The Cross-domain Information Exchange Framework (CIEF)

Paul Shaw, SPAWARSYSCOM, Paul.Shaw@navy.mil
Dr. David J. Roberts, iBASEt, Inc., droberts@ibaset.com

Abstract

The Cross-domain Information Exchange Framework (CIEF) is an architectural framework designed to support critical information exchange to assist or automate DoD (Department of Defense) mission oriented tasks. It is also an operational design for the publication, location, and subscription to information in the correct mission context and monitor the operational use of information in that context.

1. Introduction

The Cross-domain Information Exchange Framework (CIEF) is the result of the analysis of current Information Exchange Data Models (IEDM) and semantic methodologies used in academic, government, and industry efforts to support Service Oriented Architectures (SOA).

An underlying premise of CIEF is that present web services and Internet registry approaches do not satisfy Joint/DoD information exchange requirements. These commercial attempts are so limited in their service descriptions that the requestor would need to know much more about a service to be able to use it in any meaningful way. Additionally, commercial SOA and associated web services do not place in high regard DoD requirements for assured delivery, timeliness, prioritization of information, and many mission derived requirements.

CIEF research has determined a need for a SOA infrastructure that supports missions, corresponding actions, and that has contextual references (who, when, why, etc.) and is based on the information exchange requirements of the DoD mission. Further, this new infrastructure must be vendor agnostic with regard to the applications or services being supported. Be it a Microsoft SharePoint® COI, a Lotus Notes® group, or any Community of Interest (COI), the core requirement for interoperability with currently deployed enterprise collaborative and information management applications must be supported.

2. Background

Service Oriented Architectures (SOA) have followed a path of three major stages: hardware level integration, system level integration, and currently are evolving into resource based services. To the requesting client, these stages have been largely transparent with the occasional denied request of "out of range" information. It is in this evolving environment of SOA the co-partner of the World-wide Web, or Internet has also matured. In fact the early Internet (Web 1.0) was based on the connection of physical devices using connectionless protocols and very clever backbone addressing scheme that "knew" the location of potential information resources. Add indexing engines and Web 1.0 was more than a useful evolutionary step.

We are current probing the functionality of Web 2.0 Internet application that largely connects information resources in useful processes. From buying tickets, and booking a hotel, and securing a rental car in one easy transaction, to more sophisticated business to business transactions that simplify the purchasing and arrival of inventory, Web 2.0 can be seen as another level of information integration and services built around the performance of tasks.

3. The Problem

While it is true that isolated pockets of Web 2.0 information exchange do not exist, these pockets are typically hand crafted, specialized enclaves of well understood functionality and purpose. Within these specialized knowledge domains, such as the battlefield information exchange, Web 2.0 processes and supporting lexicons do exist. For example, the NATO Generic Hub, Version 5 (GH5) offers a concrete example of a clear and unambiguous information exchange to build and maintain a battlefield operational picture based on dynamic information flow. The NATO GH5, though, represents decades of

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
21 MAY 2008	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
The Cross-domain Information Exchange Framework (CIEF)	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **SPAWARSYSCOM, San Diego, CA**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
AFCEA-GMU C4I Center Symposium "Critical Issues In C4I" 20-21 May 2008, George Mason University, Fairfax, Virginia Campus, The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**UU**	**66**	

understanding of the requirements and processes necessary to maintain a common battlefield operational picture. Note: This major failing of Web 2.0 applications of their inability to scale beyond their reference domain will be addressed later in this paper.

What is apparent in environments like the NATO GH5 is the need for a commonality of purpose and intent. The focus of such efforts is the automation or assistance of mission tasking. It is not enough to focus on information exchange boundaries; rather the "purposeful intent" of the information exchange must be well understood. Additionally, the underlying purpose of the information exchange is often abstracted from raw information into more focused forms [1].

It is this context of mixed intent and meanings that underlies the problems associated with the cross-domain exchange of information. For information to "make sense," common meanings (lexicon), formats (syntax), and "purposeful intent" must be part of a common exchange process. Most of us understand the meaning and format part of the information exchange equation, but purposeful intent is a much more subtle variable.

For example, a simple question such as, "How are you?" could invoke a variety of responses based on both the context of the question and the intended response.

"I am fine," is one typical response, if the query was meant as a greeting.

"I had a fever of one hundred and two last night," could be another response to a doctor's query.

"I am waiting for a friend," could be yet another response to a stranger in a bar.

It should be obvious by now that the Web 2.0 access road to a SOA highway of understanding and cooperative services must deal with the problem of describing the purpose of information. It is in this context of task and mission orientation that the Cross-domain Information Exchange Framework (CIEF) was proposed by the authors.

4. CIEF Fundamentals

Dictionary.com defines "intent" as: 1.) the state of a person's mind that directs his or her actions toward a specific object, or 2.) meaning or significance. Further research of "intent" leads us to the *Online Etymology Dictionary's* definition of: "purpose," c.1225, from O.Fr. entente, from L.L. intentus "attention".

It is this "purposeful intent" within a mission task oriented profile that is fundamental to CIEF, to include:

a. Information is owned to the extent that is attributed to a source and access controls. While the disclosure of information to unauthorized individuals is to be avoided, less obvious may be the governance surrounding information from personal privacy issues to distribution restrictions (e.g., Defense Technology Accessibility Codes (DTAC)) to business rules. To ignore any of these areas is to "break trust" or ignore policy and governance.

b. The contextual reference of information must be understood. Temporal (when) and spatial (where) constraints provide the basis for higher level concepts such as trust, reliability, and applicability. Information with its associated provenance and context should be part of a common core of understanding.

c. The level of encapsulation of information must be understood. From raw sensor data to detailed action plans, the level of detail that is exposed must be part of the exchange process.

5. The Playing Field

CIEF was designed to assist in the performance of mission oriented task or to completely automate those tasks. It is within that matrix of person and machine interactions that various types of tasks are performed:

From: \ To:	Person	Machine
Person	**Social Engineering / Business Rules** • Establish Trust • Team Building • Value Ranking	**Publish Information** • Normalize Data • Categorization • Key Indicators
Machine	**Subscribe to Information** • Assisted Search • Pattern Association • Alerts	**Compute / Transform** • Voluminous Data • Rule-based • Repetitive Actions

Figure 5. Information Flow

The above matrix can be thought of as a roadmap of human / machine interaction, with the initiation of

person to person information core to human society. CIEF supports person to person information exchange based on the establishing of trust relationships (social engineering) through an understanding of the credentials of participants, organizational placement, and some assertion of reliability.

Person to machine information publication or "push" of information to unknown persons is guiding principle of Internet news feeds and a well reasoned publication method. CIEF "pushed" information, though, is dependent on the concept of a well described mission-oriented context and well understood interest groups.

The corollary of machine to person information exchange is person to machine exchanges, or the subscription to information. This could also be described as the intelligent "pull" or assembly of information from selected contexts. Within this matrix area, CIEF will address information encapsulation, or the representation of information at the correct level of abstraction. For example, if the request for information is, "Are you ready?" and the response is, "Six." That may be a correct response, if "Six" has a context-based meaning (e.g., scale of one to ten) at a shared level of abstraction.

With the thousands of potential sources of information and hundreds of millions of pieces of information, it is obvious that the "heavy lifting" of machine to machine (no human intervention) is a core requirement of the CIEF architecture. The searching, sorting, merging, pattern recognition, detection, alerts, and all the tasks required to assemble information into meaningful collections will require a level of machine autonomous processing in the CIEF architecture... unfortunately, machines to not resolve ambiguity well. For the "heavy lifting" of information processing and analysis to occur via machine to machine, an exquisite (or even painful) level of detail is often required. For example, and simple command such as, "Turn the light on when people enter the room," seems like a computable task, but, in fact, a great deal of specificity is missing.

- Do I turn the light on when the second person enters the room (noting the instruction did specific a plural context)?
- What constitutes "entering?"… When the forward edge of the body enters? Center of a body?
- What if the lights are already on?

And the list of "petty questions" goes on…

Similarly, what would pass as a simple mission associated task may have enough specificity to be understood by humans, but not computable by machines. Thus, an early focus area of the CIEF project will be the development of a core lexicon and minimal sets of patterns (syntax) to address the level of specificity required for machine processing of information.

6. The Requirements for Web 3.0

To move from the negotiate context of Web 2.0 services to a Web 3.0 service environment where information is self-describing and ready for consumption, clearly needed are operational requirements to determine completeness and monitoring of progress. An initial list of such requirements is presented in the following sub-sections:

6.1 Multi-directional Information Flow

An environment that is based on the semantic mitigation and integration of information must support the flow of information from many sources and potentially feedback to those sources. In this environment the lines between consumers of information and the producers of information can become blurred or even co-dependent. For example, a continuous feed of incorrect information might require feedback to the source to correct a filtering attribute. In such an information exchange, the source and sink of information could be an ongoing arbitration.

6.2 Common Process for Publication and Subscription

Travelers quickly learn the value (or lack of) translation dictionaries. It is not enough to translate our words into foreign words. Some process to address our context, references, and intent to the target language is necessary to truly impart meaning. Some common ground of universal context must be the starting point for cross-domain information exchange and part of a Web 3.0 specification.

6.3 Manage Locally and Use Globally

Web 2.0's major contribution was the application of business rules and shared context in transactional information exchanges. And while a major tenant of Web 3.0 is openness, or self-describing information, the lessons of Web 2.0 cannot be left behind. Information will continued to be owned and managed by the sources of that information, and to ignore this simple fact is to invite chaos. That is not to say that information publishing standards in Web 3.0 are not open, but those standards must accommodate business rules, privacy concerns, and security levels (need to know). Web 3.0 is not a license to throw away good business processes, but rather, present new methods of incentivizing good behaviors. (Note: This will be addressed later in this paper the Activity Tracker discussion.)

6.4 Dynamic Reconfiguration of Constituent Parts

If Web 3.0 is to scale to multiple information domains and millions of information sources, the core design must accommodate frequent and substantive change. Design methodologies such as On Demand Schema (ODS) [2] popularized by the authors of this paper represent one such core architecture that grows contextual associations as a corollary to addition of content and new rules.

7. The Missing Pieces of Web 3.0

It is obvious that Web 3.0 will "live" on the same operational backbone as the current Internet (or Global Information Grid (GIG) for the Department of Defense (DoD)). What will distinguish Web 3.0 services will be the ease in which information is published, located, and consumed. The following sub-sections address specific pieces that must be part of Web 3.0:

7.1 Partitional Architecture

It is a natural human tenancy to clustering information in hierarchical ontologies. Most information clustering schemes from the Dewy Decimal System at the library to the DoD Metadata Registry (DMR) have been successful in using hierarchical decompositions of information. The success of such schemes is based largely on single dimension decompositions, such as topic areas, or mission focus areas. Additionally, hierarchical schemes facilitate data management functions such as: error correct, duplication identification, etc.

The structural strengths of hierarchical information management (single dimension) are also its weakness. In environments where information must be viewed from multiple dimensions of interest (e.g., task, time, location) a partitional architecture is more appropriate.

In a partitional ontology information is clustered in sets that have common attributes. For example, vehicles could have tracks or wheels, and they could also be military or civilian, and they could also be armored or unarmored. A vehicle that is wheeled, civilian, and unarmored could easily be represented as a set union of three common attributes. The same entity could not be represented in a hierarchical ontology.

7.2 Operational Definitions of Time and Location

Past definitions of time and location are inadequate to describe the operation context of tasks in a Web 3.0 environment. From a fixed representation of time we must take a page from quantum mechanics and understand that we must be able to describe the movement of time relevant to location and events. A working example of this requirement is the production control concept in Just in Time (JIT) inventory control. It is not enough to know where an inventory item is located. We must also know when it is needed and what time events must occur for the part to meet its appointed assembly time. The following Entity Relationship diagram addresses a notational representation of time:

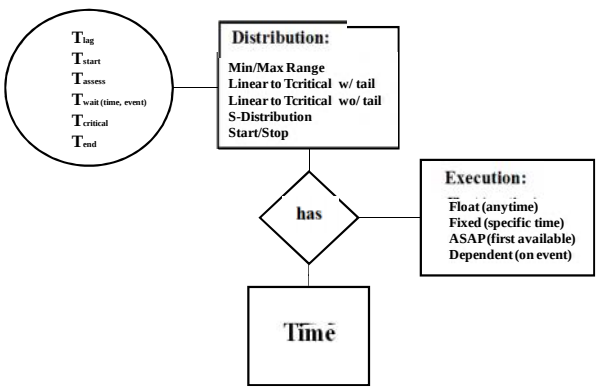


Figure 7.2.1 Time Utilization

The previous ER diagram could be used to define temporal constrains that bound a mission task. Event context such as urgency or how event resources could be distributed could be addressed.

Similarly location information could be used to define the spatial context of information as it relates to time:

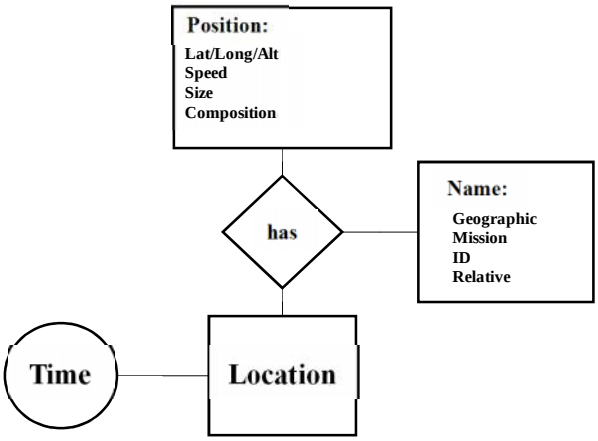


Figure 7.2.2 Location Context

While both figures are notional, they do address the requirements of tasks defined within the context of time and location.

7.3 Publication and Subscription Tools

Either file it or pile it are two methods for information storage. Either take the time to classify, index, and store information using a known ontology, or throw the information into a heap and later use heuristics to "figure out" information relationships.

While both methods have their proponents, common sense would dictate that the closer the classification of information to the source (i.e., subject matter expert), the more accurate the classification of information. While heuristic analysis may be the only options for unstructured or disparate information, again, common sense would suggest the development of publication and subscription tools to assist in the "tagging" of information using shared ontologies for Web 3.0 services.

7.4 The Right People Doing the Right Jobs

Web 3.0 should not be considered the intellectual property of any single skill domain. Rather, Web 3.0 should be viewed as a collaborative effort of task level subject matter experts, semantic engineers, database designers, communications specialist, and the list should continue to include multiple disciplines.

Additionally, the effort should be partitioned along the lines of complementing skills. For example, the development Entity Relationship (ER) diagrams to address the association of information clusters could be a collaboration of domain subject matter experts (define the process), cognitive engineers (patterns and ontologies), database designers (data structures), communications specialist (connectivity to the data), policy specialist (business rules), security specialist (governance), legal specialist (privacy and law), and other "to be named" specialists that could impact the efficacy of the ER diagrams.

8. The How's of CIEF

Although much of the theoretical underpinning of CIEF is based in semantic technologies (classification and inferencing), the mechanics of CIEF are based on a web-based registry service model that profiles tasks and requests information that matches those requests to the appropriate information source [3]:

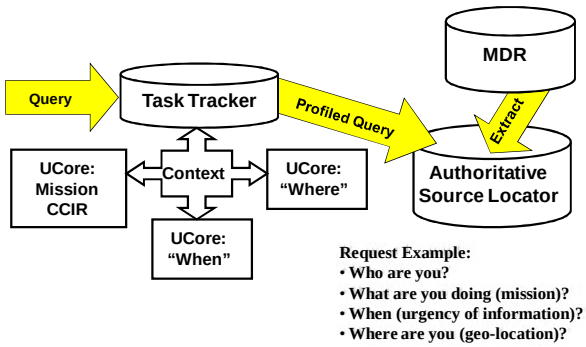


Figure 8.1 Query Profiled

In the query profiling process the ontology of the mission tasking, and the time and location context are established based on a shared Universal Core (UCore) lexicon and syntax (Note: UCore is a parallel effort to CIEF). In the illustrated example the query is submitted to an Authoritative Source Locator (ASL) web service that offers content and connection metadata has been extracted from the DoD Metadata Registry (MDR). (Note: CIEF is not limited to any single source for information content or location.)

Once the context who, what, when, and where are established, a reasonable assertion of intent can be inferred [4]. With an understanding of the intent of the information request, with appropriate security

considerations, the appropriate time and location based information sources can be indicated:

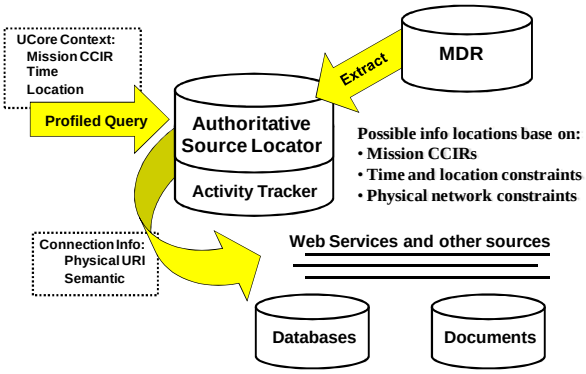


Figure 8.2 Query Mediation

To complete the requirement of the "correct information at the correct time" the CIEF Activity Tracker will provide metrics and best source locations based on historical and dynamic network assessments.

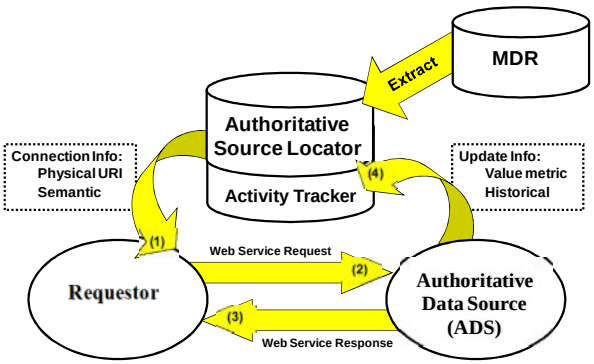


Figure 8.3 Query Response

Additionally, the CIEF architecture is not limited to purely SOA based information exchanges, but will also support both access and insertions in traditional data repositories, document system, and newer Internet methodologies such as New Feeds, and ultimately, an "intelligent" service bus [5].

9. Now and the Future

CIEF is both an architecture and an operational framework, with an implementation plan. The plan is in its initial phase of business case analysis, use case analysis, schema design, and deployment strategies. CIEF was accepted in 2007 by the GIGlite.org, sponsored by the World-Wide Consortium for the Grid (W2COG.org), as its operational framework, and is currently being evaluated for implementation by several DoD organizations.

Key to deploying an architecture such as CIEF is some commonality in the publication and consumption of information. In that regard, CIEF will leverage existing the National Information Exchange Model (NIEM), current mission centric Information Exchange Data Models (IEDM), DON Universal Core (UCore) models, and current mission-level task analysis.

In addition to the near-term deployment of CIEF, additional features and capabilities are planned, such as: semantic cluster analysis tools, information visualization software, and other information publication and subscription tools. With integration of SOA and semantic technologies CIEF is clearly on the path to deploying a Web 3.0 environment for DoD and the larger Internet community.

References

[1] Rand Corp., "Out of the Ordinary", Rand Monograph, 2004.

[2] Roberts, David J., "A Practical Implementation of Semantic Technologies… the Virtual Knowledge Repository (VKR)", Semantic Web Applications for National Security Conference, sponsored by DARPA, Washington, D.C., May 2005.

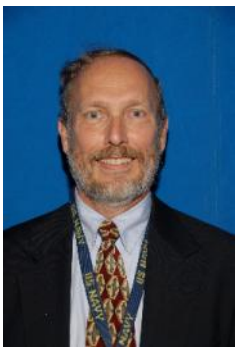
[3] Hayes-Roth, Fredrick, "Model-based Communication Networks and VIRT: Orders of Magnitude Better Information Superiority", IEEE MILCOM 2006 Proceedings.

[4] Hayes-Roth, Fredrick, "Two Theories of Process Design for Information Superiority: Smart Pull vs. Smart Push", Information Sciences Dept., Naval Postgraduate School October, 2005.

[5] Cox, Brad J., "Paving the Bare Spots: Towards an Enterprise-wide Defense Service Bus", White Paper of the Binary Group, Bethesda MD.

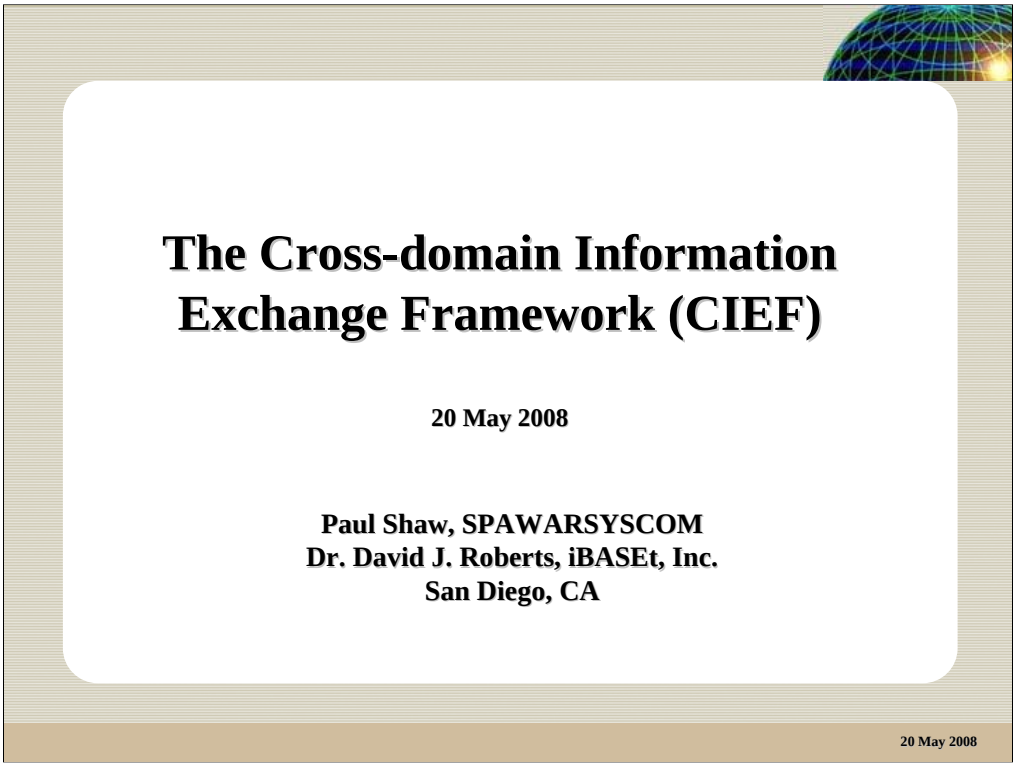
Paul Shaw is currently the Navy Functional Data Manager (FDM) for Command and Control (C2) and is based in San Diego at the Space and Naval Warfare Systems Command.

Mr. Shaw has spoken widely at conferences and seminars on semantic technologies and innovative approaches to data management within the DoD.



