in some systems (such as the heliophysics example) is not applicable to cybersecurity. Cybersecurity, with rapid changes in technology and the adaptive behavior of attackers, defenders and users, is not such a system. In systems that are not isolated, stationary, and recurrent, collecting more information does not equate to having more knowledge. '[N]o society can predict, scientifically, its own future states of knowledge'. (*The Poverty of Historicism*, p. vii)

Events which evade detection using historical precedent are dealt with in Semantic eScience's "Activity Based Intelligence" process which is graphically illustrated on p. 39:



The red "Discover" box is the heart of the matter.

- How much "time on target" does the attacker have before "Discover"?
- At what point of the attack does "Discover" occur? Before the attacker has control of targeted capabilities? After the attacker has accomplished its objectives?
- Is "Discover" the result of Semantic eScience or some other process? Is the attack discovered by a third party, such as the FBI using methods other than Semantic eScience? Is the attack discovered by non-cyber problems (such as credit card fraud)? Is the attack discovered by its damage objectives, such as a defaced website or fake tweets or reading secrets in the newspaper? Is Semantic eScience primarily a retrospective forensics tool?
- What courses of action are made available by Discover? The timeliness of Discover strongly influences what courses of action are available. The OPM course of action of providing credit reporting is not ideal but was the available course of action **after** the data was compromised.

It must be noted that attackers go to great lengths to obscure the provenance of their activities. See, for example:

- JP 3-12(R), Cyberspace Operations, "Because cyber-personas can be complex, with elements in many virtual locations, but normally not linked to a single physical location or form, significant intelligence collection and analysis capabilities are required for the joint forces to gain sufficient insight and situational awareness (SA) of a cyber-persona to enable effective targeting and creation of the JFC's desired effect."
- Miller, Matthew, et al. "Why Your Intuition About Cyber Warfare is Probably Wrong." http://smallwarsjournal.com/jrnl/art/why-your-intuition-about-cyber-warfare-is-probably-wrong
- FBI Affidavit in the matter of Chinese cyber-spy Su Bin, http://online.wsj.com/public/resources/documents/chinahackcomplaint0711.pdf
- Snowden and Manning demonstrate how difficult it is to find the abuse of credentials when the actor is knowledgeable about the victim's processes. Cyberattackers devote significant effort to understanding the victim in order to evade detection.
- Common elements, if any, are often discovered too late to be of value. For example, the Anthem/OPM/ Premera/Empire Blue Cross/ Carefirst attacks had common elements which were discovered long after the perpetrators achieved their objectives.

The value of Semantic eScience, alone or in conjunction with other analytic tools, can only be determined with evidence from field experience. The Philosophy of Science suggests that in cybersecurity (which is not well-isolated, stationary, and recurrent), predicting of the future will be elusive. Nevertheless, rapid responses to unpredicted attacks can be of inestimable valuable. Current response times are, as Verizon observes annually in its DBIR, abysmal. If Semantic eScience can consistently and sustainably drive response times to seconds and minutes instead of months and years, its value will be established.

2. "Semantic eScience of Security infrastructure is an evolutionary approach that applies a scientific foundation to an organization's cyber ecosystem to support continual evolution of the organization's ability to analyze, assess, mitigate, and monitor the cybersecurity of their cyber ecosystem." (p. 40)

In point 1, the limitations of the scientific foundations of Semantic eScience were discussed. The use of the term "evolutionary" to describe Semantic eScience further undermines the concept that Semantic eScience is predictive. The logical inconsistency of evolving systems and prediction was discussed by Popper in *The Poverty of Historicism* under the topic "Is There a Law of Evolution? Laws and Trends." Popper warns, that while trends exist, trends (such as evolution) are not predictive.

Cybersecurity is a human engagement which is not in the nature of meteorology or heliophysics. Neither the weather nor the Sun observe the actions of scientists and change their characteristics in order to defeat predictions. In cybersecurity, the adversaries are constantly changing in response to defenses. See, RAND's *The Defender's Dilemma: Charting a Course Toward Cybersecurity.* http://www.rand.org/content/dam/rand/pubs/research_reports/RR1000/RR1024/RAND_RR1024.pdf

This cybersecurity system is made worse by the unhelpful actions of authorized users. Iacovos Kirlappos, Adam Beautement, M. Angela Sasse; *"Comply or Die" Is Dead: Long Live Security-Aware Principal Agents,* Financial Cryptography and Data Security, Volume 7862 Lecture Notes in Computer Science pp 70-82. "[Y]ou can have the greatest technology and greatest defensive structure in the world, but in the end, never underestimate the impact of user behavior on defensive strategy." Adm. Mike Rogers, NSA Director.

