



The Cross-domain Information Exchange Framework (CIEF)

20 May 2008

Paul Shaw, SPAWARSSYSCOM
Dr. David J. Roberts, iBASEt, Inc.
San Diego, CA

20 May 2008

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 SPAWARSSYSCOM, San Diego and a **roadmap for the implementation** of a shared core semantic objects and processes.

Outline...

- **Some of the Problems**
- **Definition: Shared Understanding**
- **CIEF Goals**
- **What is CIEF?**
- **CIEF Value Propositions**
- **CIEF Business Case Analysis (BCA)**
- **Why a Registry Architecture for CIEF?**
- **CIEF Query, Mediation, and Response**
- **CIEF Implementation Plan**
- **CIEF Tools**
- **Summary**

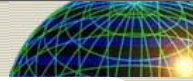
20 May 2008

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.

What Are Some of the Problems?...



- **You can find anything on the Internet... Somewhere in 43,256 hits (e.g., Google)**
- **What does the data mean?... Data doesn't mean the same thing to all people (lack of context)**
- **Information levels are mixed from raw to summarized, and of "varying" quality**

Bottom line: Web services and Internet based systems may be fine for buying and selling shoes, but do not encompass DoD mission requirements of: security, Quality of Service (QOS), semantic mediation, valued sources, etc. for information exchange.

20 May 2008

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.

Who's Definitions?...

“Shared Understanding”...

- **To the Engineer: a shared code/de-code model**
- **To the Computer Scientist: an Information Exchange Data Model (IEDM)**
- **To the Cognitive Scientist: shared semantic objects (content) in an ontological model (context)**

All discipline views have value and all must be part of any attempt to build “shared understanding.”

20 May 2008

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.

CIEF Value Propositions...

- **Provide Valued Information at the Right Time (VIRT*) and place**
- **De-clutter the workspace... one information view versus many**
- **Support information sharing based on “need to know” and Communities of Interest (COI)**
- **Provide a simple method to integrate additional information systems without “changing everything”**

* Dr. Rick Hayes-Roth, Naval Postgraduate School (NPS)

20 May 2008

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.

Related CIEF Goals...

- **Existing legacy information system must be leveraged in any new efforts (e.g., GIG)**
 - Use globally, but manage locally
 - Any new architecture must address a method of integrating everything from COBOL based systems to current web services (i.e., Software Development Tool Kits)
- **Accommodate constant change to the core data objects and associated schema (e.g., On Demand Schema (ODS))**
- **Value added at all project phases (< 6 mo.), not just “at the end”**

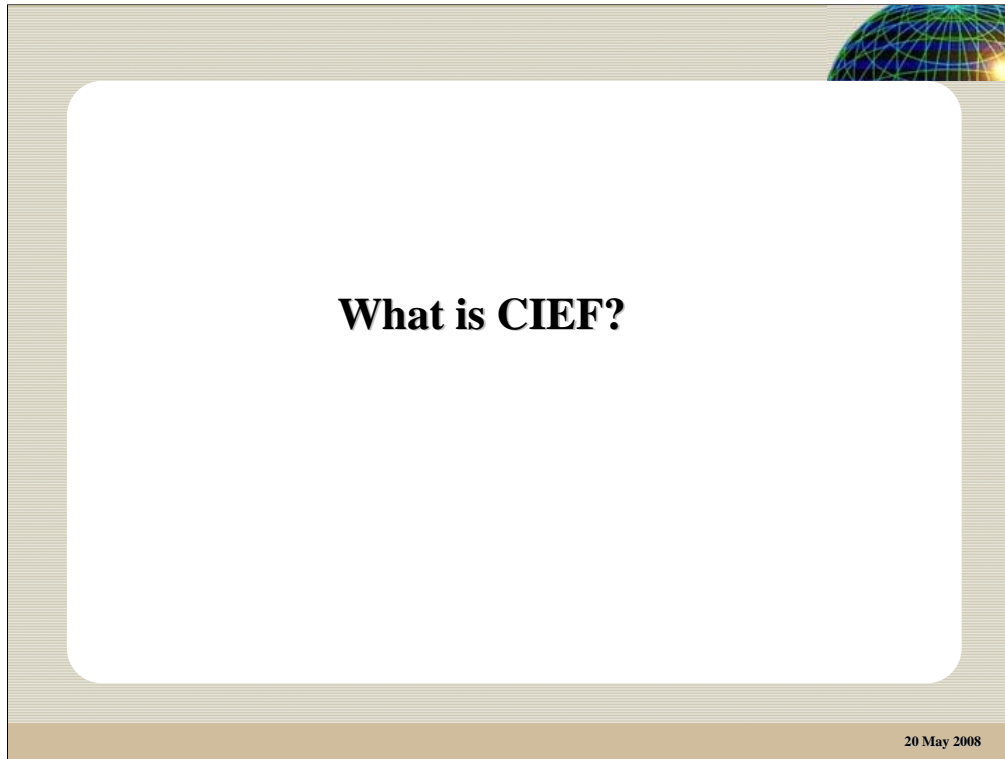
20 May 2008

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 in some Pollyannaic 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 CIEF Architecture...