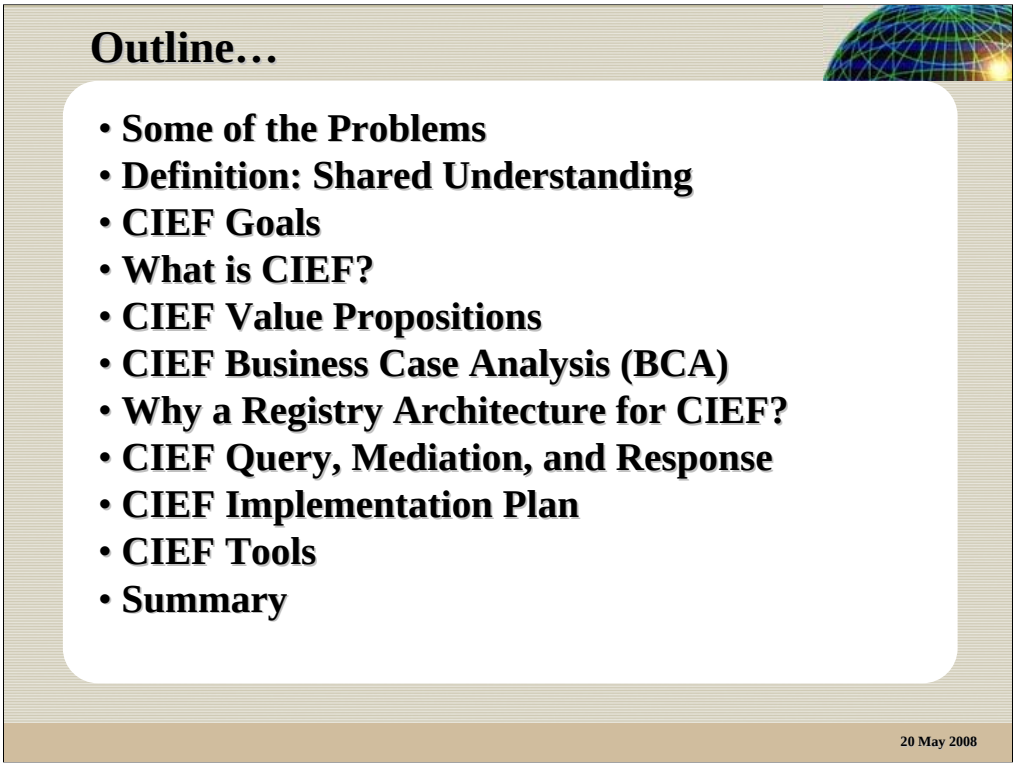
Dr. David J. Roberts is the Chief Scientist on the Cross-domain Information Exchange Framework (CIEF) project and also supports the SPAWARSYSCOM, San Diego.

Dr. Roberts has presented CIEF and other semantic based Information Management systems at numerous conferences and technical workshops.



The Cross-domain Information Exchange Framework (CIEF) is the result of an analysis of current Information Exchange Data Models (IEDM) and semantic methodologies used in academic, government, and industry... Both advantages and shortcomings.

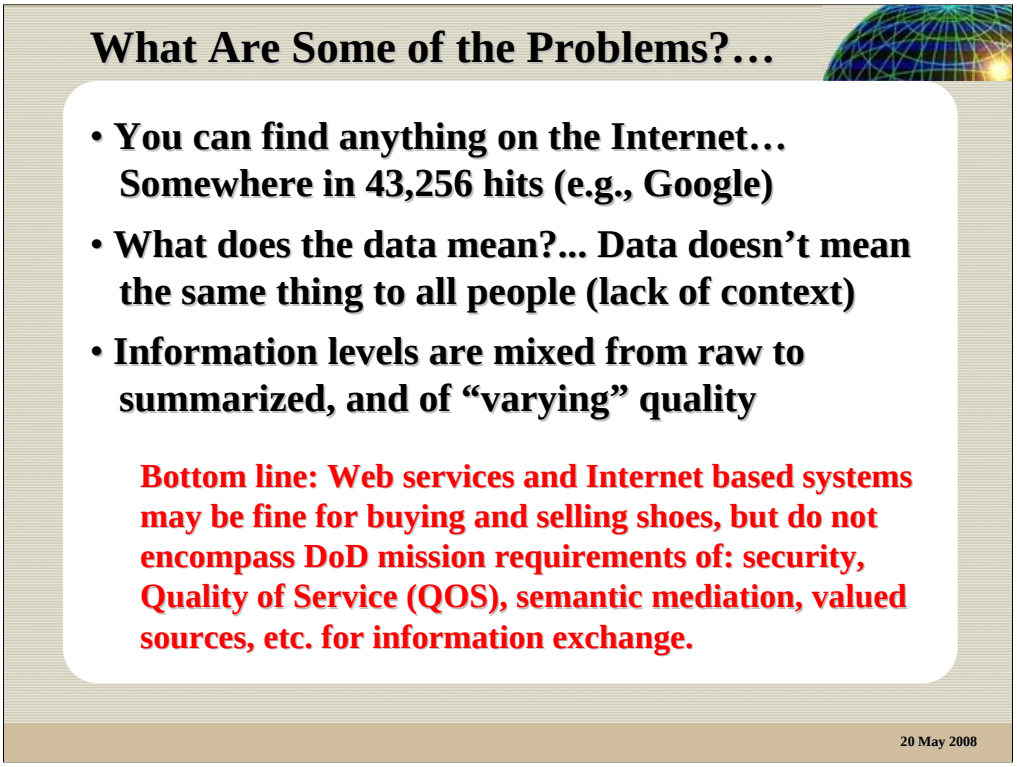
The following presentation should be viewed as both as an overview of research conducted at SPAWARSYSCOM, San Diego and a roadmap for the implementation of a shared core semantic objects and processes.



This brief will touch on the highlights of the Concept of Operation (ConOps) of CIEF and a high level overview of the implementation roadmap.

A major theme of this presentation is that a shared core of understanding is more than a simple lexicon or format. Rather, this brief will layout a process for building Communities of Interest (COI) based information exchanges on the foundation of existing IT systems, leverage powerful new commercial tools (i.e., 200,000 news feeds), and deploy new technologies that simplify very complex environments (i.e., Semantic Technologies).

Per the opening page, additional detailed materials are available, to include costing, staff requirements, and schedule information.



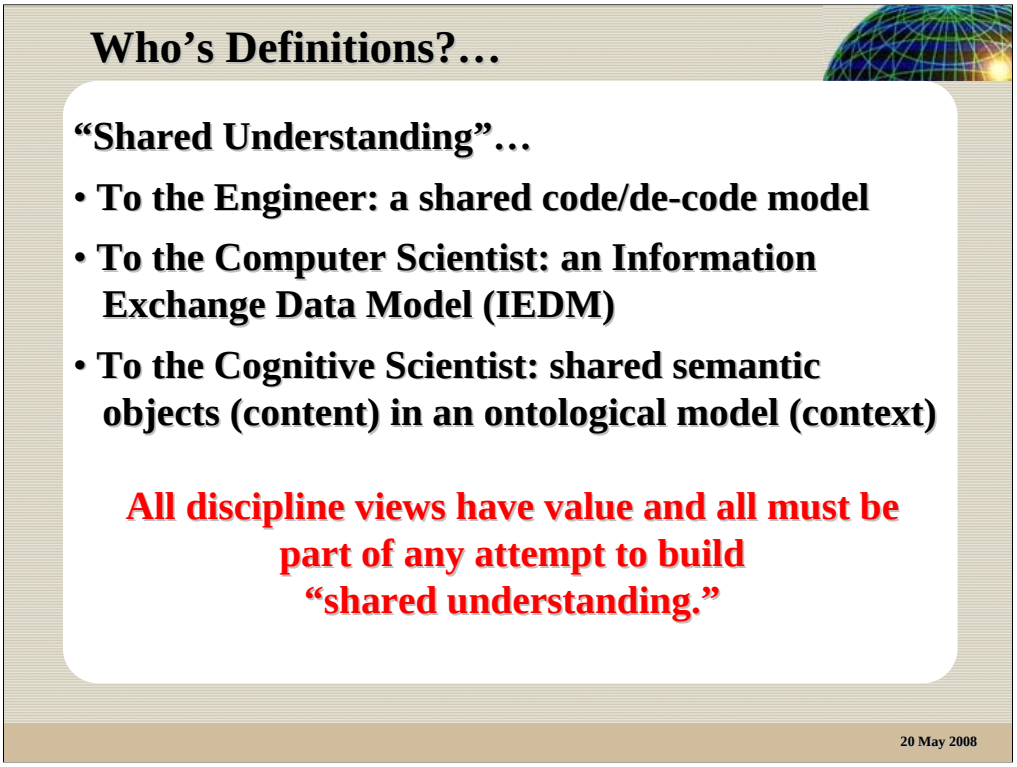
The core problem with current search and find methodologies lies at the roots of the Internet and DoD GIG… the open exchange of information by anyone that meets minimal format standards. There is little validation of posted information or correlation to any common semantic integrity.

In other words, information on the Internet is "all over the map" with regard to meaning, value, or even commonality of presentation style.

Even DoD sources and services lack common lexicons, formats, or even operate at the same level of detail. The assembly of raw data into useful information is still the work of analyst and subject matter experts (SME) to fuse, correlate, and "make sense" of all of it.

Additionally, commercial standards and expectations do not address key requirements of DoD missions or information needs.

Bottom line: The commercially oriented Internet is not going to solve many of DoD's information requirement. We (DoD) will have to solve our own problems.



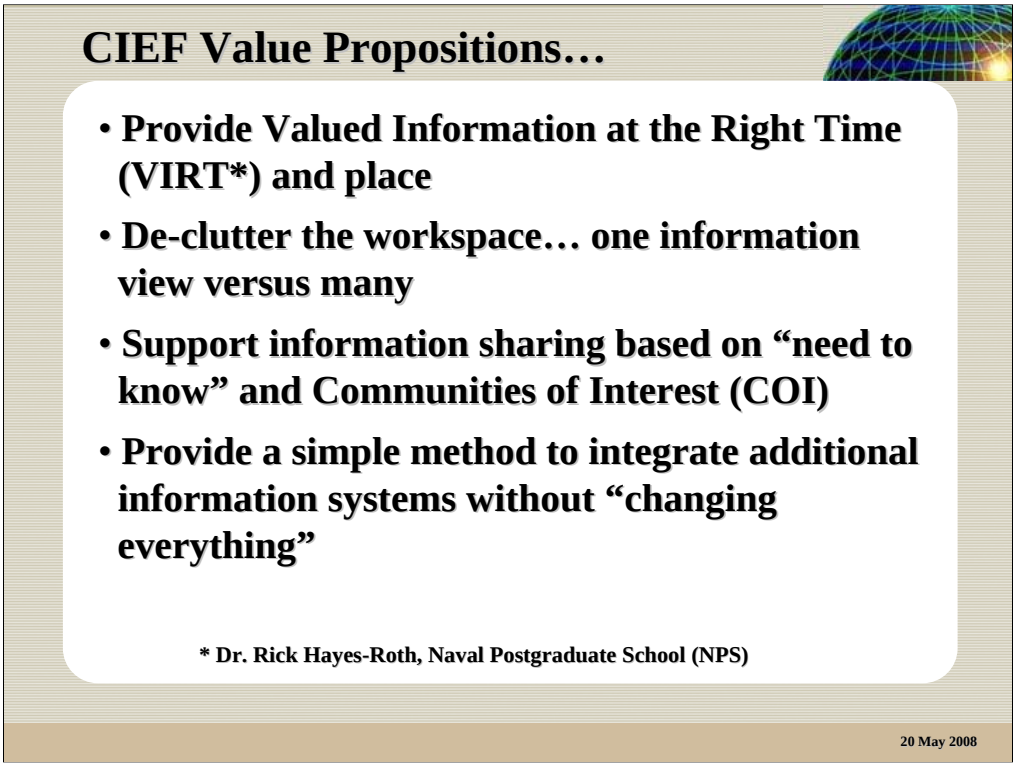
A core of shared understanding is a start of intelligent information exchange but…

An engineer logically will focus on the physical connectivity, sequence protocols, and perhaps more of a message format and report approach.

An IT professional will address schema definitions to include field and table specification, and process flow. List of values (LOVs) and other delimiting specifications will be considered.

The Cognitive Scientist may take a more "Ontological Web Language (OWL)" approach and build models of shared reasoning and understanding.

Bottom line: All approaches must be part of the solution from the physical parameters of connectivity, to a computable schema, to an understanding of the context of information exchange.

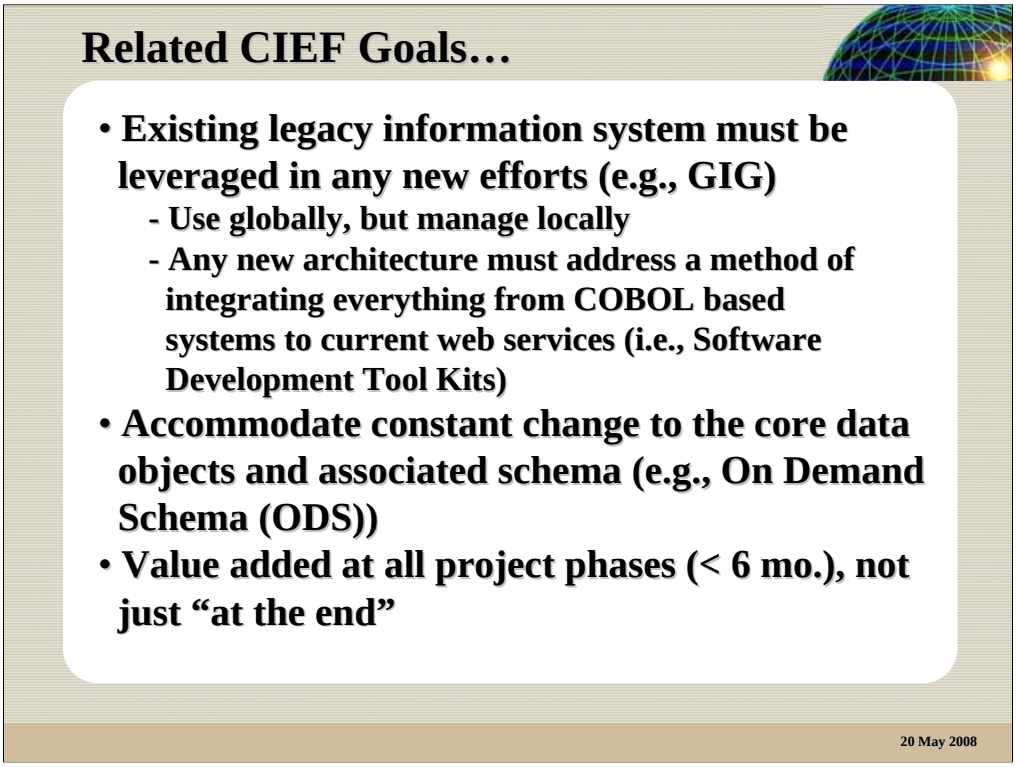


The first question that should be asked in any deployment of new or "better" technology is, "Why?" What problem is being solved? Why is it worth the investment of time and effort? At the end of the day, what is the payoff?

Correspondingly, we must determine what is the value propositions of the new and better technology, and define metrics to determine the degree to which we have met our stated goals.

Thus, our value proposition must be not only attainable goals, they must also be measurable.

Similar to the "real world" of shared understanding, CIEF must be robust enough to accommodate change… And that change could range from every minute to a more moderate pace of every day. Without a mechanism to integrate and synchronize change, CIEF will fail.

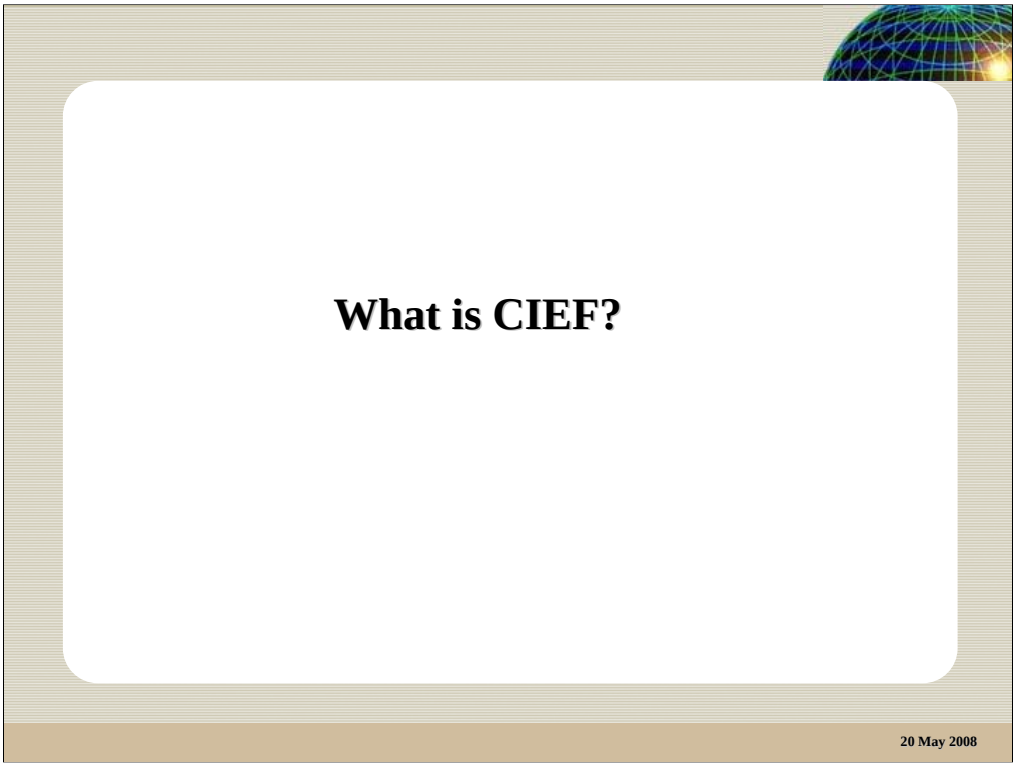


At least half of all IT systems in the world today are COBOL based. While that may be shocking to some, it is a testament to the value of past efforts and the mentality of not fixing things that are not broken.

Any new scheme of linking global information must incorporate legacy and existing IT systems. Additionally, the integration or sharing of information does not negate the ownership and control of that information. Sharing will not occur is some pollyannic environment of love and peace. Business rules and access controls must be part of any universal process for information exchange.

Any definition of a shared core of understanding must accommodate change management, daily use (version control), and the distribution of updates, corrections, and new concepts and terminology.

Short, well defined deployments are recommended to both validate "as you go" and to establish perceived value to the client base.



CIEF is universal framework for the exchange of all Global Information Grid (GIG) information, but it is also:

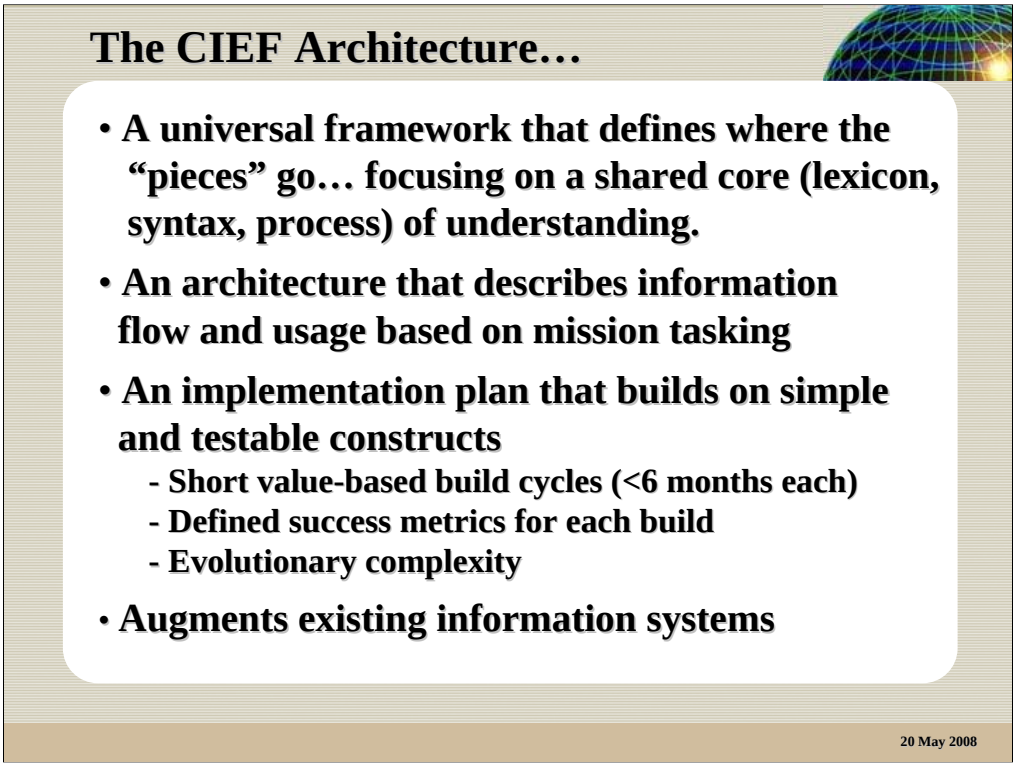
• An integrate methodology… that includes legacy and existing IT systems

• Based on natural language and biological models of information exchange

• A realistic philosophy of how to share information… based on business rules and DoD access controls

• An implementation plan… with clearly defined steps and milestones

But most importantly, it is achievable… the complexity of the effort builds in a stepwise deployment.



The Cross-domain Information Exchange Framework (CIEF) outlines a realistic process to build a core of shared lexicon, formats, and processes. CIEF also addresses real world problems in funding and managing large information integration efforts, and other focus areas such as:

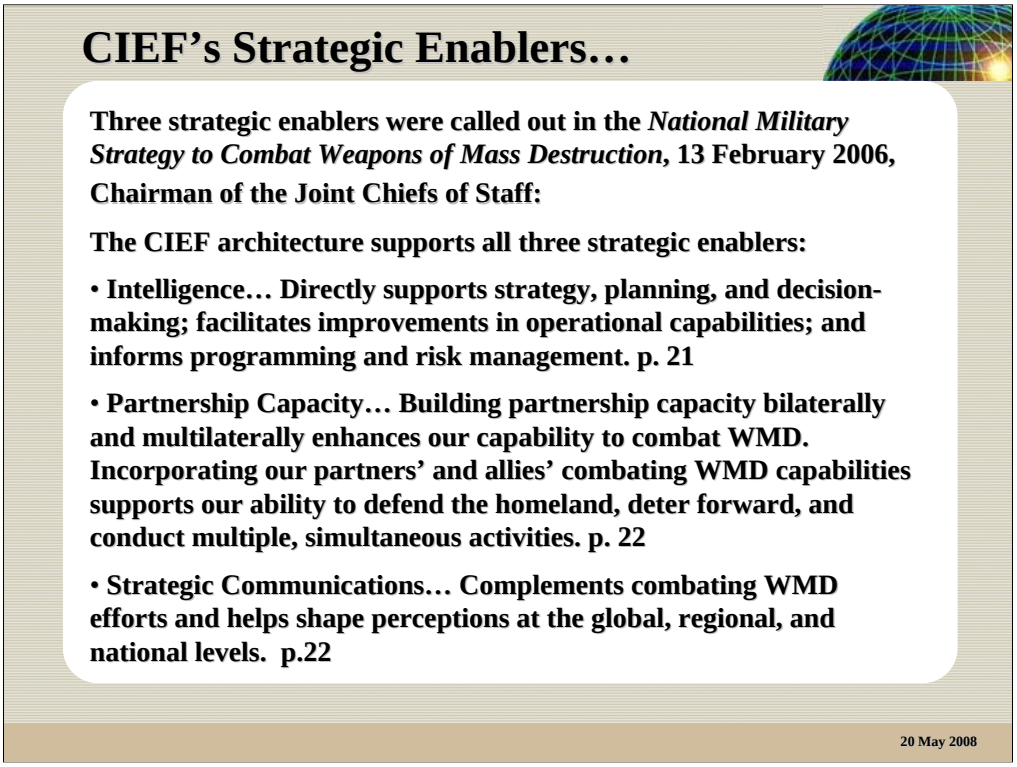
Where does the information come from? Legacy systems? Existing commercial web pages and sources?