Cybersecurity incidents may play an enabling role in a problem that is remote to the cybersecurity "ecosystem." The recent indictments in the SEC's press release investigation were driven by impacts of cybersecurity that were remote from the cyber breaches. http://www.sec.gov/news/statement/press-conference-remarks-massive-hacking-trading-scheme.html.

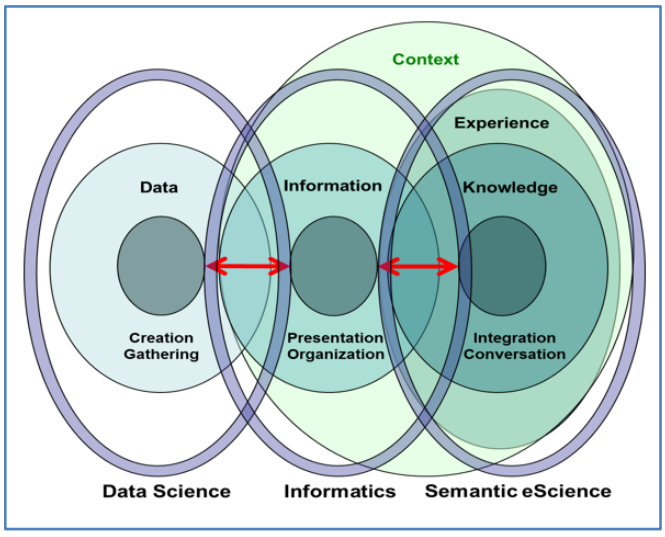
DHS made an unfortunate choice of words when they used the term "ecosystem" in *Enabling Distributed Security in Cyberspace – Building a Healthy and Resilient Cyber Ecosystem with Automated Collective Action.* The term "ecosystem" is not used in Joint Doctrine. By using the term "ecosystem" DHS invited the divergence of the DHS lexicon from the rest of the government. In JP 2-01.3 *Joint Intelligence Preparation of the Operational Environment* the term "Holistic View of the Operational Environment" is used to describe the comprehensive Operational Environment.

3. "Semantic eScience, the web science and technology stack, is NOT a replacement or competition for the Joint Doctrine. In fact, it should fully support it from a technology standpoint, depending of course on the data model and datasets being added." (Forum comment)

Semantic eScience is poorly articulated with Joint Doctrine in that Semantic eScience does not use the terminology or framework of Joint Doctrine. This observation is not to suggest that Semantic eScience is not a powerful tool for data analysis – only that Semantic eScience should be put into the context of Joint Doctrine.

Where does Semantic eScience fit in the Joint Doctrine scheme? Semantic eScience is an intelligence tool. As such, Semantic eScience must be consistent with JP 2-0 *Joint Intelligence*. By aligning *Science of Cybersecurity Developing Scientific Foundations for the Operational Cybersecurity Ecosystem* with Joint Doctrine, the Joint Doctrine data-information-intelligence model will replace the corresponding model in *Science of Cybersecurity Developing Scientific Foundations for the Operational Cybersecurity Ecosystem*.