- A universal framework that defines where the “pieces” go... focusing on a shared core (lexicon, syntax, process) of understanding.
- An architecture that describes information flow and usage based on mission tasking
- An implementation plan that builds on simple and testable constructs
 - Short value-based build cycles (<6 months each)
 - Defined success metrics for each build
 - Evolutionary complexity
- Augments existing information systems

20 May 2008

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.

CIEF's Strategic Enablers...

Three strategic enablers were called out in the *National Military Strategy to Combat Weapons of Mass Destruction*, 13 February 2006, Chairman of the Joint Chiefs of Staff:

The CIEF architecture supports all three strategic enablers:

- **Intelligence...** Directly supports strategy, planning, and decision-making; facilitates improvements in operational capabilities; and informs programming and risk management. p. 21
- **Partnership Capacity...** Building partnership capacity bilaterally and multilaterally enhances our capability to combat WMD. Incorporating our partners' and allies' combating WMD capabilities supports our ability to defend the homeland, deter forward, and conduct multiple, simultaneous activities. p. 22
- **Strategic Communications...** Complements combating WMD efforts and helps shape perceptions at the global, regional, and national levels. p.22

20 May 2008

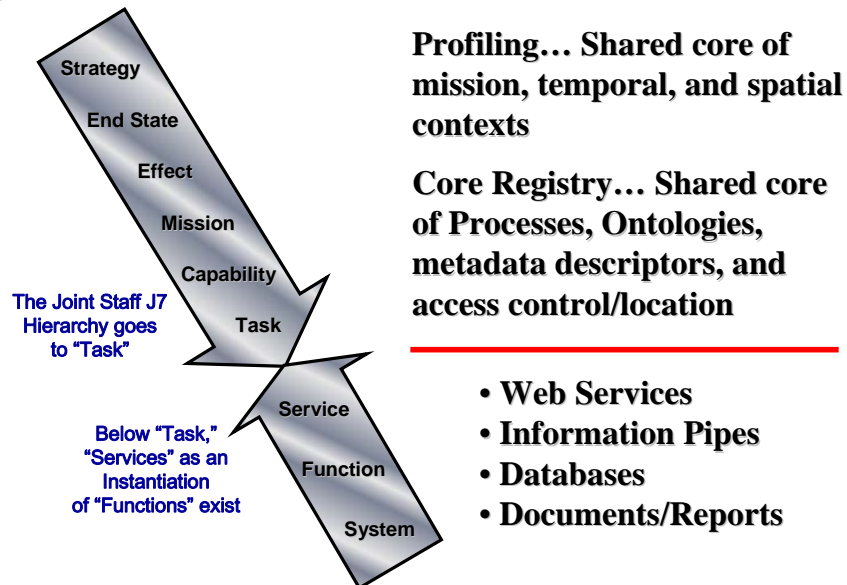
It is useful to think of new and evolving technologies as enablers... allowing us to solve complex problems that may have previously been “too hard”.

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 Focuses on the Mission Task Level...



20 May 2008

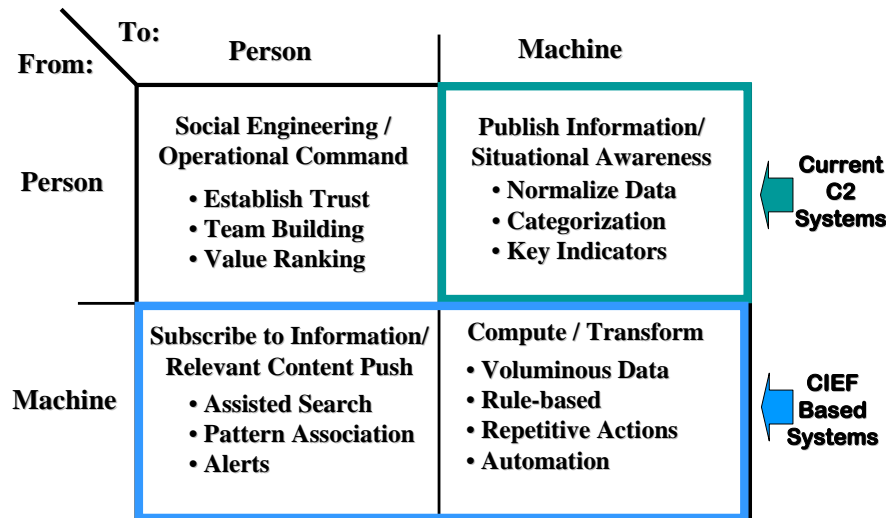
CIEF bridges the Service Oriented Architecture (SOA) gap between mission oriented tasks and relevant services and data sources.