How can "normal" people use this new semantic technology?

How can I "easily" integrate my information or system into this framework?

How do I get "credit" for providing for the common good?

Bottom line: CIEF takes into consideration many operational and "real world" integration and development issues through an implementation of the "business rules" that exist in operational environments.

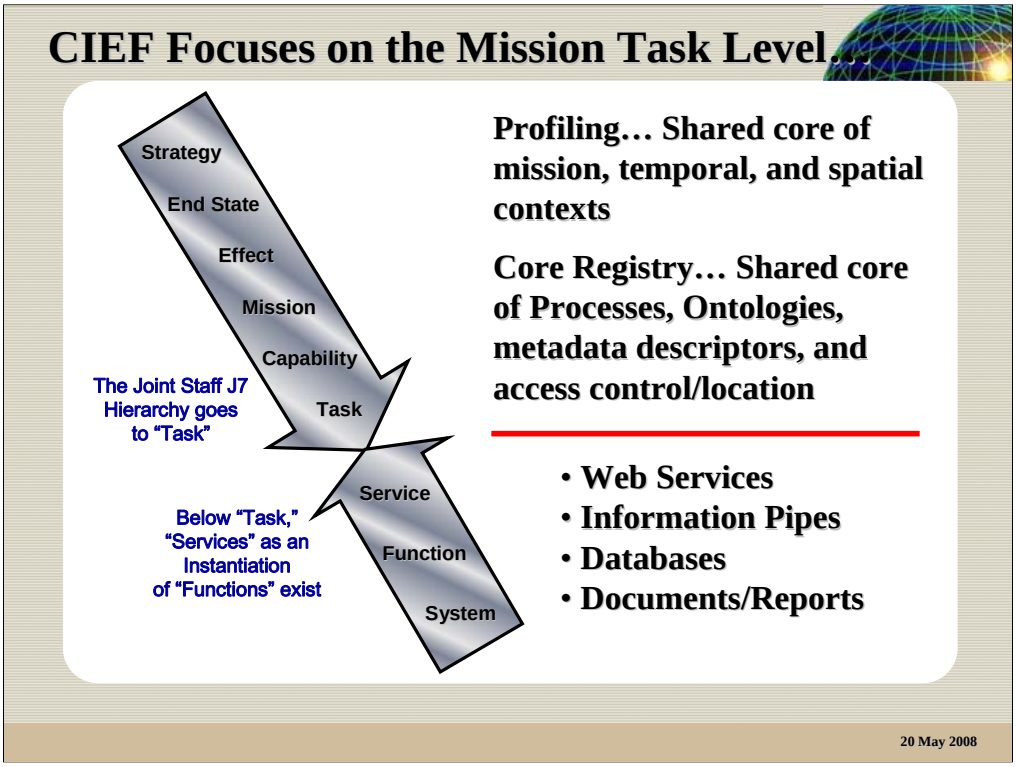


It is useful to think of new and evolving technologies as enablers… allowing us to solve complex problems that may have previously been "too hard".

In the cited JCS document, three areas of intelligence, partnership capacity, and strategic communications were called out as areas where IT efforts should be directed.

CIEF address and provides a logical implementation path for the cited enabling areas. For example, CIEF support not only the fusion of raw and sensor data, it also makes accommodations for more complex planning and decision making information.

Note: Per the previously cited Rand Monograph (Out of the Ordinary, 2004), CIEF was designed to support current DoD/Joint missions and associated tasking. CIEF was not designed in the abstract… rather, the focus was supporting DoD/Joint missions.

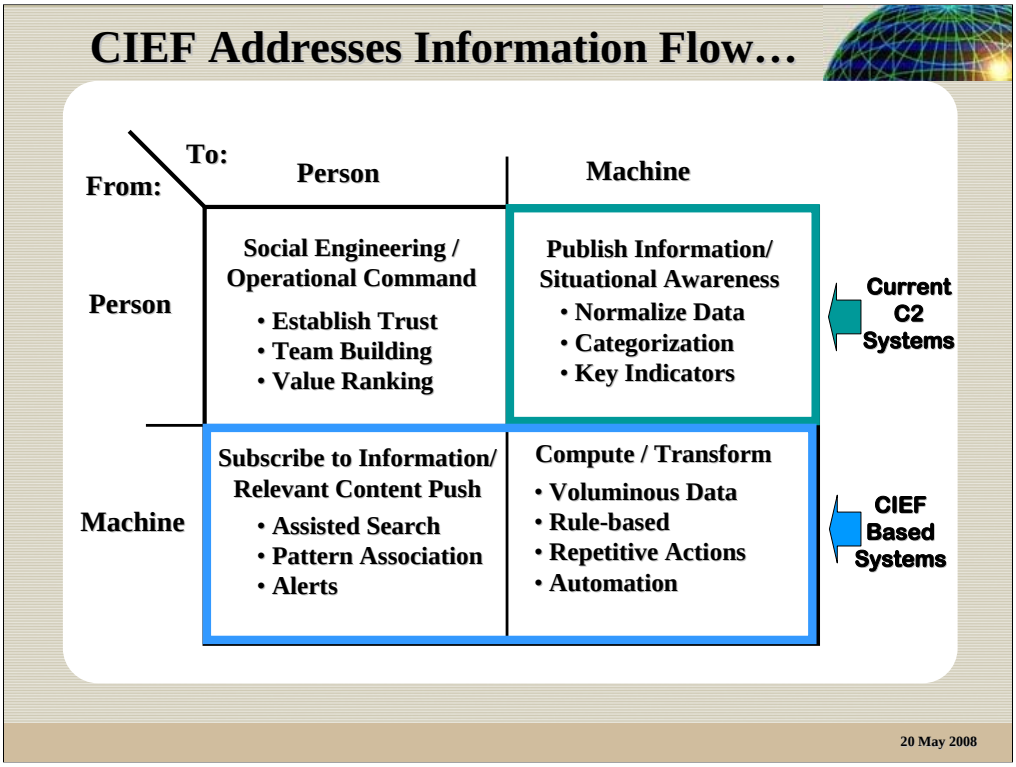


CIEF bridges the Service Oriented Architecture (SOA) gap between mission oriented tasks and relevant services and data sources.

In addition to providing a focused mission-based information exchange, semantic mediation is also provided to include:

- **Lexical translation (common "dictionary" of words and terms) to a shared core for a Community of Interest (COI)**
- **Syntax (format) transformation to a shared core**
- **Information level of abstraction (detail of information… raw, summarized)**
- **Security access control based on mission parameters and level of authorization (rule based access)**

Bottom line: CIEF acts as an intelligent filter to target information exchanges to mission tasking at the correct time and location.



CIEF will support the following information exchange methods:

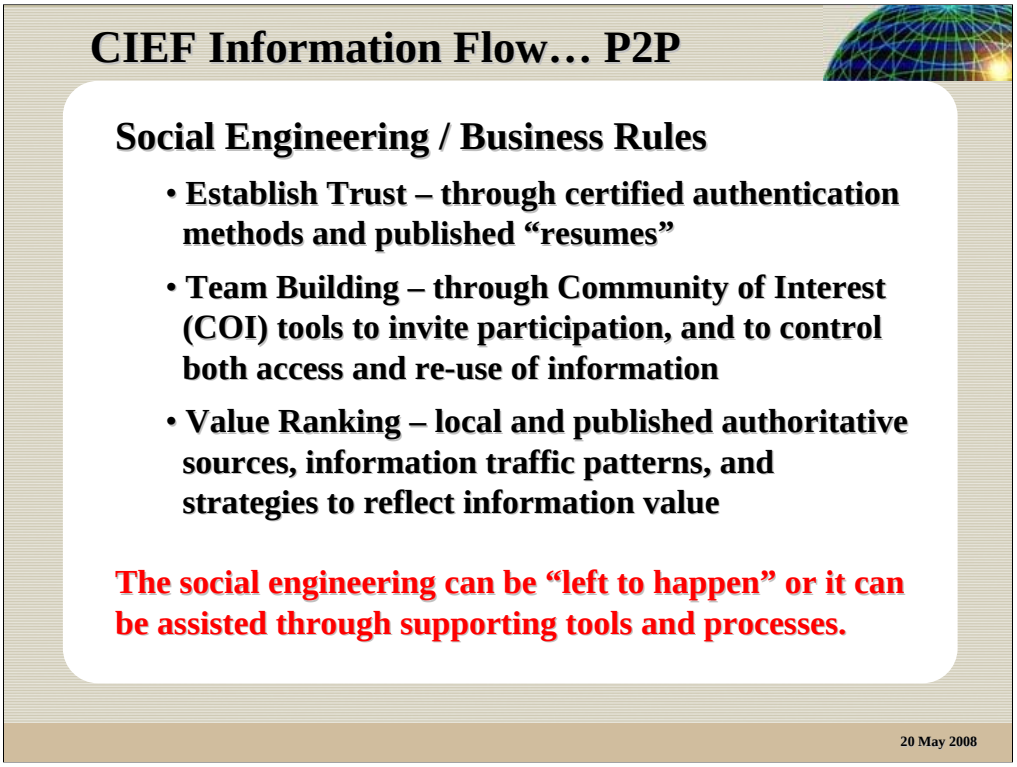
From basic information exchange: Person to person… social engineering (establish trust relationships)

To current technologies: Person to machine… publish information ("push" information to unknown persons, but known interest groups)

To the future technologies and architectures:

- Machine to person… subscribe to information ("pull" from selected context)

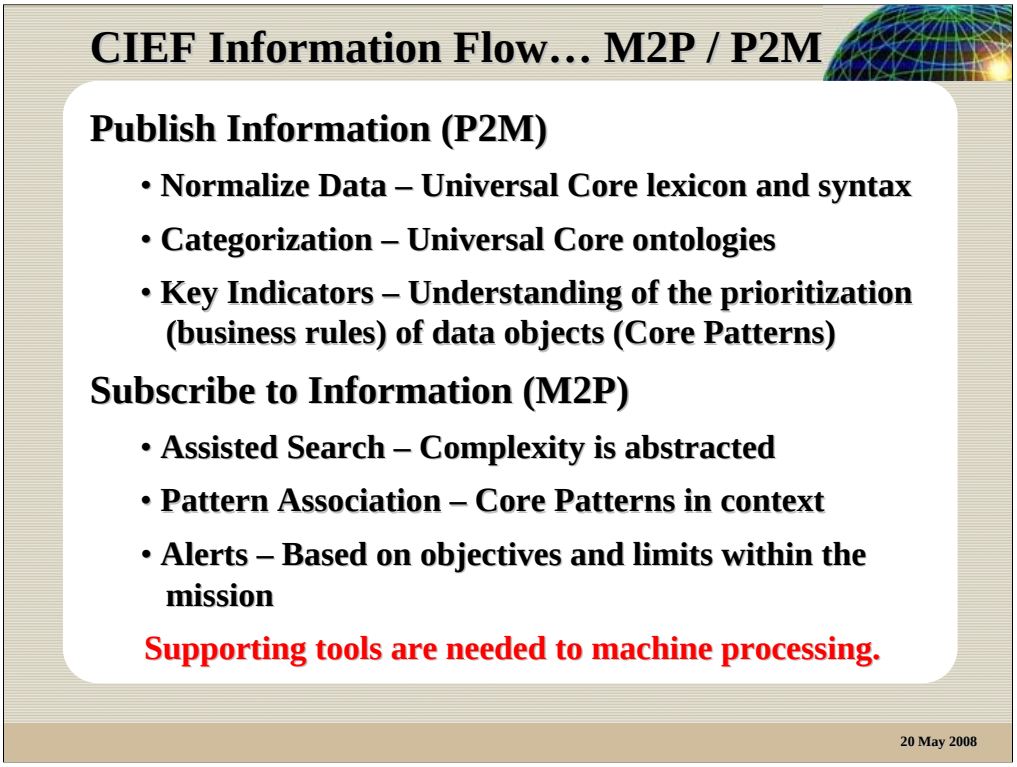
- Machine to machine… heavy lifting (search, sort, merge, pattern recognition, detection, alert)



Person to person… social engineering (establish trust relationships) through an understanding of the credentials of participants and some assertion of reliability. Trust is generally established through human bonds, or an established code (i.e., military chain of command).

Information is "controlled" and has value. Within a Community of Interest (COI) the rules for use and distribution of information must be established and codified. Software tools can assist in this process of maintaining established processes, and surfacing circumstances that are "out of bounds."

Understanding the value of information is key to managing it. Trusted or valued sources should be protected (replicated, fire-walled, etc.) to ensure their availability. Information exchange strategies should be built on the perceived value and corresponding flow of information (the purpose of the CIEF Activity Tracker to be detailed latter in this brief).



Person to machine… publish information ("push" information to unknown persons, but known interest groups). Note: This is a critical capabilities within the Intelligence community.

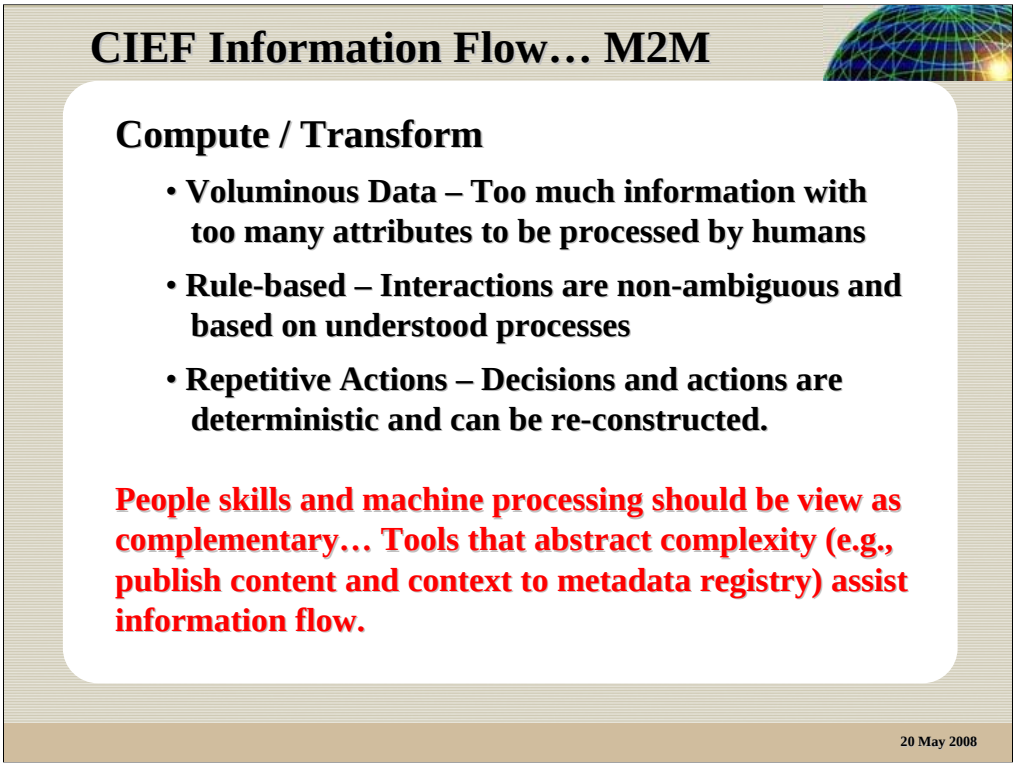
Machine to person… subscribe to information ("pull" from selected context)

Two major focus areas: 1.) the formalization of information semantics and syntax; and 2.) tools to assist in that process.

A major inhibitor in the formalization of information is that the subject matter experts (SME) are not cognitive scientist

(Note: They may be rocket scientist).

Tools that can be used by "normal" people are required to assist in the codification and categorization of information.



Machine to machine… heavy lifting (search, sort, merge, pattern recognition, detection, alert)

Machines to not handle ambiguity well. For the "heavy lifting" of information processing and analysis to occur via M2M, an exquisite (or painful) level of detail is required.

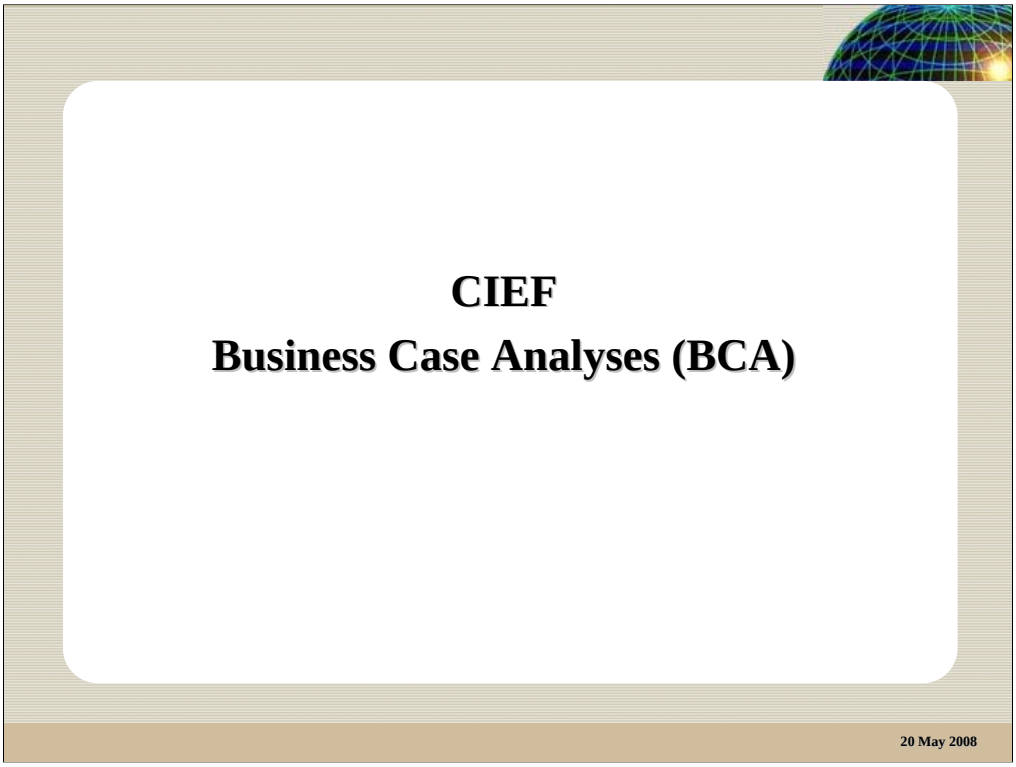
Example: "Turn the light on when the people enter the room," seems like a computable task, but, in fact, a great deal of specificity is missing.

- Do I turn the light on when the second person enters the room (noting the instruction did specific a plural context)?

- What constitutes "entering"… When the forward edge of the body enters? Center of a body?

- What if the lights are already on?

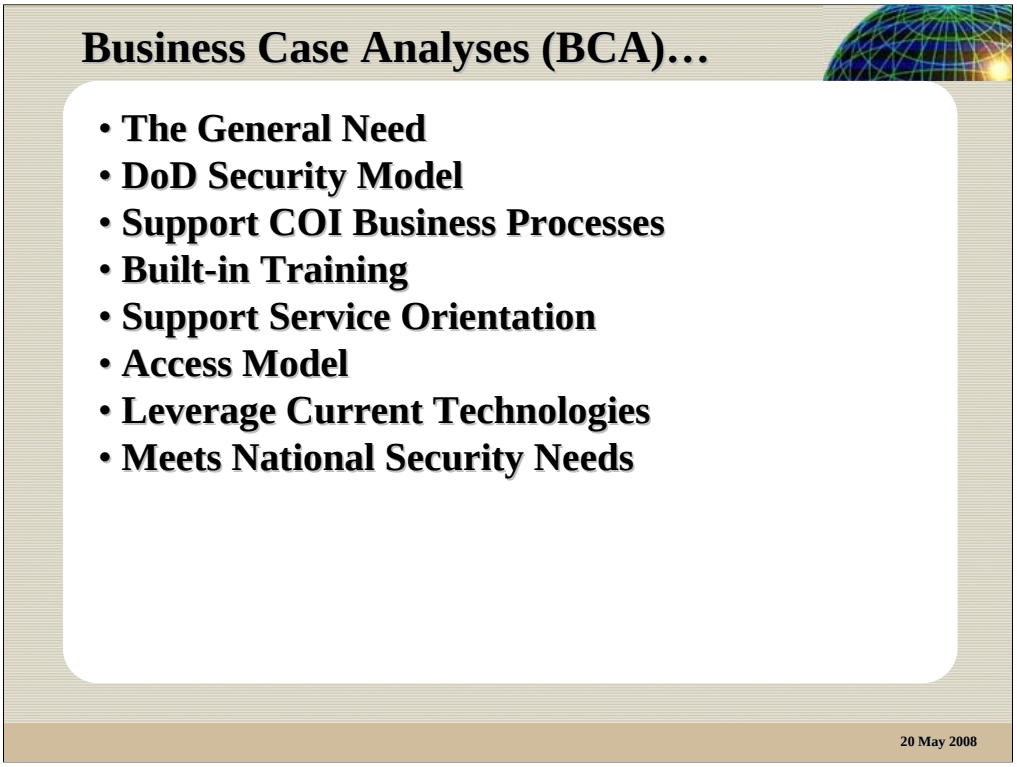
And the list of "petty questions" goes on…



The following Business Case Analyses (BCA) address obvious areas of value in the deployment of CIEF. Secondary, or more subtle areas of value were not addressed but should be considered in the deployment schedule. For example, what is good information in a timely manner worth? What is the impact of no information on a suspected terrorist?