Semantic eScience



JP 2-0, *Joint Intelligence*

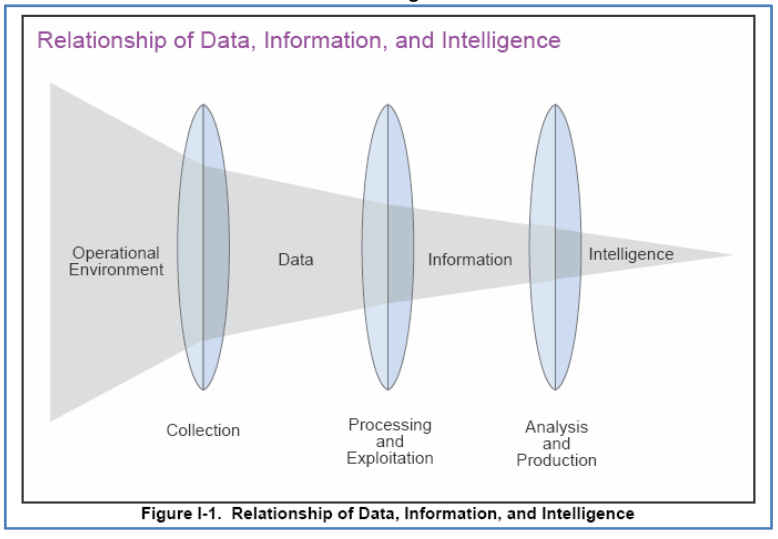
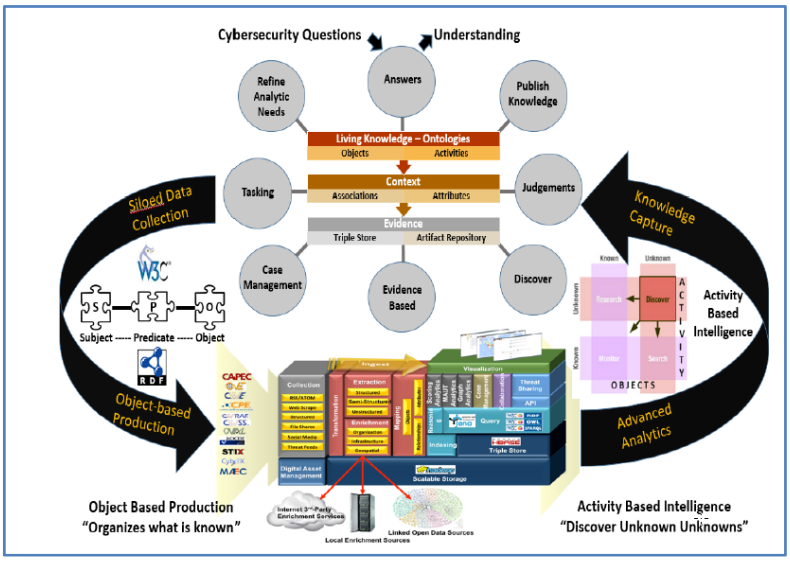


Figure I-1. Relationship of Data, Information, and Intelligence

Similarly, the Principles of Joint Intelligence, including the Attributes of Intelligence Excellence, should be used as the measure of effectiveness, replacing the concepts of Measurable Security and Human Factors in *Science of Cybersecurity Developing Scientific Foundations for the Operational Cybersecurity Ecosystem.*

Finally, the data-centric Semantic eScience ecosystem must be subsumed into the Joint Doctrine's Holistic View of the Operational Environment.

Semantic eScience



JP 2 01.3 *Joint Intelligence Preparation of the Operational Environment*

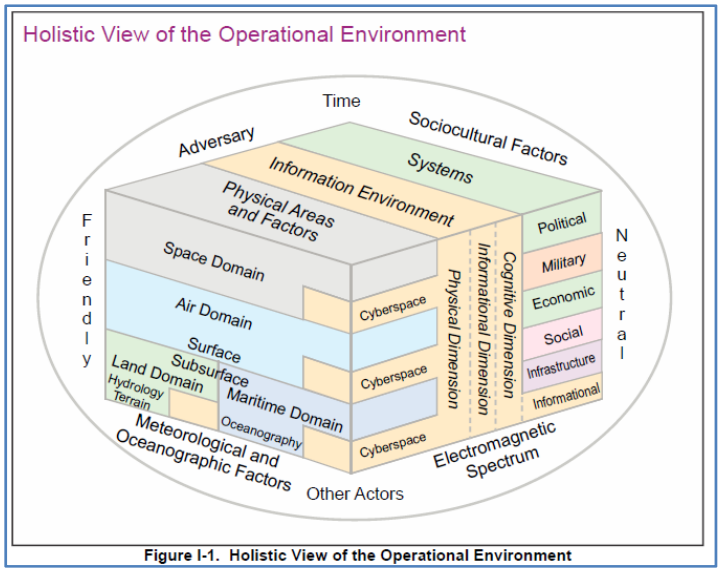


Figure I-1. Holistic View of the Operational Environment

When cybersecurity is viewed as part of the Holistic View of the Operational Environment, non-data processing solutions to cybersecurity problems can be found. For example, solutions to the growing scourge of email impersonation fraud (https://www.fbi.gov/news/stories/2015/august/business-e-mail-compromise) can be found by changing business processes (https://www.wellsfargo.com/com/fraud/fraud-schemes), not in data processing systems.

After aligning Semantic eScience with Joint Doctrine, if Semantic eScience is validated as improving cybersecurity as determined using the Principles of Joint Intelligence, then Semantic eScience should be added as an Analytic Tool in JP 2-01, *Joint and National Intelligence Support to Military Operations*.