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 Addresses Information Flow...



20 May 2008

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)

CIEF Information Flow... P2P

Social Engineering / Business Rules

- Establish Trust – through certified authentication methods and published “resumes”
- Team Building – through Community of Interest (COI) tools to invite participation, and to control both access and re-use of information
- Value Ranking – local and published authoritative sources, information traffic patterns, and strategies to reflect information value

The social engineering can be “left to happen” or it can be assisted through supporting tools and processes.

20 May 2008

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).

CIEF Information Flow... M2P / P2M

Publish Information (P2M)

- **Normalize Data** – Universal Core lexicon and syntax
- **Categorization** – Universal Core ontologies
- **Key Indicators** – Understanding of the prioritization (business rules) of data objects (Core Patterns)

Subscribe to Information (M2P)

- **Assisted Search** – Complexity is abstracted
- **Pattern Association** – Core Patterns in context
- **Alerts** – Based on objectives and limits within the mission

Supporting tools are needed to machine processing.

20 May 2008

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)

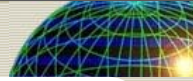
Two major focus areas: 1.) the formalization of information semantics and syntax; and 2.) tools to assist in that process.

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.

CIEF Information Flow... M2M



Compute / Transform

- **Voluminous Data** – Too much information with too many attributes to be processed by humans
- **Rule-based** – Interactions are non-ambiguous and based on understood processes
- **Repetitive Actions** – Decisions and actions are deterministic and can be re-constructed.

People skills and machine processing should be view as complementary... Tools that abstract complexity (e.g., publish content and context to metadata registry) assist information flow.

20 May 2008

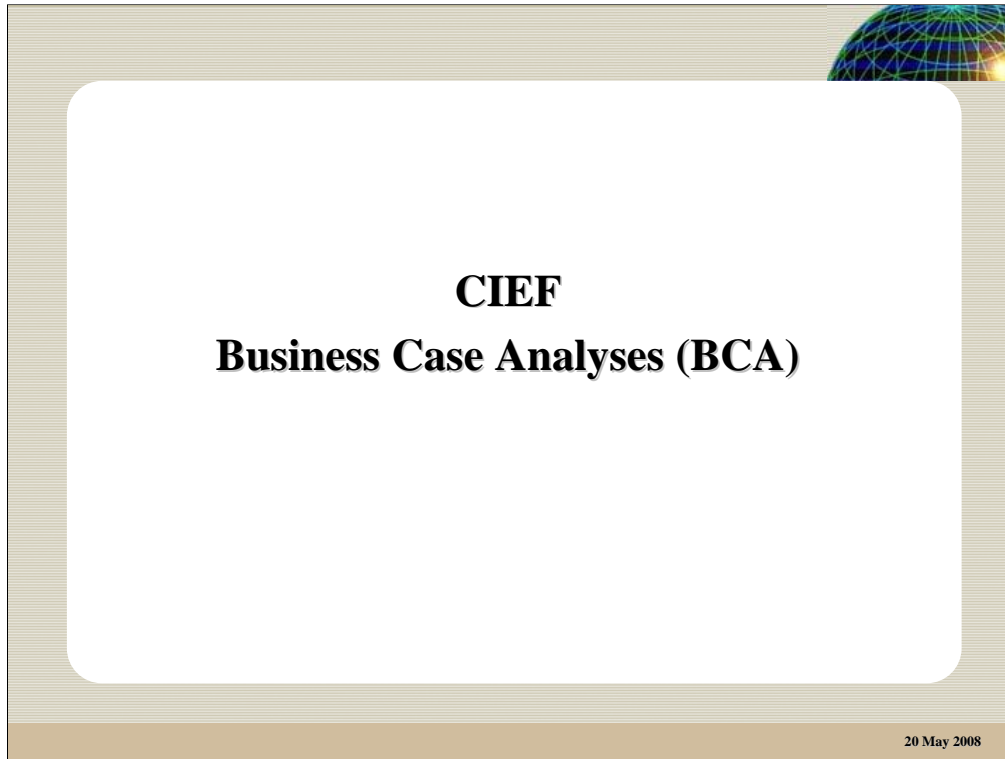
Machine to machine... heavy lifting (search, sort, merge, pattern recognition, detection, alert)

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...

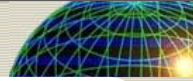


The following Business Case Analyses (BCA) address obvious areas of value in the deployment of CIEF. Secondary, or more subtle areas of value were not addressed but should be considered in the deployment schedule. For example, what is good information in a timely manner worth? What is the impact of no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

Business Case Analyses (BCA)...