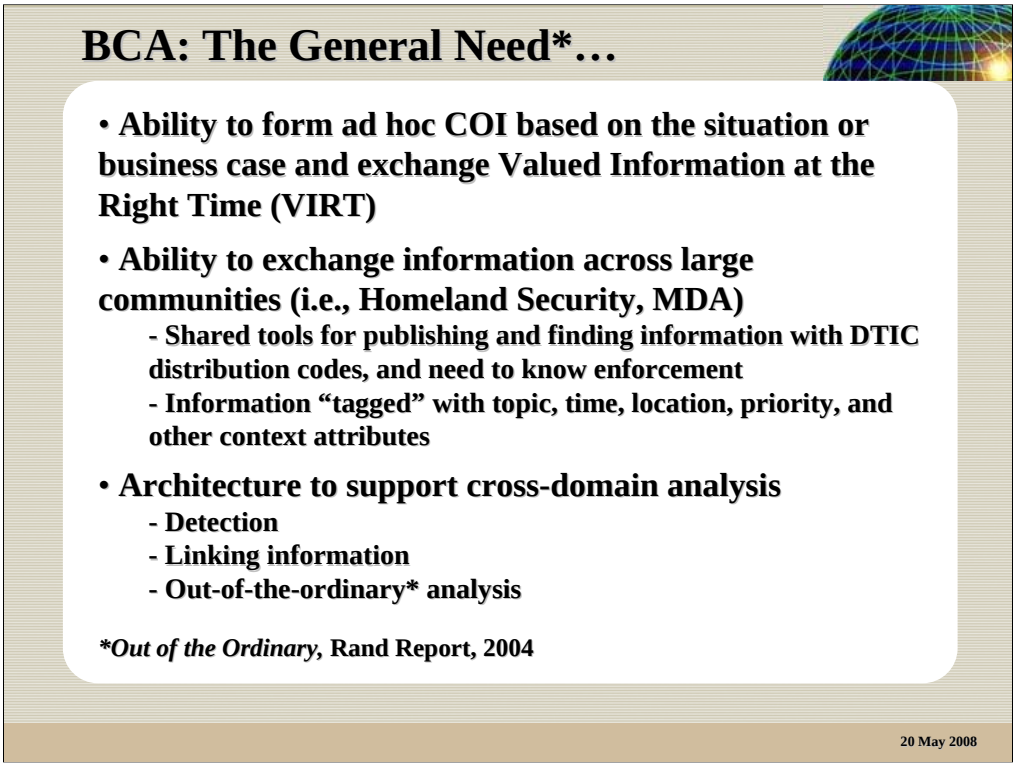
CIEF will not only aid in the integration of disparate information, it will also assist in real data strategies that address: the value of information (usage), predictive value (accumulated histories), support dynamic access control to information based on context, and other "intelligent" information exchange processes.

Note: CIEF's underpinnings are semantic technologies and biological systems that reflect "real world" information use and distribution.



Per the previous slide, the BCA presented are literally the tip of the iceberg, but should be sufficient to address concerns over the ROI of the effort.

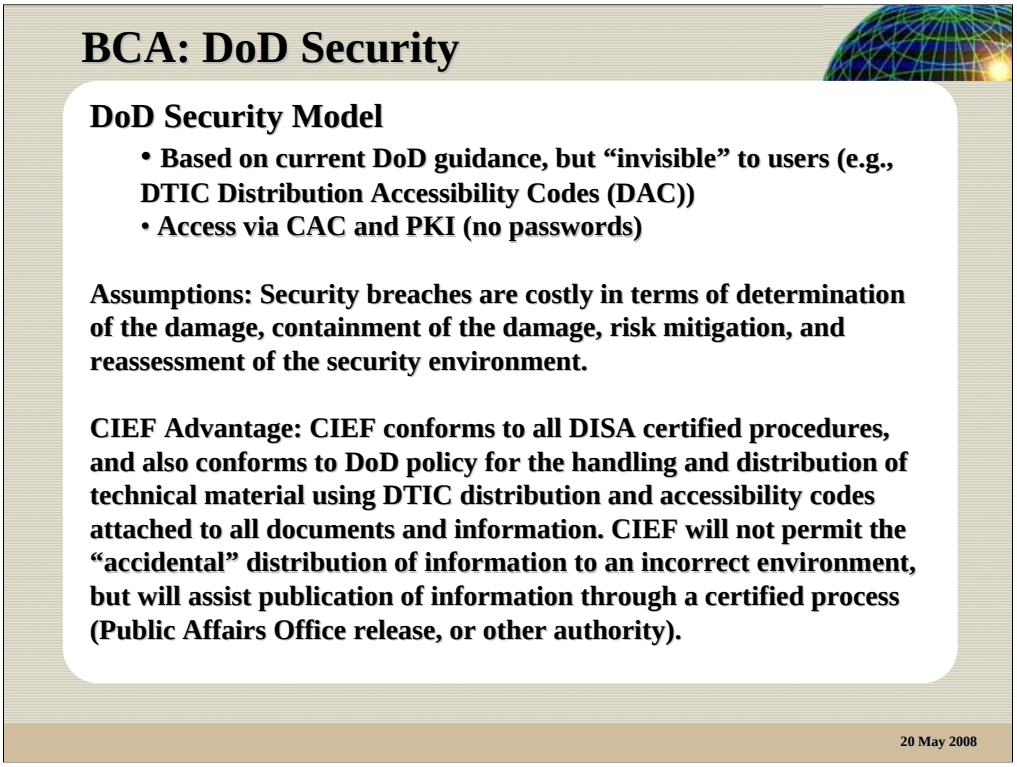
The last BCA of "Meets National Security Needs" addresses the major value of CIEF. This area of concern was the main focus of the initial design and subsequent improvements.



The adoption of a Community of Interest (COI) orientation was based on early discussions with the Joint Operational Effects Federation (JOEF) and other Joint programs. Typically, the problems of information exchange were described at a COI to COI level where definitions, formats, and even processes for handling information differed.

Additionally, the scale or size of the information set pointed to a problem that needed to be addressed in the initial system design. While efforts to integrate hundreds of data sources and millions of records have met with some degree of success (e.g., NIEM), the CIEF architecture addressed orders of magnitude more information and in a "near-real time" environment.

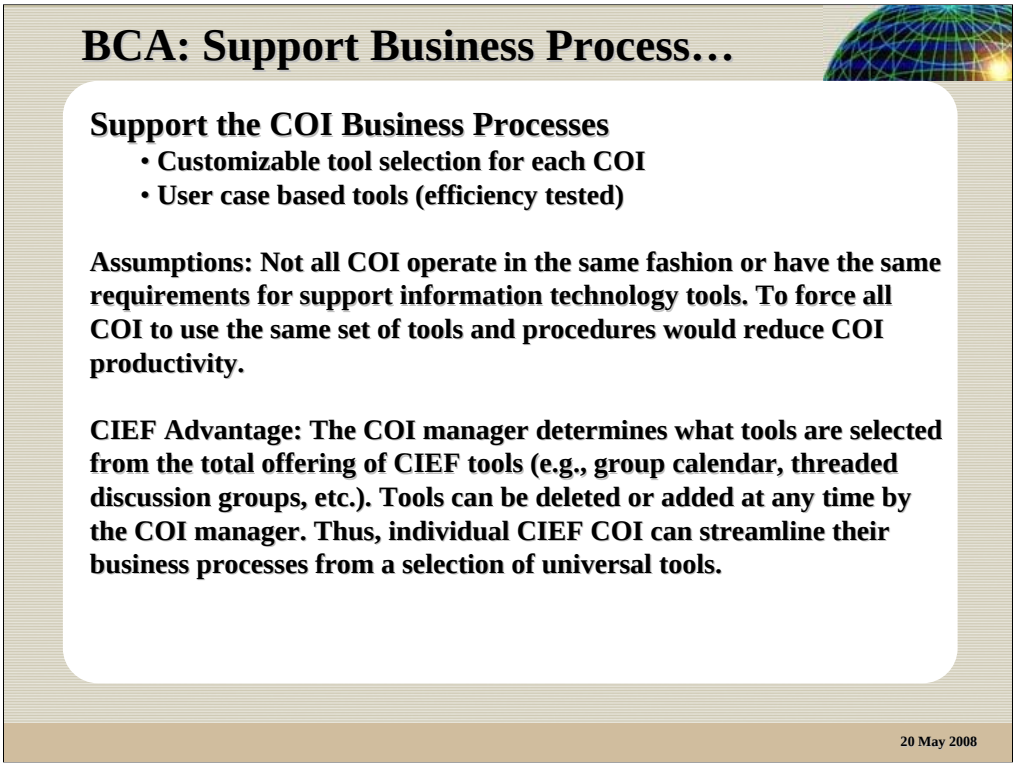
CIEF solves the complexity and scaling problems by breaking COI information exchange into many well defined and structured processes.



Based on previous efforts and good IT practices, security cannot be left to the end of a system design process. Rather, the CIEF architecture considered DoD regulations (i.e., DTIC DAC) and requirements in the exchange of information in both a single security level and across multiple security levels.

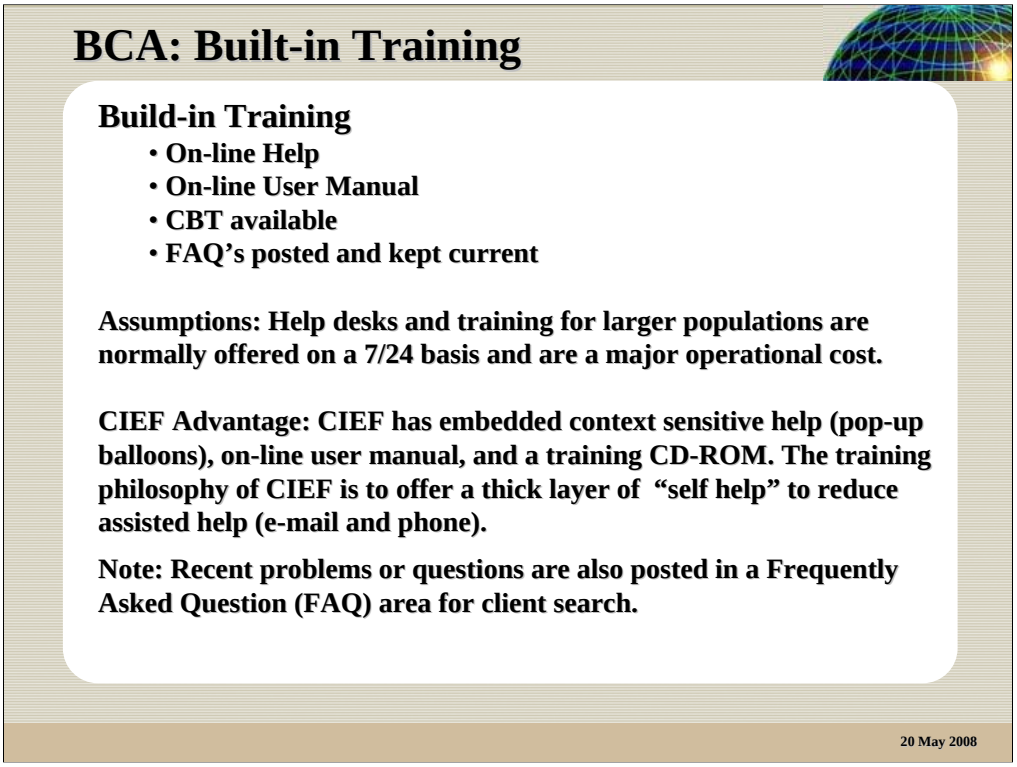
Initial CIEF builds will only address single security levels in its deployment (system high), and will address cross security level transfer of information in subsequent builds. The reason for the delay in deploying multi-level security is the lack of certified data level schemes (data level software encryption with no security hardware), and the complexity of the rule set for context based security.

CIEF deployments will parallel efforts in multi-level security with single level security with DTIC DAC with encryption at the distribution layer (e.g., Public, DoD Only, NATO, etc.).



CIEF was design to support the mission (business process). While tools and processes will be draw from a common repository of shared functionality, it will be up to the manager of the COI to determine what tools will be appropriate, and the corresponding "default" configuration of the COI (i.e., DTIC DAC level, mission tasking, etc.).

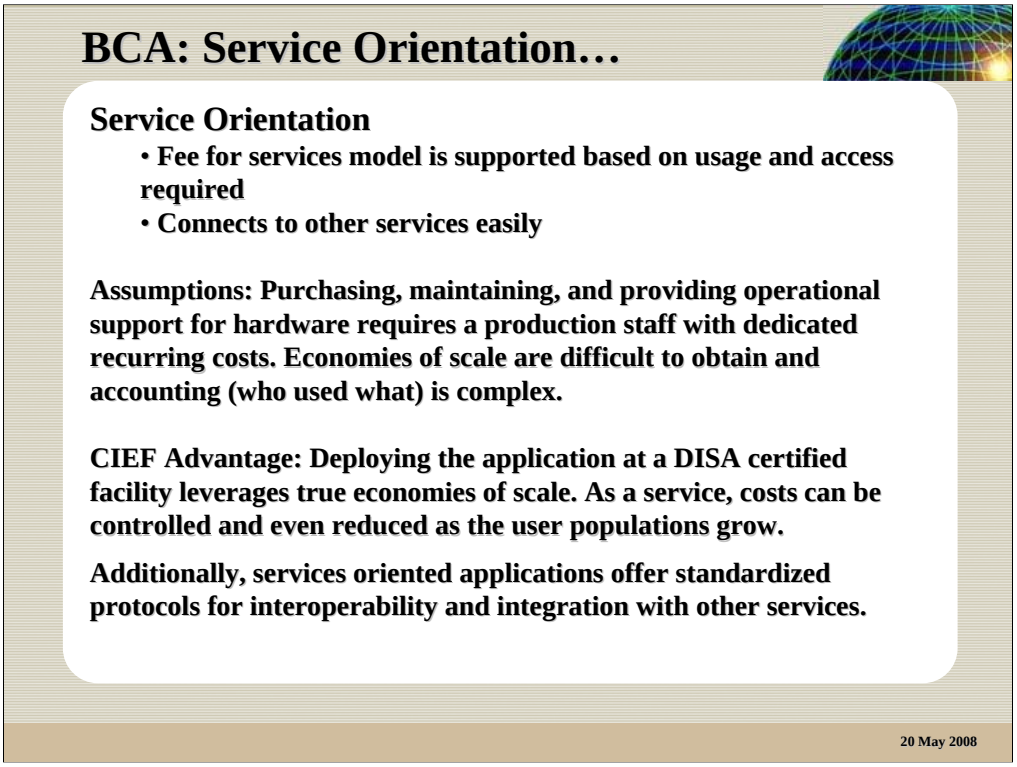
Typical CIEF tools are: Phone Book, Calendar, Threaded Discussions, Meeting "Rooms", Special Interest Groups (SIG) Collaboration, Workflow Routing, Archive Manager, Publication Manager, Search Assistant, and several existing web service applications, such as: the Technology eXchange Clearinghouse (TXC), the Virtual Knowledge Repository (VKR), and the Information Factory (data integration and publication).



The most common reason for communications errors in most information exchange environments is configuration errors. While automation in system configuration can reduce errors, training can further reduce operational errors.

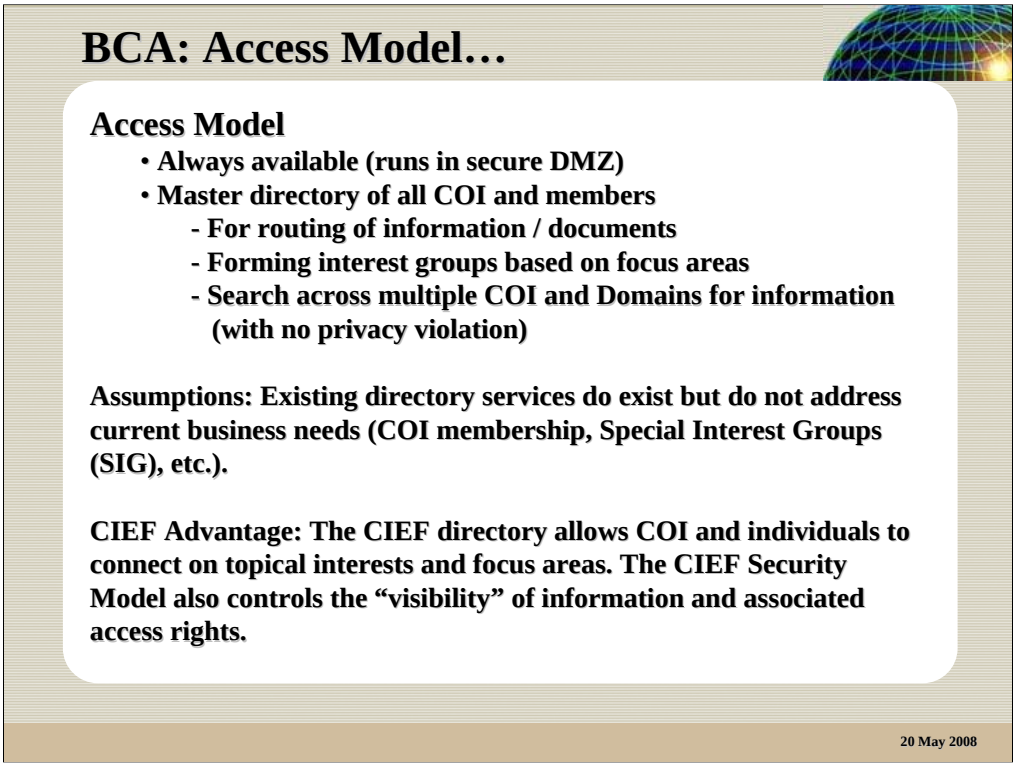
CIEF training will use a "Just in Time (JIT)" approach that is sensitive to the context of the problem. Additionally, best practices will be presented in both text and visual presentations.

The overall philosophy of CIEF training is, an upfront expenditure in training resources will yield many times the benefits in client usage and system performance.



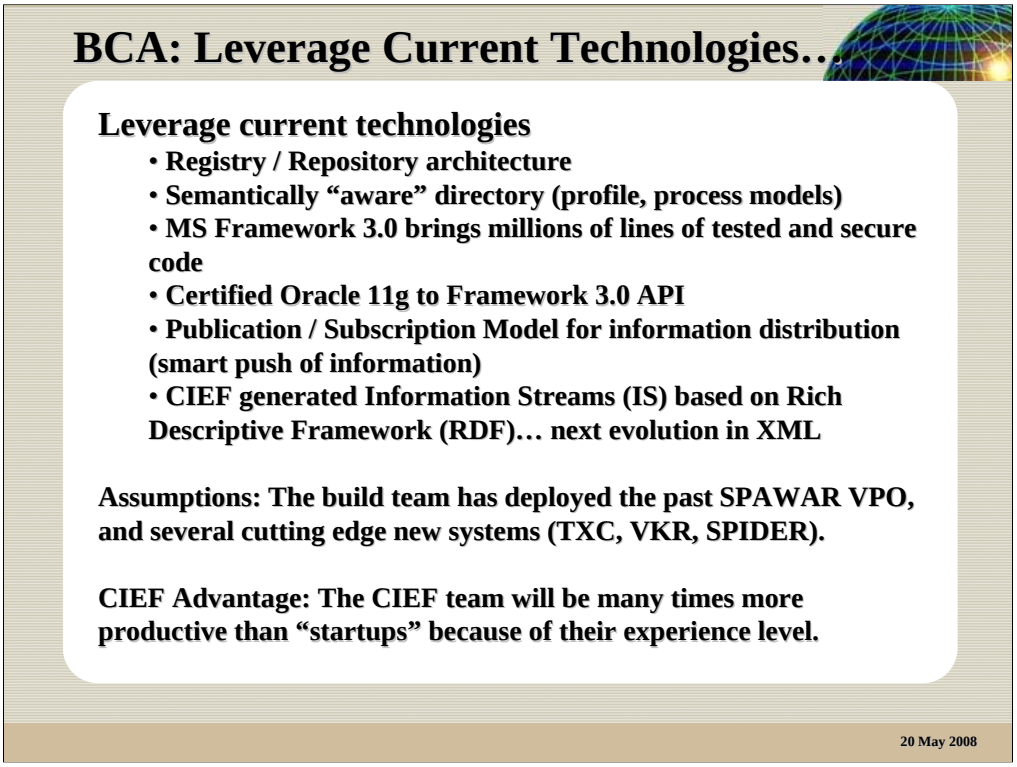
CIEF is built around a clearly defined registry structure that lends itself to centralized administration, such as at the DISA Mega Centers. The Admin registry that was designed to assist in data strategies and valuation of information can also be used to accumulate usage information for billing in a fee for service environment.

CIEF is also extends current web service architectures to support the integration of current web services with existing IT systems that may require Quality of Services (QOS) and security considerations (typically not offered by web services).



CIEF was designed to be deployed in a fire-walled DISA enclave, typically called a DMZ. While CIEF would be accessible through the Internet, that access would be controlled through DoD certified methods.

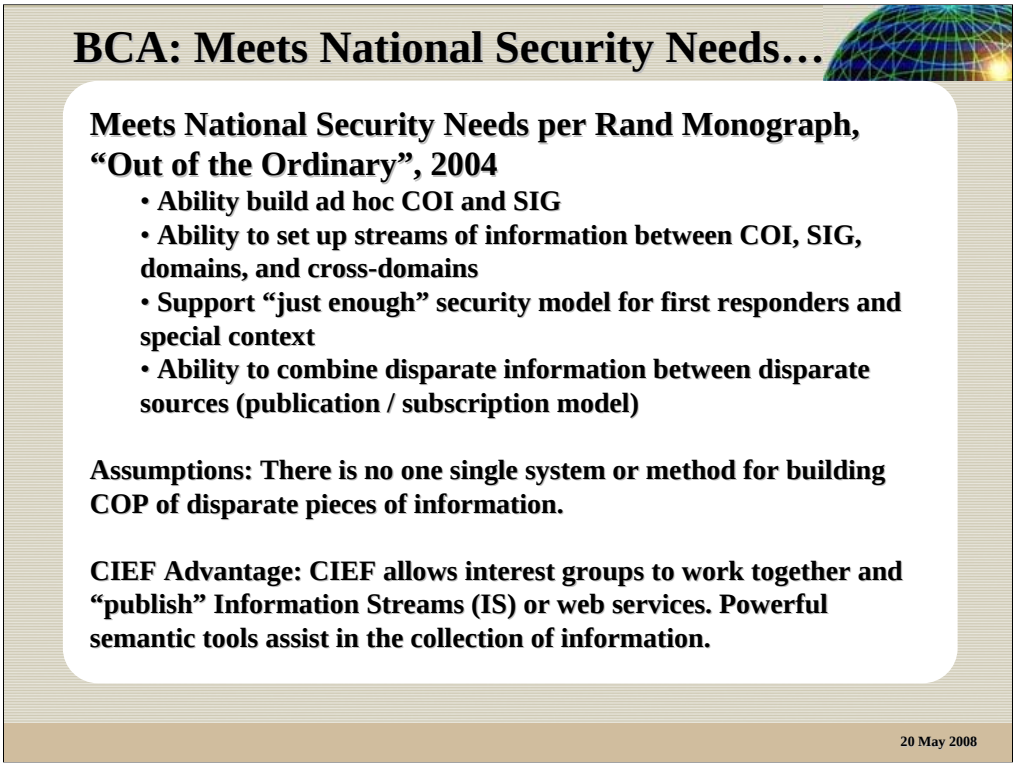
CIEF was designed with a balance of controlled access and ease of access to information. Also considered was the privacy of personal information based on an approved usage model and granting of rights by the owner of the information.



While CIEF is based on semantic technologies, it also leverage current industry software frameworks and COTS databases.

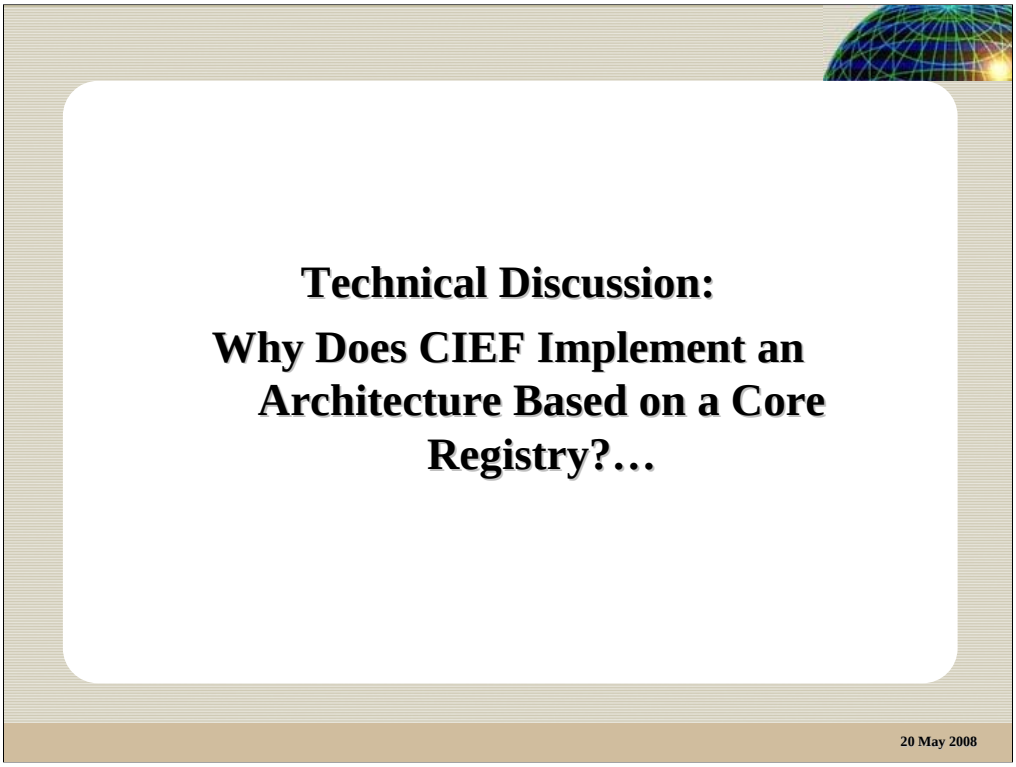
CIEF also extends industry standards to include DoD requirements for QOS, security, and domain specific information.

One of the goals of CIEF is to use proven technologies in a well integrated system. For example, with minor adaptations the Really Simple Syndication (RSS) news feed specification can be adapted to meet DoD requirements for security, privacy, and domain specific information. Note: This technology is not only proven, is being used in the publication of over 200,000 Internet information sources.



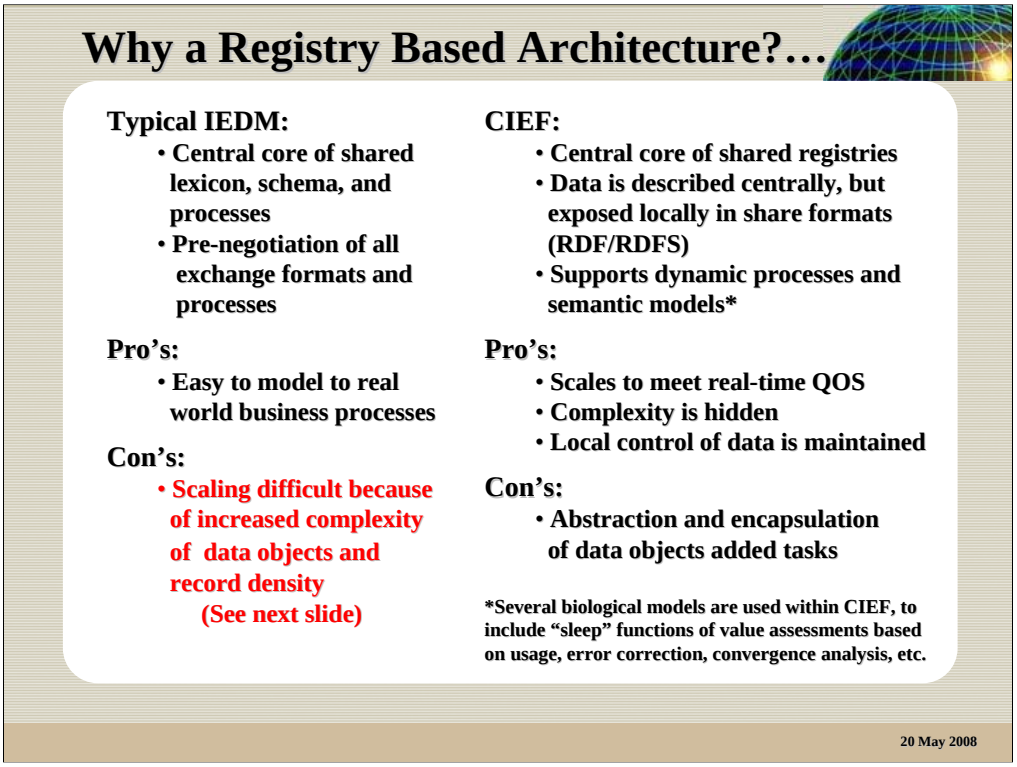
CIEF was originally designed to meet the requirements presented in the Rand Monograph.

Thus, CIEF's design was driven by national mission requirements, rather than abstract or general requirements.



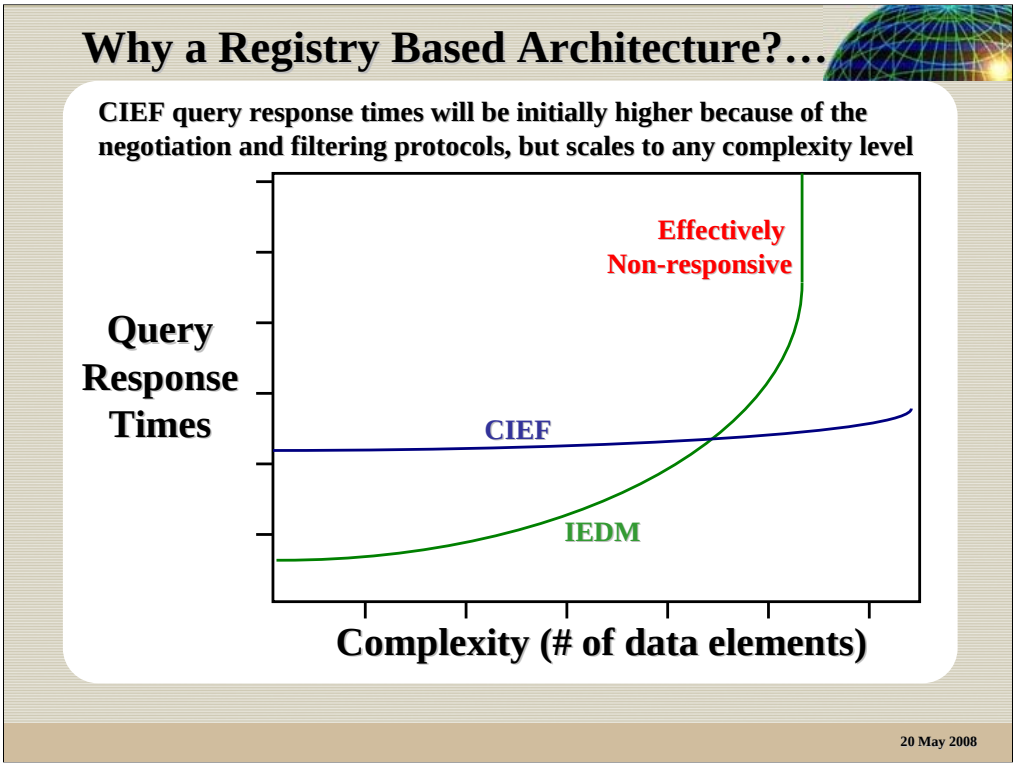
The design orientation of CIEF should be understood to appreciate how it differs from other information exchange methods.

The following section attempt to present a balanced view of the strengths and capabilities of the CIEF architecture.



CIEF differs primarily from current Information Exchange Data Models (IEDM) in its focus on the mission context and corresponding filtering of the Valued Information at the Right Time (VIRT).

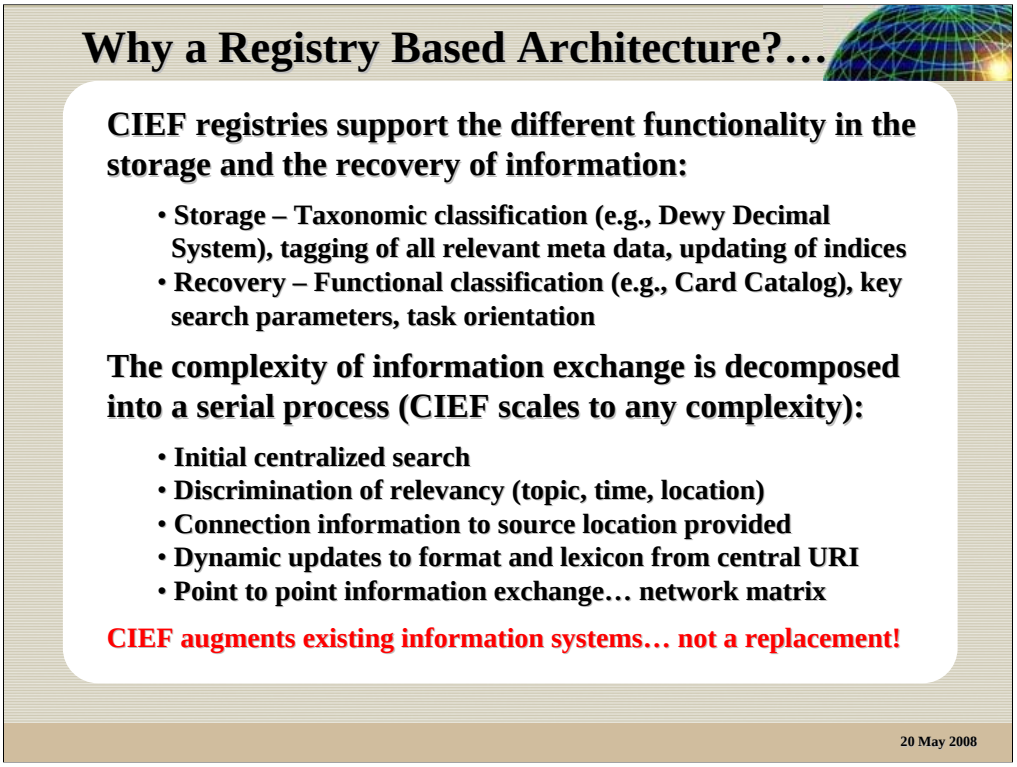
Beyond the basic architecture, CIEF offers a framework that can scale beyond simple exchange models to include all information and data sources of the Global Information Grid (GIG).



Typically IEDM based architectures are useful for a contained domain of information, with specifically constrained contexts. As their usefulness grows and corresponding complexity (number of data elements), trade-offs of accuracy or precision must be made. Failure limit complexity will result in excessive computation times and a non-responsive system.

CIEF handles complexity through the constraining of initial states (profiling requests), abstraction and encapsulation of data (raw, working, archived), and tracking of relevant information on the physical network.

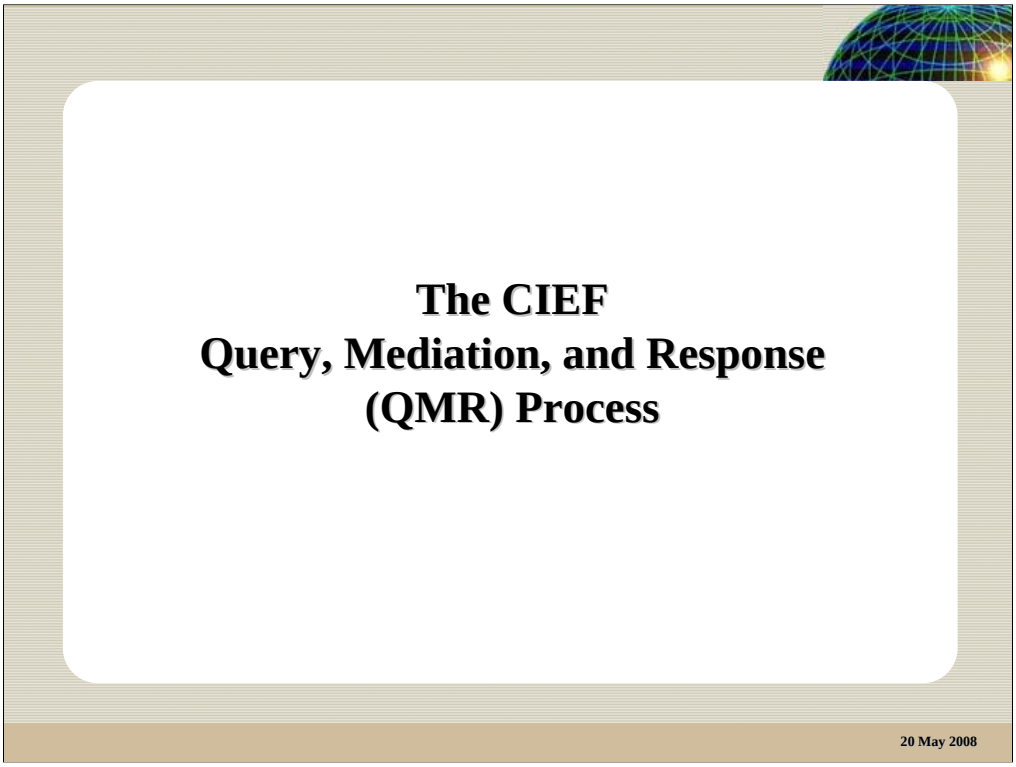
While information exchanges in CIEF will carry a fixed overhead of profiling and negotiation, that overhead will not be impacted by the number of data elements in the overall system.



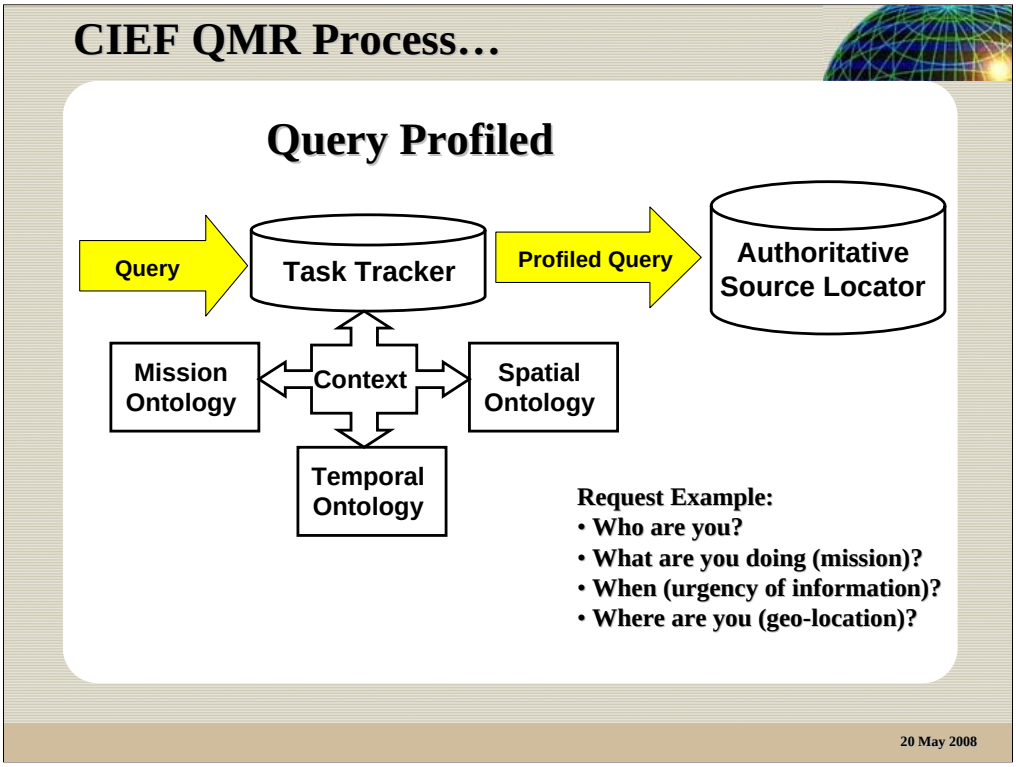
Library Science 101… the way you store information is typically not the way you retrieve information.

CIEF uses a storage process (shared ontologies) to classification and putting things in the right place, similar to the Dewy Decimal System used in libraries. This process is efficient with regard to: the reduction of duplication, validation through multiple sources, and other "mechanical" aspects of data management.

The retrieval of information in CIEF is mission/task based and has uses very different set of attributes to recall information (e.g., Library Card Catalog). This orientation takes into consideration the mission tasking in determining the context of the information request.



The following diagrams can be thought of as Over View Ones (OV 1) and demonstrate the concepts of operation of the core CIEF processes.

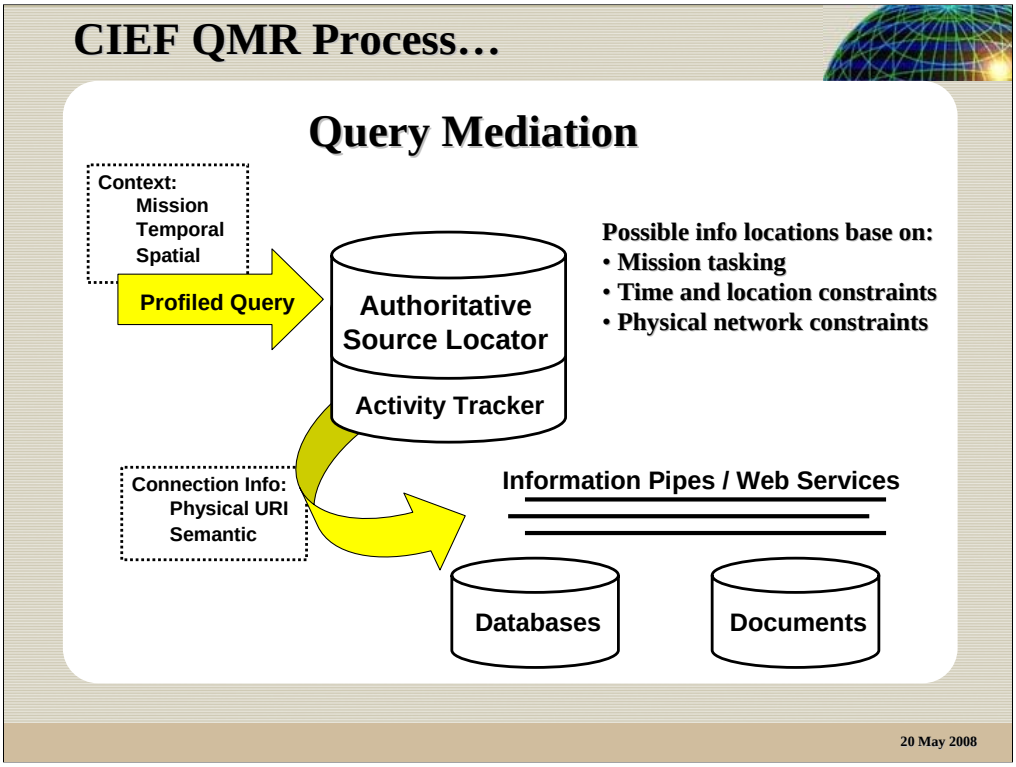


Concept of Operations (CONOPS) of a Profiled Information Query:

1. Query is submitted to the Task Tracker.

2. Based on the identity of the requestor and the content of the request (e.g., organization, mission, security level), ontological information is added to the request.

3. The profiled query is submitted to the Authoritative Source Locator to search for the appropriate information match and possible locations.

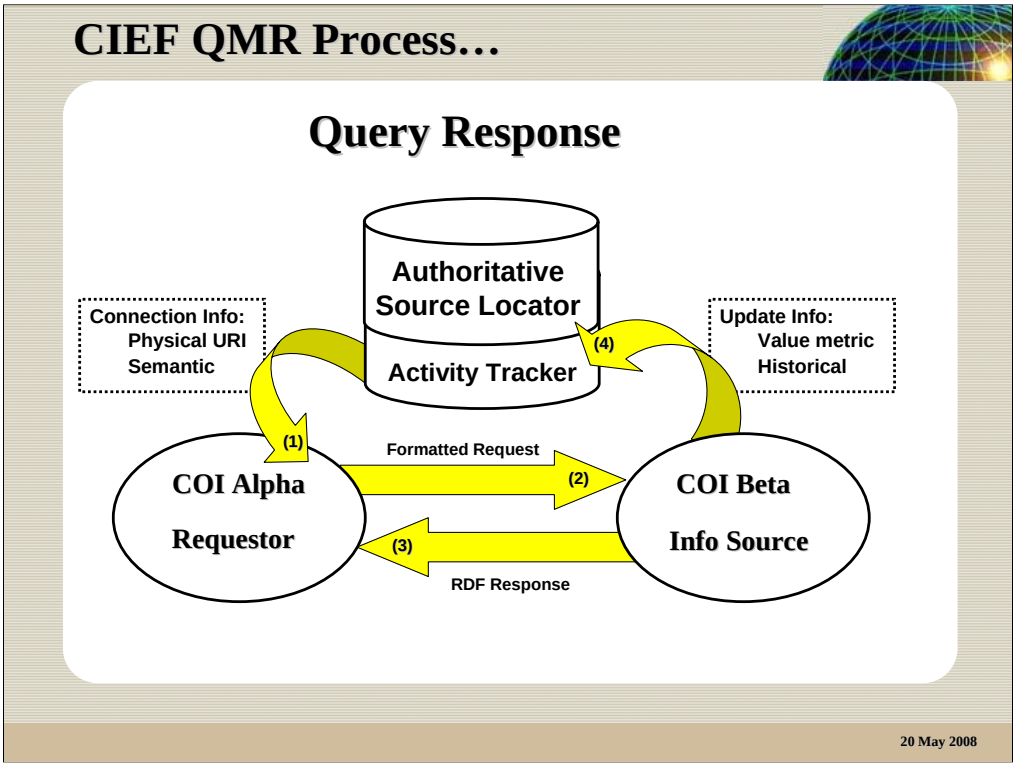


CONOPS of Query Mediation (continued):

1. Profiled query is submitted to the Authoritative Source Locator (ASL) with the additional context metadata.

2. Appropriate information provider is "discovered" in the ASL.

3. Physical access parameters (location and format) and semantic information are provided to the requestor. Note: This includes predictive information based on histories and "value" assessments.



CONOPS of a Query Response (continued):

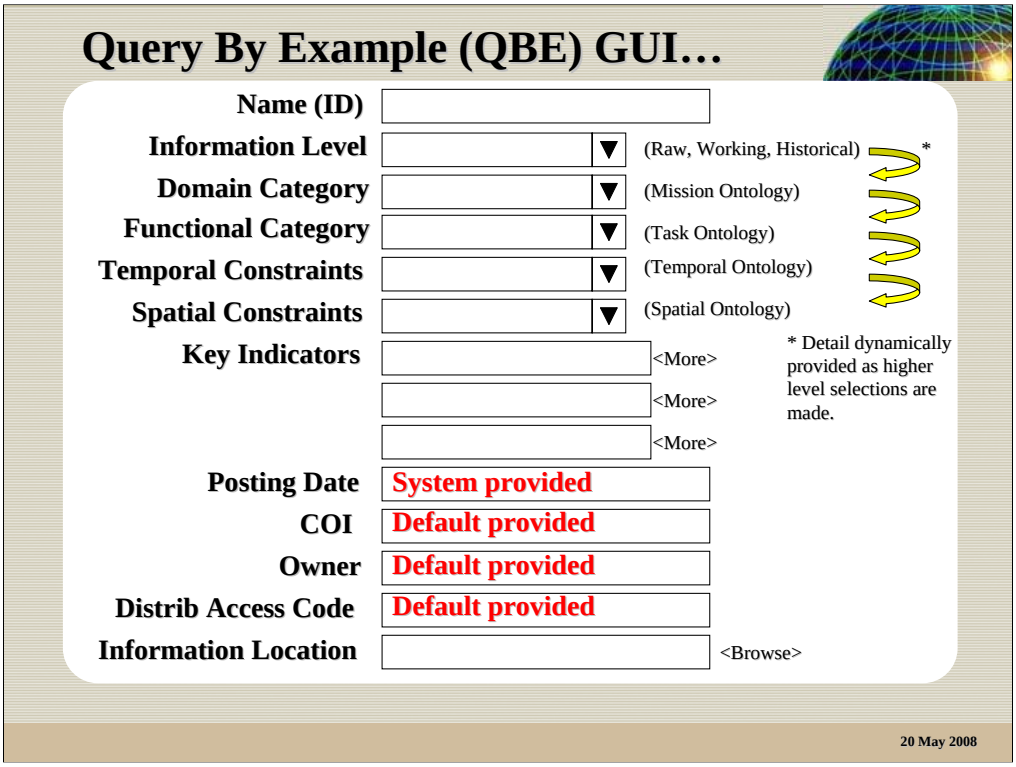
1. Physical access parameters (location and format) and semantic information are provided to the requestor by the Activity Tracker to the requesting source.

2. A RDF (Rich Descriptive Format) point to point information request to the source location is initiated.

3. An RDF packet with the information is provided to the requestor.

4. The Activity Tracker "listens" to the transaction and provides feedback on "best path" and historical information. Note: The Activity Tracker can also be interrogated for "valued sources" and other data strategy functions.

5. The requestor will use tools to "digest" the information, and possibly republish the information (e.g., from "raw" to "working").



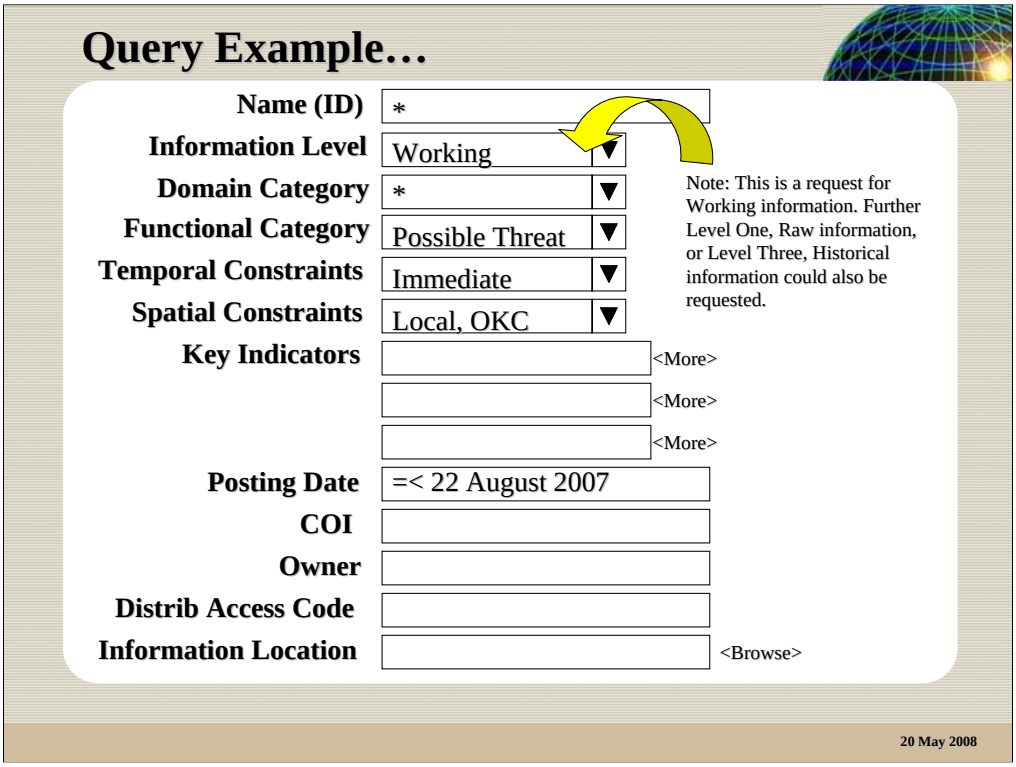
A web based tool would be used to categorize and publish metadata about information.

Initially information would be categorized by its abstraction level (raw, working, or historical) and then by relevant mission and tasks. Temporal and Spatial constraints would also be added.

Default information from the posting COI would also be added to the registry entry.

Note: From a top down perspective, as information detail is provide, the appropriate subcategories (Lists of Values (LOV)) are exposed in the selection pull-down lists. For example, as the Information Level of interest is selected, the relevant Missions are selected from the overall Mission Ontology in the Domain Category pull-down.

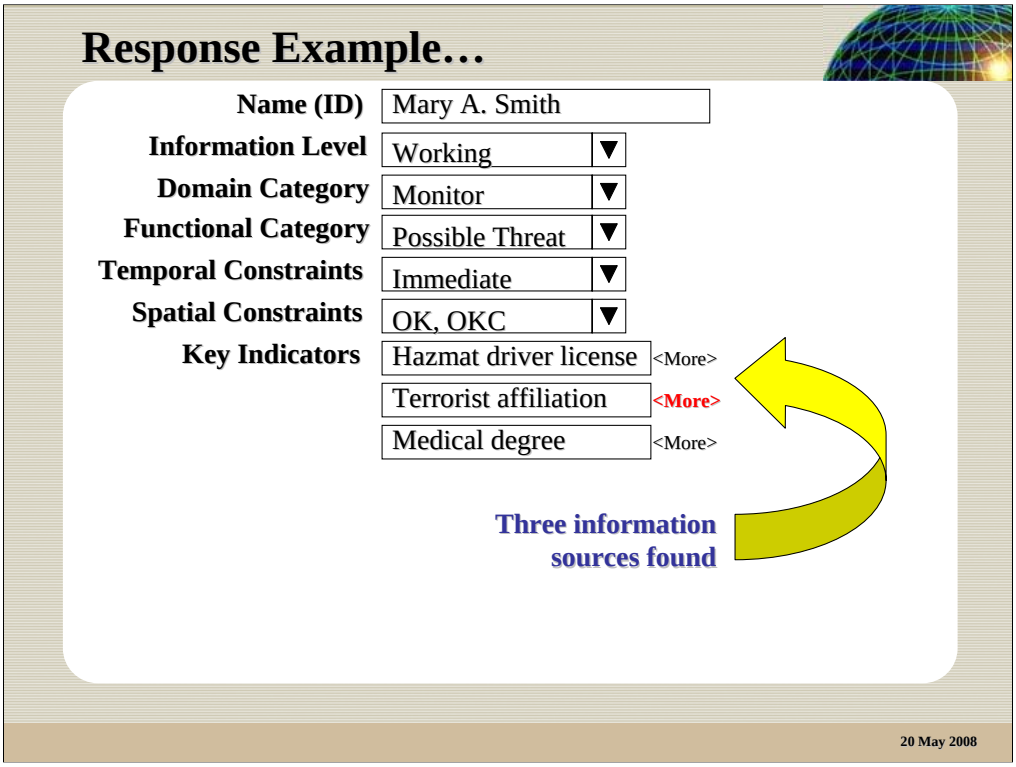
Wild card symbols in names and fields will be permitted.



Using a simple Query by Example (QBE) process, information could be requested from the Content Registry.

The query can be refined or broadened per the client's search criteria.

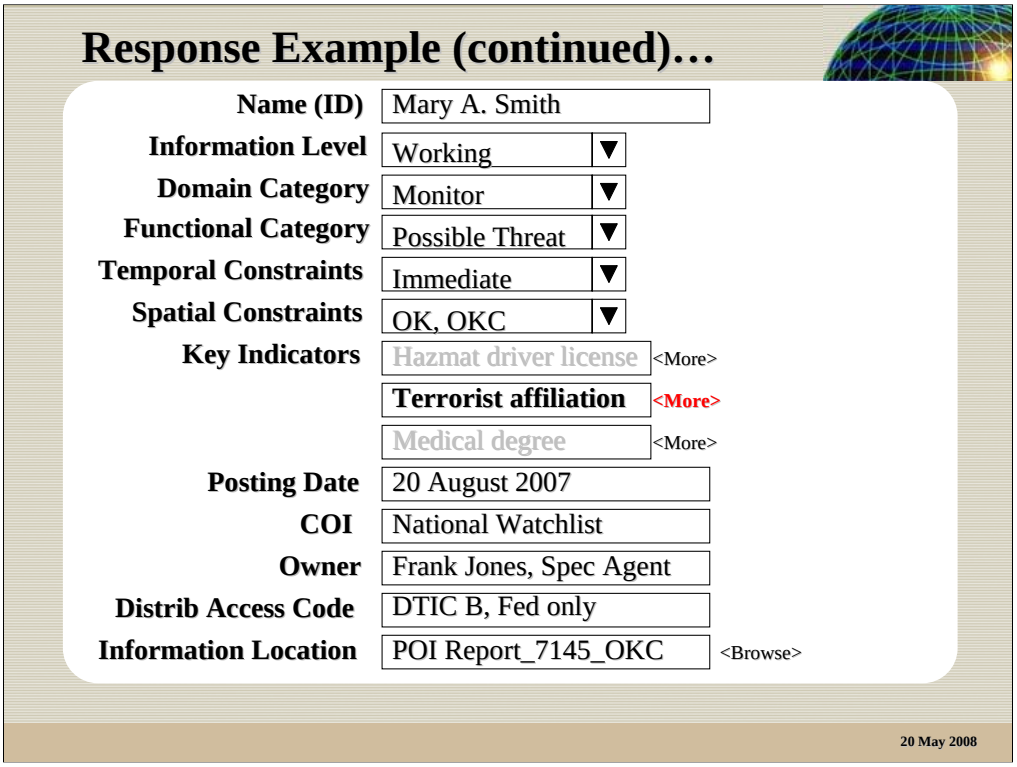
Note: Other GUI designs such as wizards, natural language input, key word searching, topic and query maps will be explored in the fielding of CIEF.



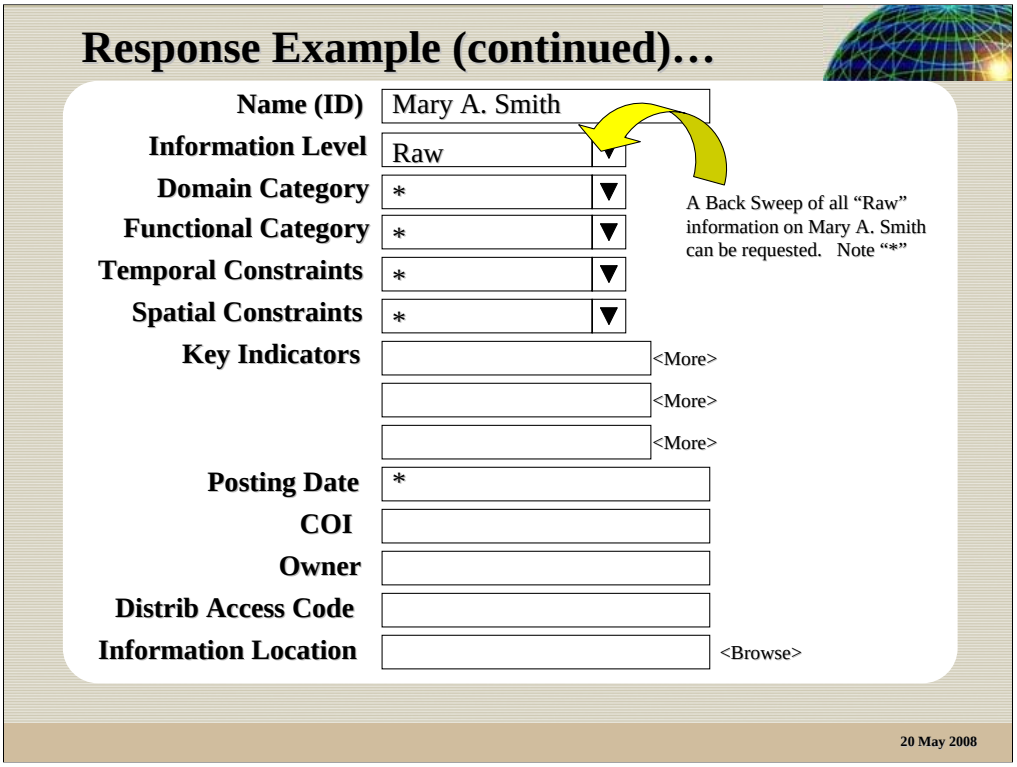
In this example, matching meta data information is provided, with a link to the actual Person of Interest (POI) report for the local Oklahoma City area.

Note that the Level Two, Key Indicators are the result of the recombinant processing of Level One information by the owning COI.

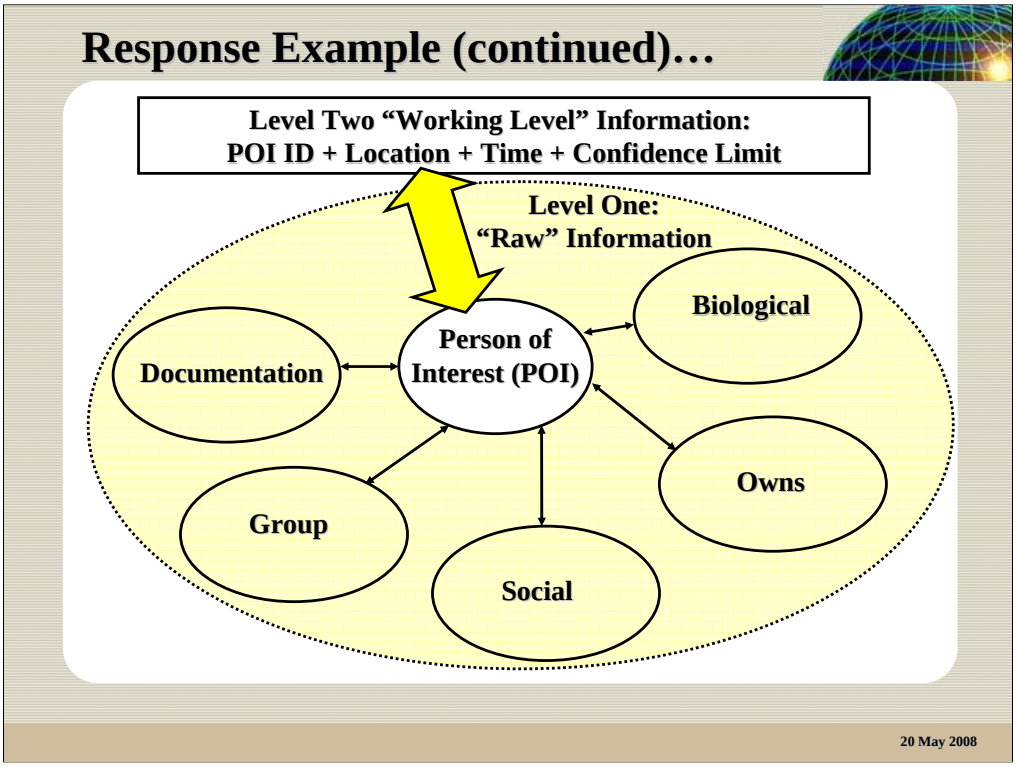
A single atypical indicator would probably not have caused concern, but the combination of two atypical indicators with the POI's background of a medical doctor was sufficient to raise the POI's threat level.



<More> information is requested from the Content Registry… dates, COI, etc.



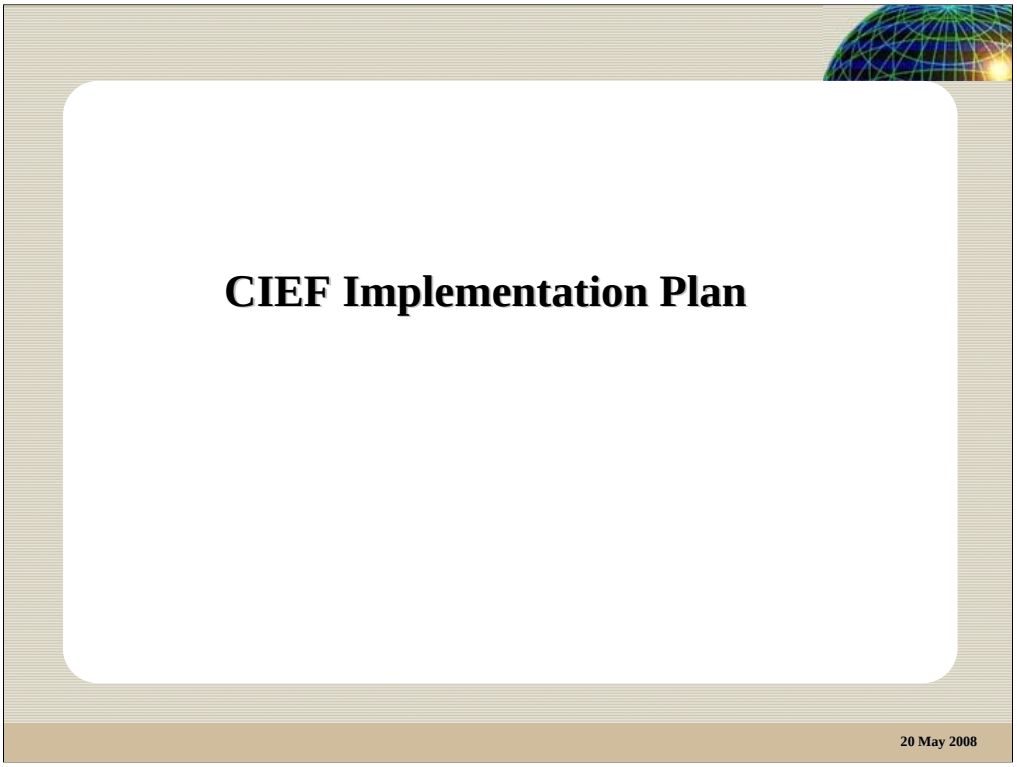
An additional Level One (Raw data) could be requested… of any and all information (*).



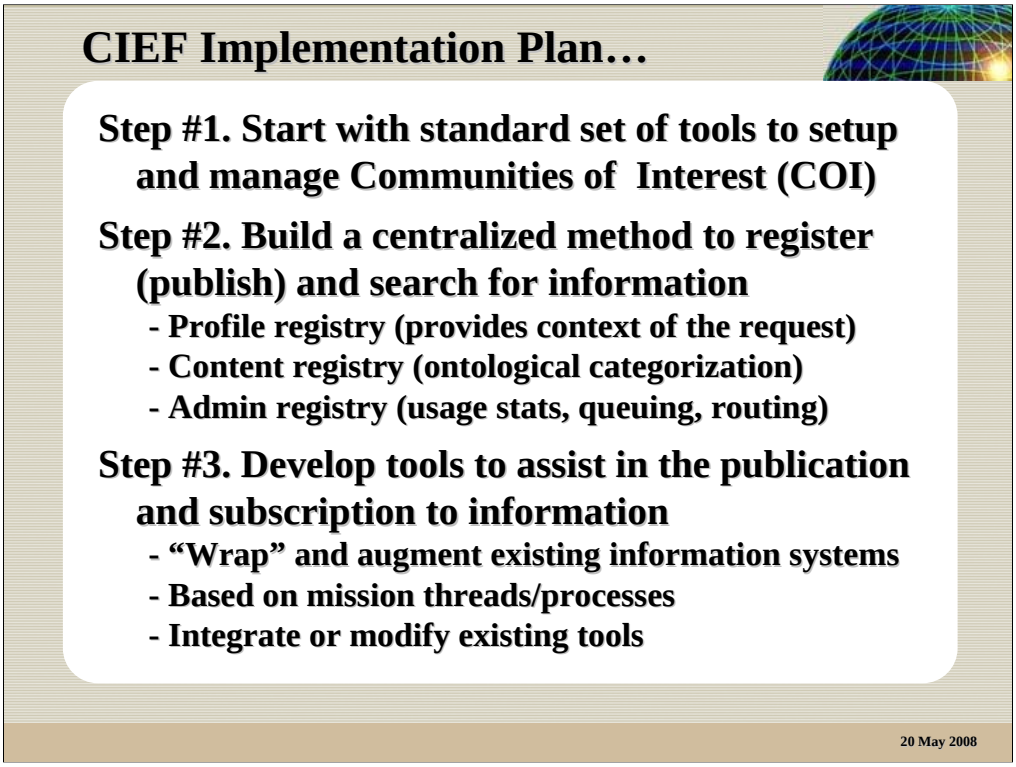
Example of the re-publication of information at a higher level of abstraction (from Raw to Working).

Although a complete picture of a Level One POI may be available (and centrally registered), only an encapsulated summary of the information is used by Level Two, within a defined pattern… thus, abstracting the information and hiding its complexity, but providing enough of the right level of information to satisfy the mission requirements.

Note: Level Two processing may add still more associated analytical information, but related to such concerns as: confidence of the information, validation of hypothesis, and related POI.



Note: The plan is somewhat modular and can be "done" in any number of sequences depending on the emphasis of the sponsoring community.

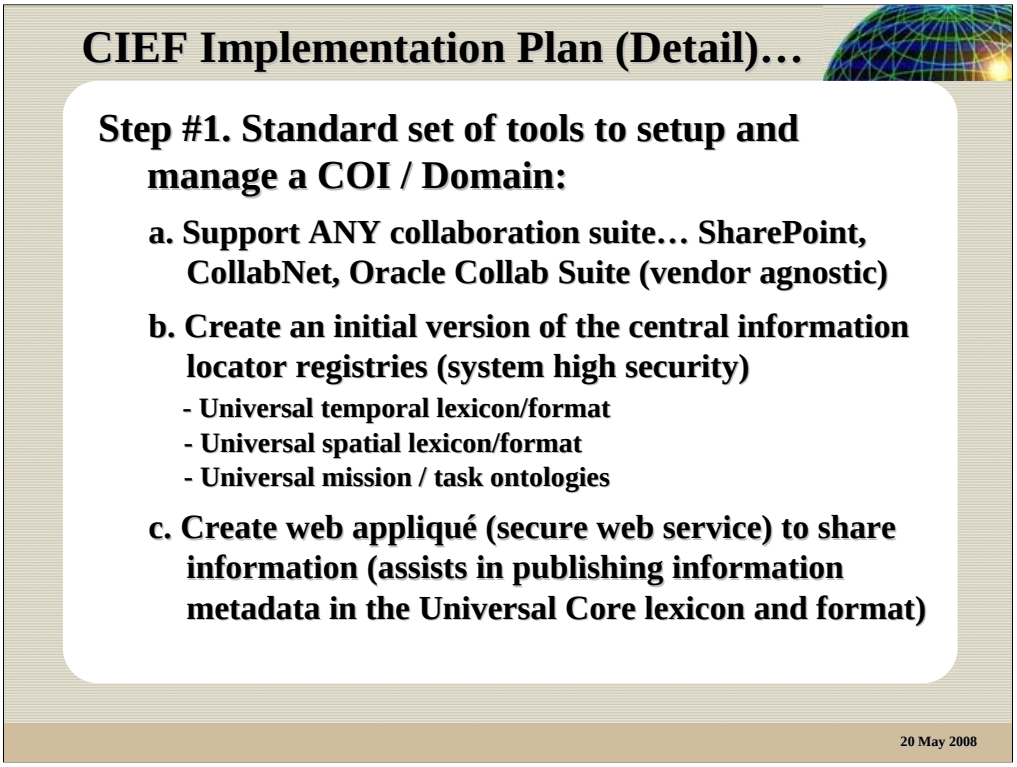


1.) Do the basics first… that means provide the tools and services to enable COI formation and some mechanisms to support them… basic document management, threaded discussion groups, and even phone and routing lists.

2.) Present UDDI services are so limited in their service description that the requestor would need to know more about the service to be able to use it in any meaningful way. Commercial web services do not meet DoD requirements for assured delivery, timeliness, prioritization of information, and many mission derived requirements. Bottom line: DoD web service registries need to be a whole lot "smarter" and focus on DoD missions.

What are needed are registry structures that support missions, corresponding actions, and have contextual references (who, when, why, etc.) based on the information exchange requirements of the mission.

3.) More on tool development on subsequent pages.

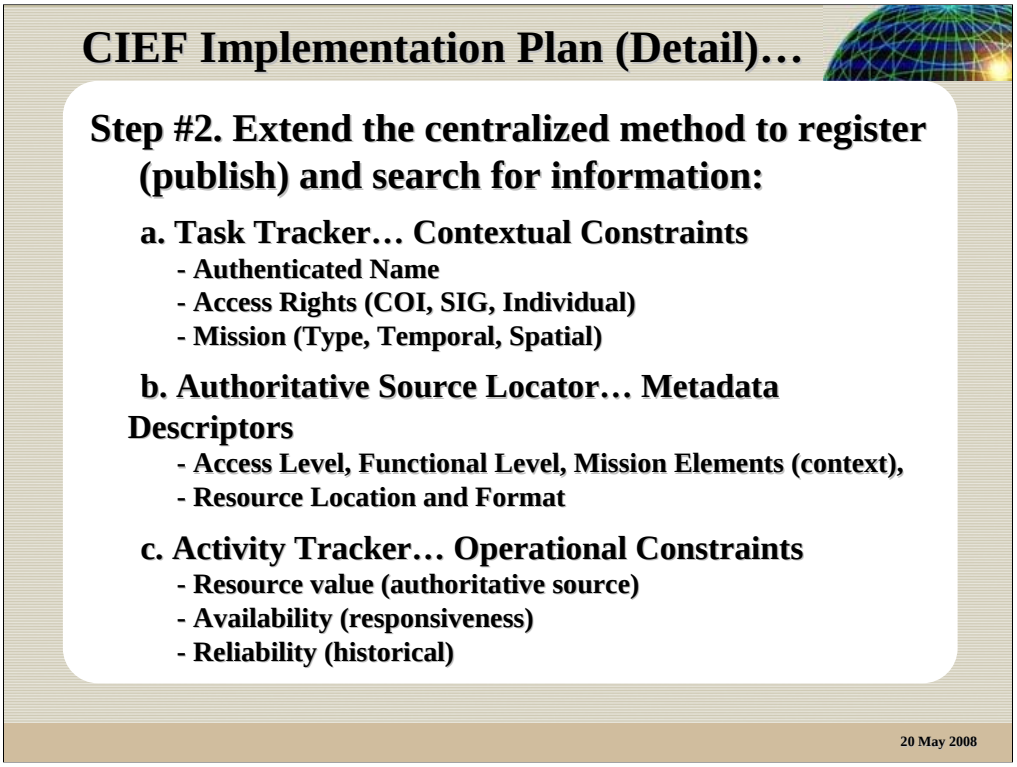


Step #1. Do the basics first…

a.) To even start to share information some method of organizing and managing information must be instituted… based on COI rules, membership, focus areas, etc. Additionally… Investments in current and past collaboration tools must be leveraged.

b.) Although information will still be managed locally (by the COI), information to be shared must be described in a centrally managed in a meta-data content registry. Tools to "publish" information will be web based and will provide enough meta-data to locate and access the local COI information.

c.) The central registry will provide mission based content metadata and physical connection information to local COI sources. Note: This is a major focus area of the Universal Core lexicon and format descriptors and the associated Universal Core mission/task ontologies.

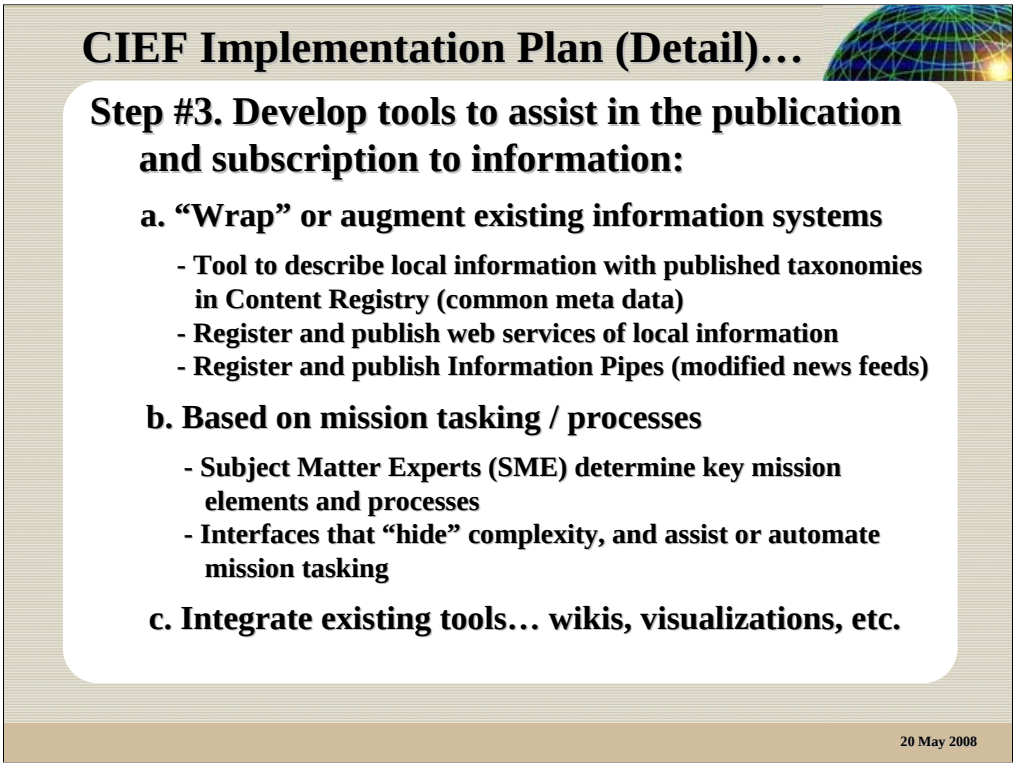


Step #2. In the central management of the Universal Core of shared lexicon, formats, and processes is necessary to build:

a.) The Task Tracker to manage the "who" of information access and subsequent rights. With permissions established, mission tasking can be decomposed into temporal and spatial information requirements.

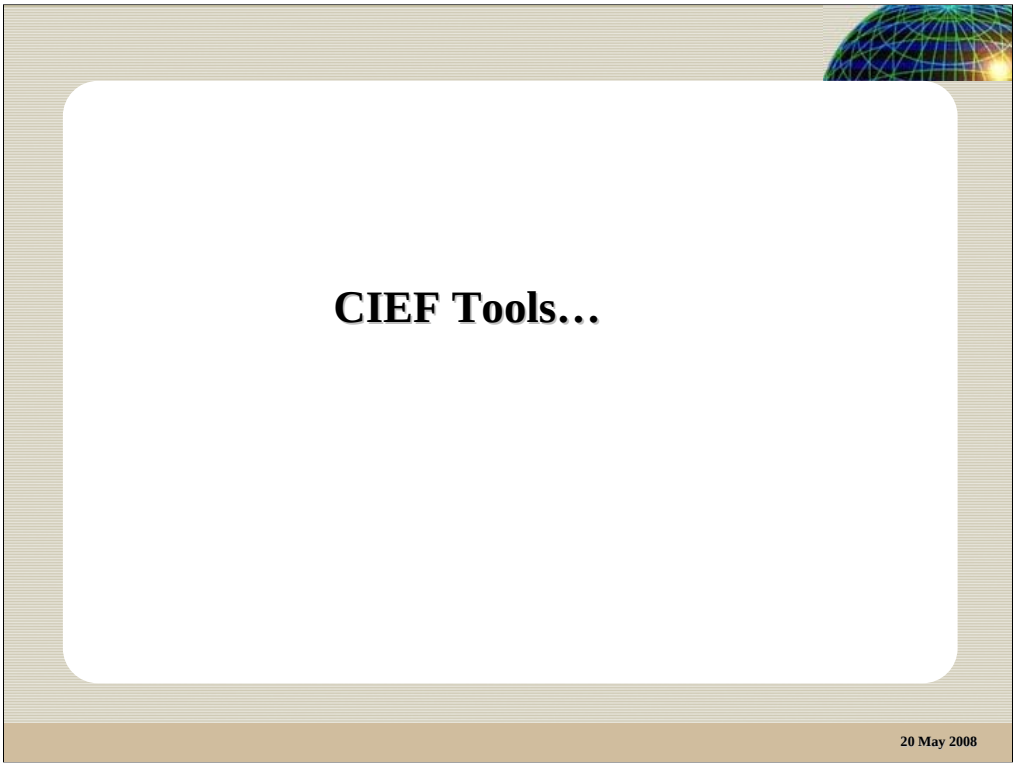
b.) The Authoritative Source Locator has "just enough" metadata description based on the mission, temporal, and spatial context of the request to "point to" the location of the information and provide connection and format information.

c.) The Activity Tracker provides Quality of Service (QOS) information to access best paths and information sources. Also, both real-time and historical qualitative information can be analyzed to determine valued sources and associated metrics.

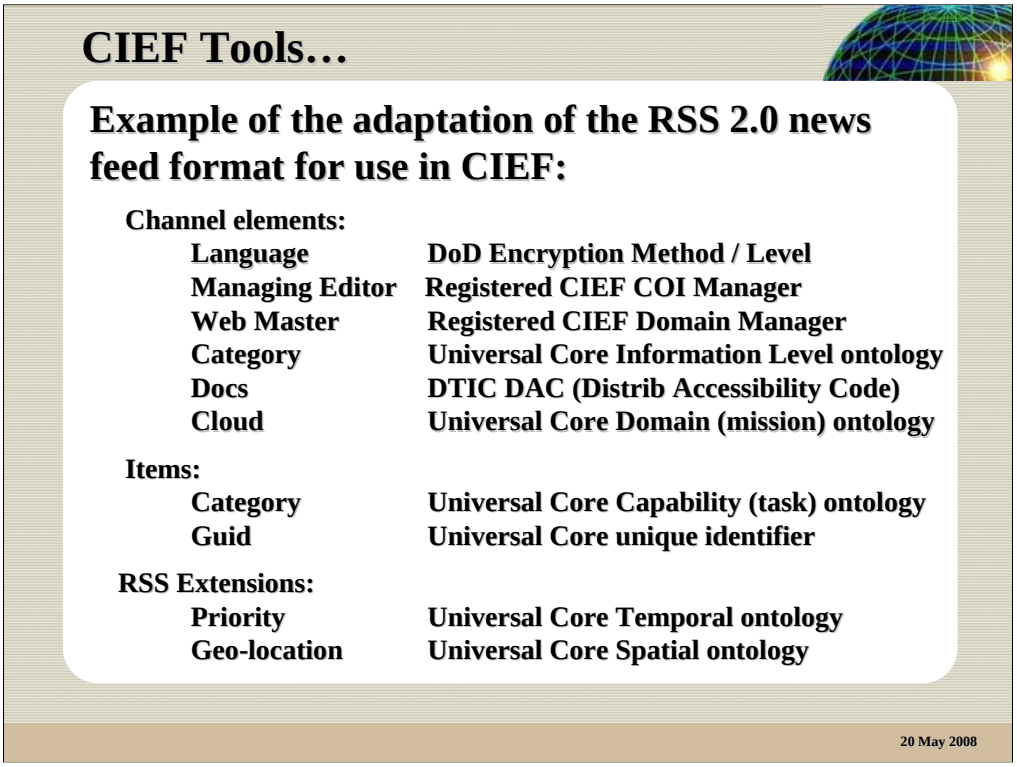


Step #3. Tools that normal humans can use must be developed to assist in the semantic categorization of information, the maintenance of dynamic ontologies (new tasks, missions, threats, etc.), and the integration of multiple source information.

The following pages illustrate such tools that have been used in commercial banking, and in DON applications.



The following are examples of CIEF adapted tools and processes…



Ten's of thousands of current RSS news feeds, news feed readers, and associated applications can be leveraged, and extended to add DoD specific functionality… mission/task orientation, DTIC distribution requirements, NSA certified Encryption, etc.

Minor modifications to the existing RSS format can be made, to include:

- Data Abstraction Level (From raw, summarized, historical)

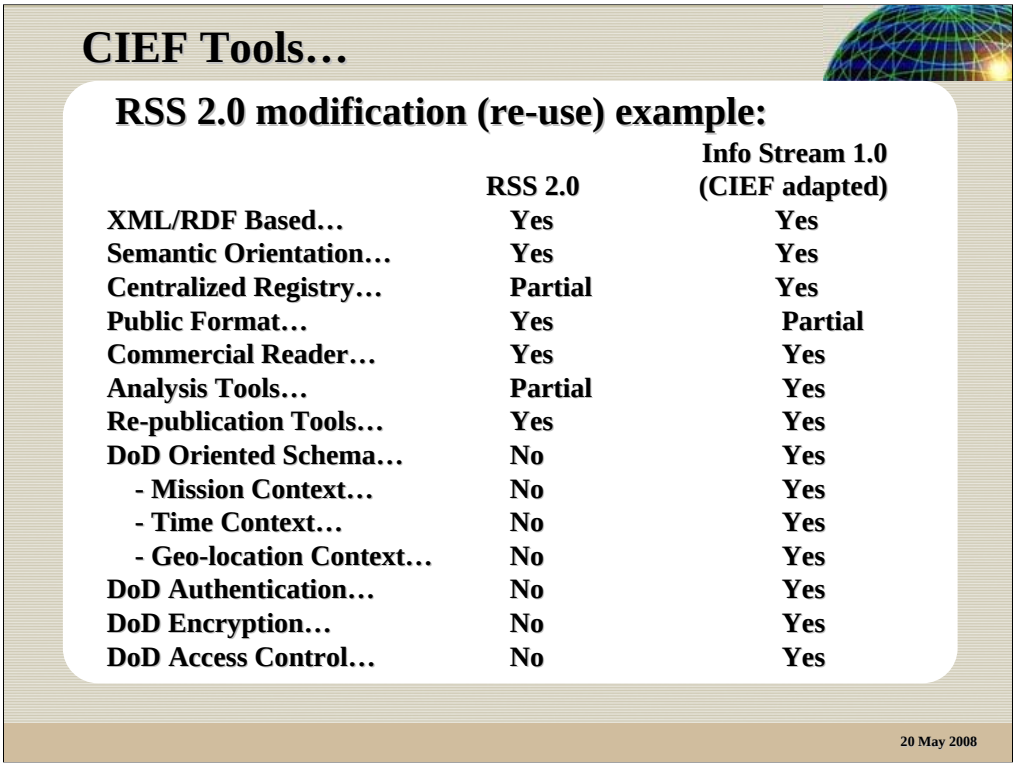
- Access control information

- Source information (COI, Domain)

- Domain information (Universal Core Domain Ontology)

- Categorization information (Universal Core Category Ontology)

- Context information (Universal Core Temporal and Spatial Ontologies)

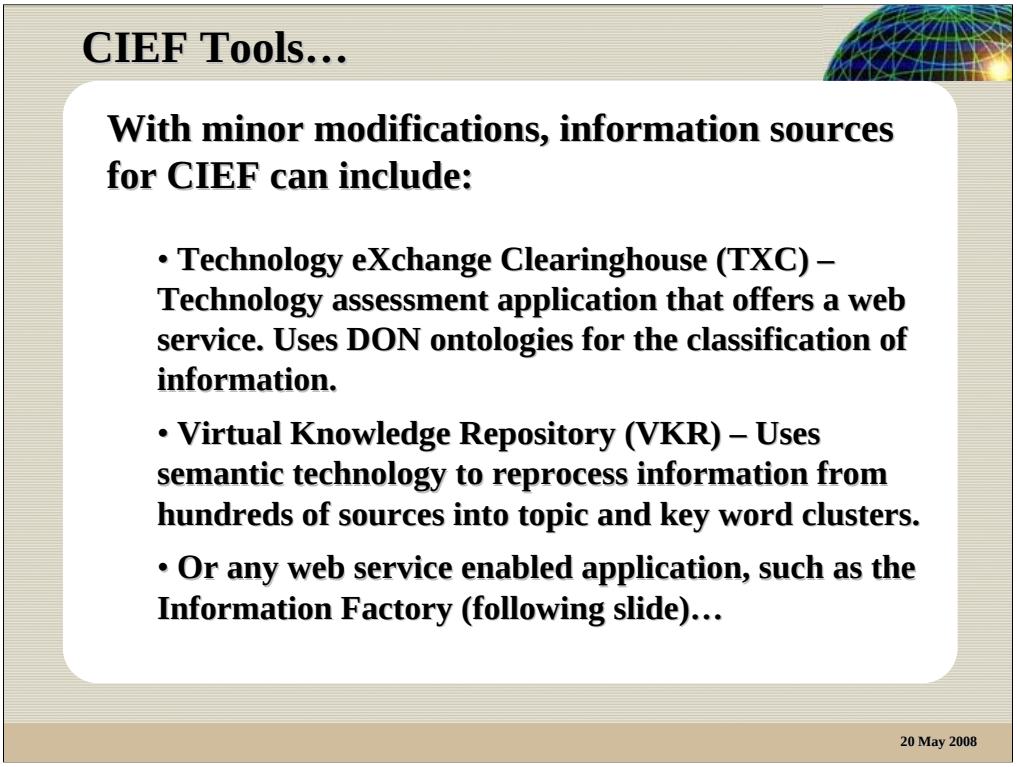


Universal Core adapted RSS news feeds can be at the top of the "well define context" stack… with metadata registries providing enough contextual attributes to "sufficiently describe" the information content.

In other words, CIEF provides a method to get "close" to the correct information without any previous knowledge or pre-negotiation.

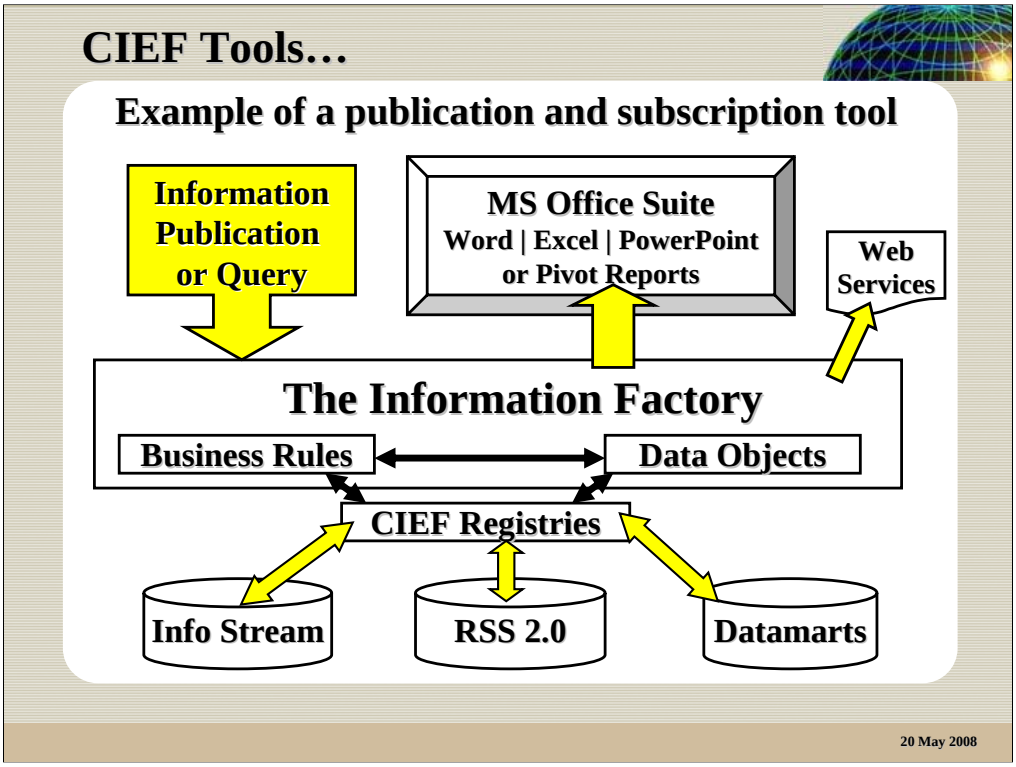
Further, information from raw sensor data to summarized reports can be combined from multiple sources of information into cohesive bodies of knowledge (recombinant information).

The ability to produce recombinant information through multiple data source integration (400+ RSS news feeds, document manage-ment systems, databases, etc.) was demonstrated in SPAWARSYSCOM's Virtual Knowledge Repository (VKR) in 2003. Note: VKR was built on semantic technology.



Both TXC and VKR represent "easy to integrate" application through web services.

Applications such as the Information Factory represent the "alternative" applications that are client-side based and required the installation and configuration of a systems analyst. The configuration process would include the encapsulation of the client's data objects and business rules that would be exposed through web services or Information Streams (CIEF modified RSS New Feeds).

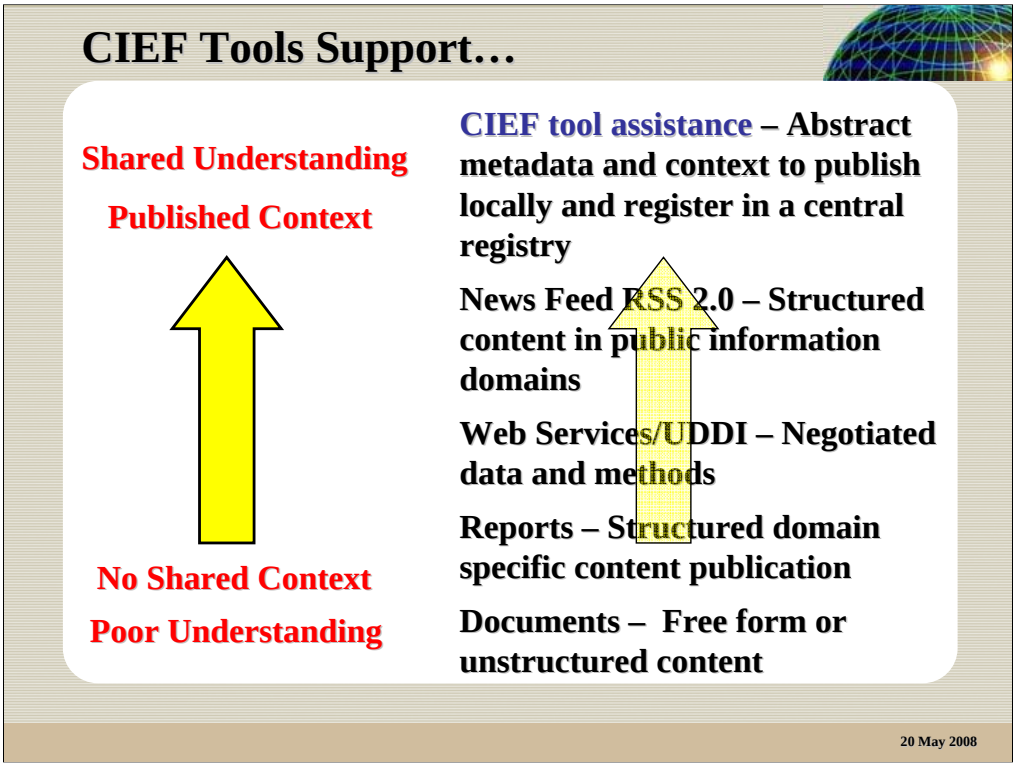


The Information Factory is an example of an application that has been built by iBASEt and that can be adapted CIEF use.

The Factory "exposes" an underlying database or information source by mapping business rules and data objects to an MS output application or web service.

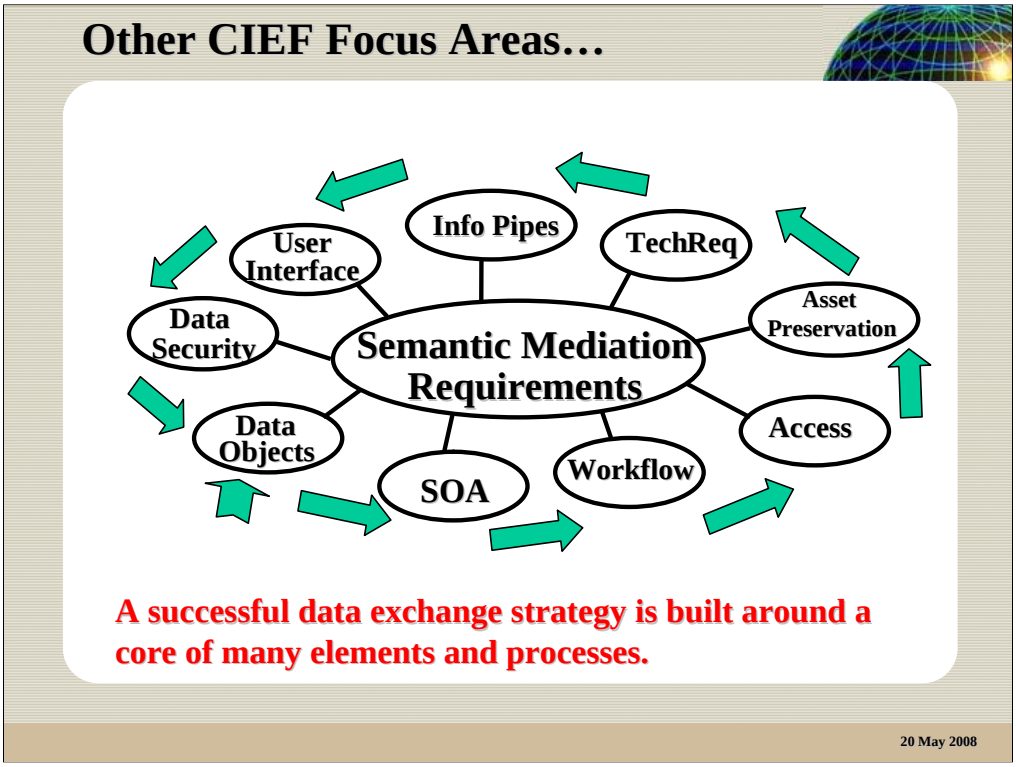
With the additional step of registering the Factory's web services and describing metadata content, the Factory could be used as a shared information publication tools.

Similarly, the Factory could access CIEF registries for the consumption, integration, and republication of information.



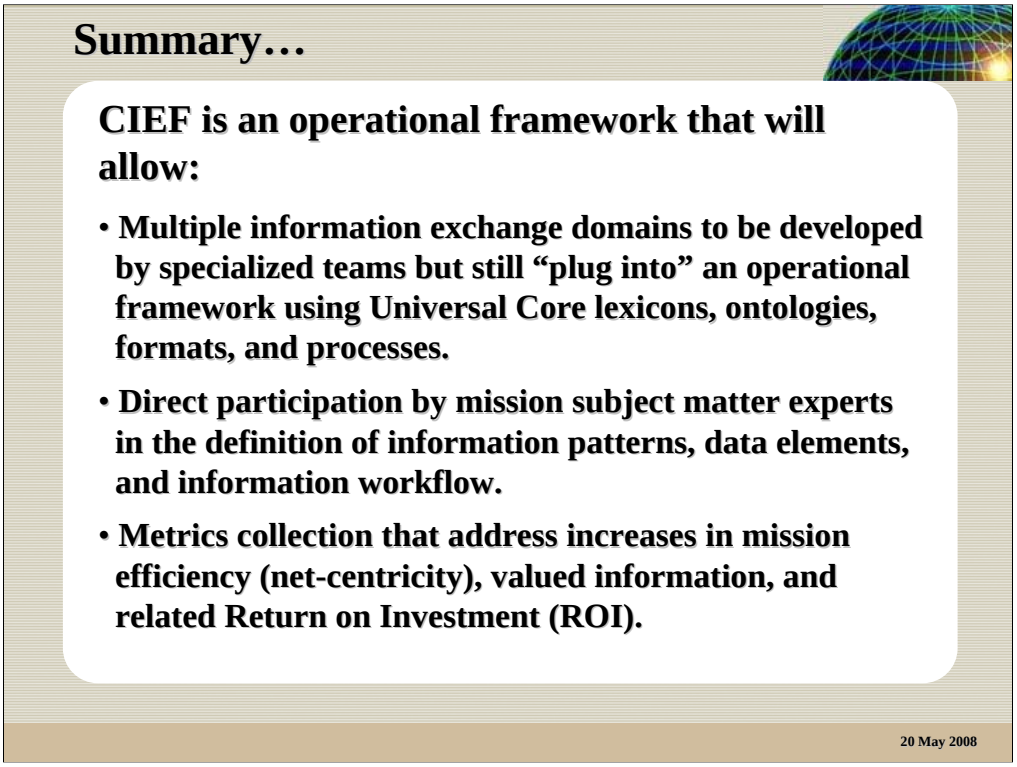
CIEF information is at the top of the "well define context" stack… The intent of the CIEF metadata registries is to provide enough contextual attributes to "sufficiently describe" the information content so it can be located and integrated into a cohesive body of knowledge to support mission tasking.

CIEF provides a process to get "close" to the correct information without any previous knowledge or pre-negotiation.



The CIEF data strategy is to incorporate all associated elements and processes for complete and unambiguous semantic mediation.

The goal is to address DoD requirements to support mission tasking… augment or automate tasks.



CIEF is both a road map and an architecture. CIEF addresses the underlying theories and technologies, and an implementation path.

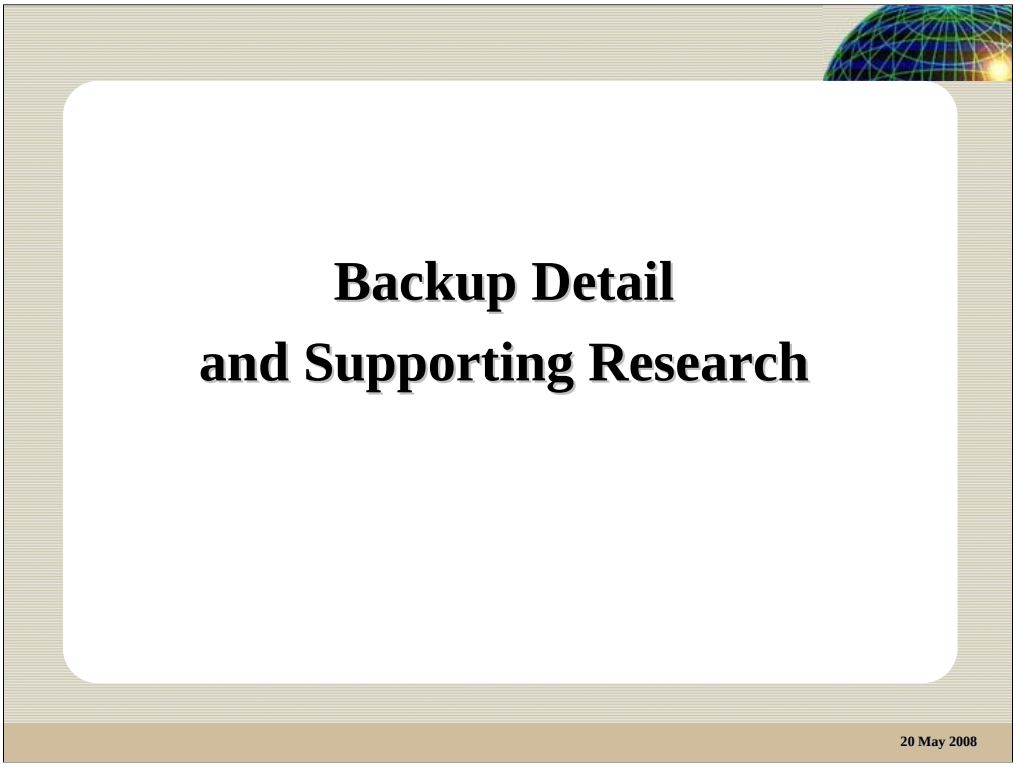
One of the primary advantages of CIEF is the direct participation of SME in supported mission areas. Only through SME will CIEF define clusters of information (data objects, patterns, etc) that "make sense." Further, defined processes will not be "thrown away," rather they will be analyzed and incorporated into an agreed upon shared methodology (part of the Universal Core).

CIEF registries will provide tiered of information processes currently available only in a piecemeal fashion… HOW INFORMATION IS USED IN A MISSION CONTEXT is the primary focus of CIEF.

CIEF will also provide information on: Valued Data Sources, Best Patterns, Activity Level of COIs, and many other metrics that will address the value and use of information… Data Strategy.

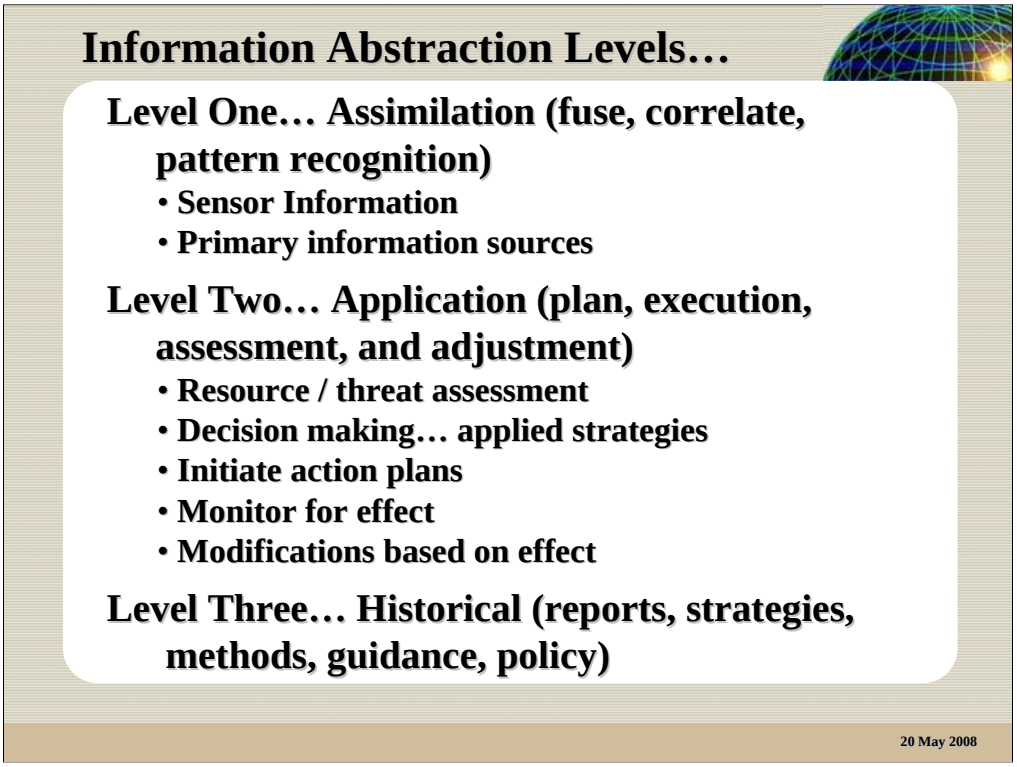
The release authority for all information is Paul Shaw, Paul.Shaw@navy.mil.

Specific technical questions can be directed to Dr. Roberts.



The following backup detail represent a small portion of CIEF research which is available upon request to any authorized Federal or DoD agency.

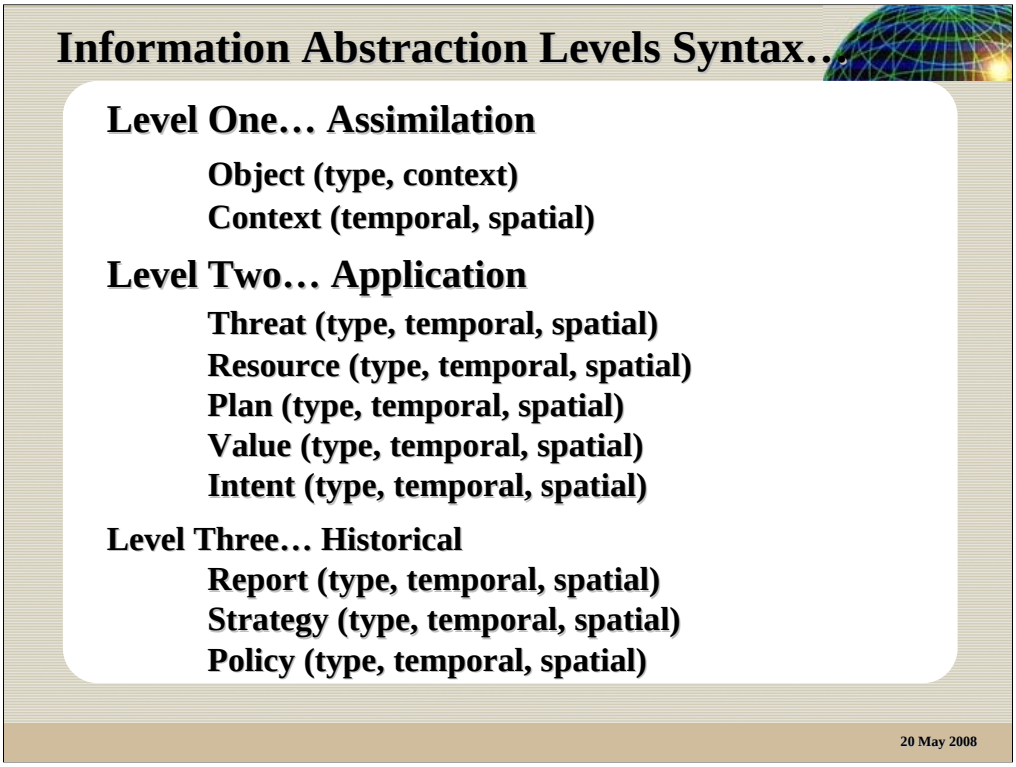
The release authority for all reports, briefs, or information is Paul Shaw, Paul.Shaw@navy.mil.



Level One are typically voluminous, direct feeds from primary data sources. The emphasis in processing Level One information is speed, identification and assessment of atypical attributes within data patterns, and the encapsulation of "raw" data for higher level processing.

Level Two is the primary "working" level for analysts and decision makers. Information from Level One is assessed for threat or other analysis criteria. Level Three information will also be integrated at this level to review past action plans and strategies.

Level Three can be thought of as a repository of past, and hopefully, best cases and subsequent strategies. Guidance and policy will also be maintained at this level through the processing of policy documents and organizational rules.



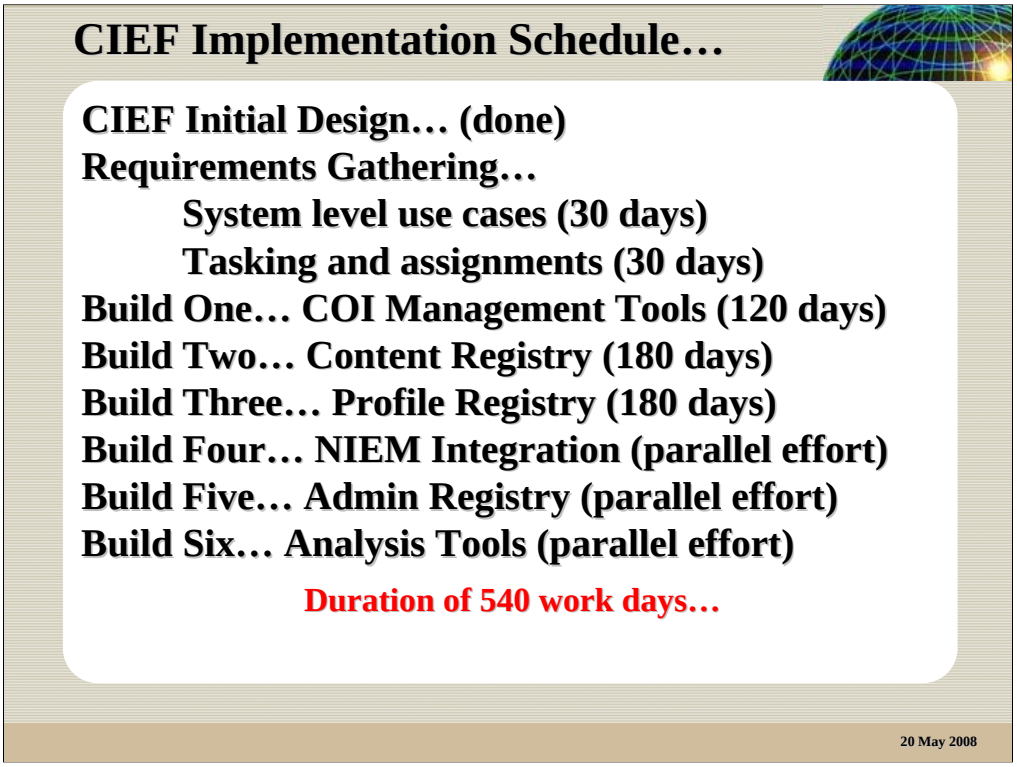
This slide represents three possible levels of abstraction of data objects to support mission areas within CIEF.

Level One is extremely concise in sub-class definitions with primarily mechanical and administrative processing. The primary emphasis of this level will be the processing of raw data from sensors, human intelligence, weather, and other primary sources of information. Note: Level One information may also include News Feeds and recombinant publishing of Level One information removing duplications, and automatically fuse data, categorize, or summarize information.

Level Two represents the processing of information by SME and intelligent automated processes (i.e., rule based systems) to be republished for use by decision makers and planners. Level Two may be republished with much of the detail encapsulated or presented with analysis evaluations (ranking, priorities, confidence limits, etc.).

Level Three information is historically derived (i.e., best cases), or based on accepted guidance or strategies.

Rather than design in the abstract, CIEF is based on a scaled rollout of well defined builds. Note: The sequence of the builds can be adapted to the sponsoring community.



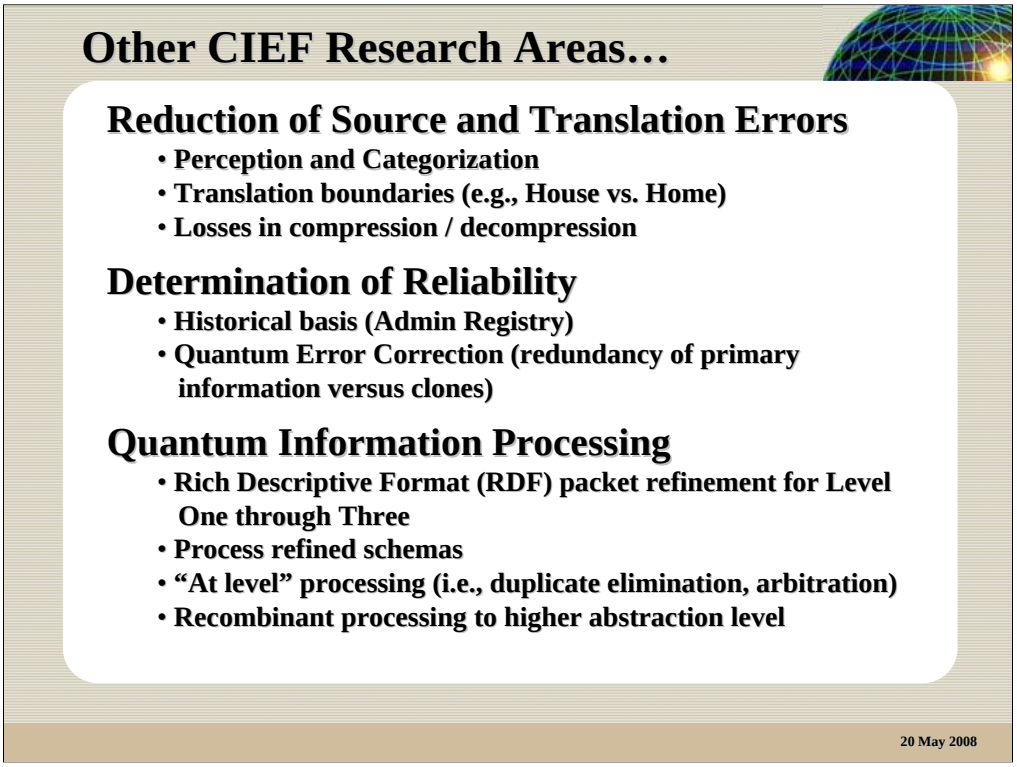
Additional information is available as an MS Excel spreadsheet to address work hours on tasks, rates, and a detailed schedule of deliverables.

Note: CIEF has been designed in a modular fashion that lends itself to multiple parallel efforts that can be integrated to a common baseline.

Members of the CIEF team have developed hundreds of software and hardware applications over the past thirty years from robotic painting systems to high security collaboration services for ten's of thousands of clients.

While the team processes may seem agile because of prior experience, the team uses proven DoD development and deployment methods. Prior efforts illustrate the well documented and controlled processes of the team.

The CIEF team will use a structured process to delineate system requirements and client expectations. Also, bi-weekly structured VTCs will ensure that the project is on track and within the defined objectives of the sponsor community. In addition to the standard DoD document set, monthly reports will summarize tasking, highlight action items, and provide a realistic view of progress to date. Further planning detail will be provided within the first thirty days of initial funding as: Project Charter, Project Schedule (MS Project), Spend Plan (MS Excel), and System Level Use Cases. The initial information will be summarized in task level quad charts.



Initially, simple topic maps and ontological classification schemes will be used in the CIEF deployment. A major goal will be in subsequent tasking to decrease source and channel errors and automate (to some degree) much of the "mechanical" processing of information.

The Activity Tracker 's historical tracking of information source and usage will be the initial method of determining the reliability of information within CIEF. Communications techniques such as Quantum Error Correction (QEC) based on primary versus cloned information will be used in subsequent CIEF builds.

A major research area of CIEF will be in the improvement of processing information packets (Quantum). RDF will be the primary representational method with a mapping to relational schema. The schema will be refined over successive builds to process both "at level" information and to support recombinant information publication.