- **The General Need**
- **DoD Security Model**
- **Support COI Business Processes**
- **Built-in Training**
- **Support Service Orientation**
- **Access Model**
- **Leverage Current Technologies**
- **Meets National Security Needs**

20 May 2008

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.

BCA: The General Need*...

- Ability to form ad hoc COI based on the situation or business case and exchange Valued Information at the Right Time (VIRT)
- Ability to exchange information across large communities (i.e., Homeland Security, MDA)
 - Shared tools for publishing and finding information with DTIC distribution codes, and need to know enforcement
 - Information “tagged” with topic, time, location, priority, and other context attributes
- Architecture to support cross-domain analysis
 - Detection
 - Linking information
 - Out-of-the-ordinary* analysis

**Out of the Ordinary*, Rand Report, 2004

20 May 2008

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.

BCA: DoD Security

DoD Security Model

- Based on current DoD guidance, but “invisible” to users (e.g., DTIC Distribution Accessibility Codes (DAC))
- Access via CAC and PKI (no passwords)

Assumptions: Security breaches are costly in terms of determination of the damage, containment of the damage, risk mitigation, and reassessment of the security environment.

CIEF Advantage: CIEF conforms to all DISA certified procedures, and also conforms to DoD policy for the handling and distribution of technical material using DTIC distribution and accessibility codes attached to all documents and information. CIEF will not permit the “accidental” distribution of information to an incorrect environment, but will assist publication of information through a certified process (Public Affairs Office release, or other authority).

20 May 2008

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.).

BCA: Support Business Process...



Support the COI Business Processes

- Customizable tool selection for each COI
- User case based tools (efficiency tested)

Assumptions: Not all COI operate in the same fashion or have the same requirements for support information technology tools. To force all COI to use the same set of tools and procedures would reduce COI productivity.

CIEF Advantage: The COI manager determines what tools are selected from the total offering of CIEF tools (e.g., group calendar, threaded discussion groups, etc.). Tools can be deleted or added at any time by the COI manager. Thus, individual CIEF COI can streamline their business processes from a selection of universal tools.

20 May 2008

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).

BCA: Built-in Training

Build-in Training

- On-line Help
- On-line User Manual
- CBT available
- FAQ's posted and kept current

Assumptions: Help desks and training for larger populations are normally offered on a 7/24 basis and are a major operational cost.

CIEF Advantage: CIEF has embedded context sensitive help (pop-up balloons), on-line user manual, and a training CD-ROM. The training philosophy of CIEF is to offer a thick layer of “self help” to reduce assisted help (e-mail and phone).

Note: Recent problems or questions are also posted in a Frequently Asked Question (FAQ) area for client search.

20 May 2008

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.

BCA: Service Orientation...

Service Orientation

- Fee for services model is supported based on usage and access required
- Connects to other services easily

Assumptions: Purchasing, maintaining, and providing operational support for hardware requires a production staff with dedicated recurring costs. Economies of scale are difficult to obtain and accounting (who used what) is complex.

CIEF Advantage: Deploying the application at a DISA certified facility leverages true economies of scale. As a service, costs can be controlled and even reduced as the user populations grow.

Additionally, services oriented applications offer standardized protocols for interoperability and integration with other services.

20 May 2008

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).

BCA: Access Model...

Access Model

- Always available (runs in secure DMZ)
- Master directory of all COI and members
 - For routing of information / documents
 - Forming interest groups based on focus areas
 - Search across multiple COI and Domains for information (with no privacy violation)

Assumptions: Existing directory services do exist but do not address current business needs (COI membership, Special Interest Groups (SIG), etc.).

CIEF Advantage: The CIEF directory allows COI and individuals to connect on topical interests and focus areas. The CIEF Security Model also controls the “visibility” of information and associated access rights.

20 May 2008

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.

BCA: Leverage Current Technologies...

Leverage current technologies

- Registry / Repository architecture
- Semantically “aware” directory (profile, process models)
- MS Framework 3.0 brings millions of lines of tested and secure code
- Certified Oracle 11g to Framework 3.0 API
- Publication / Subscription Model for information distribution (smart push of information)
- CIEF generated Information Streams (IS) based on Rich Descriptive Framework (RDF)... next evolution in XML

Assumptions: The build team has deployed the past SPAWAR VPO, and several cutting edge new systems (TXC, VKR, SPIDER).

CIEF Advantage: The CIEF team will be many times more productive than “startups” because of their experience level.

20 May 2008

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.

BCA: Meets National Security Needs...



Meets National Security Needs per Rand Monograph, “Out of the Ordinary”, 2004

- Ability build ad hoc COI and SIG
- Ability to set up streams of information between COI, SIG, domains, and cross-domains
- Support “just enough” security model for first responders and special context
- Ability to combine disparate information between disparate sources (publication / subscription model)

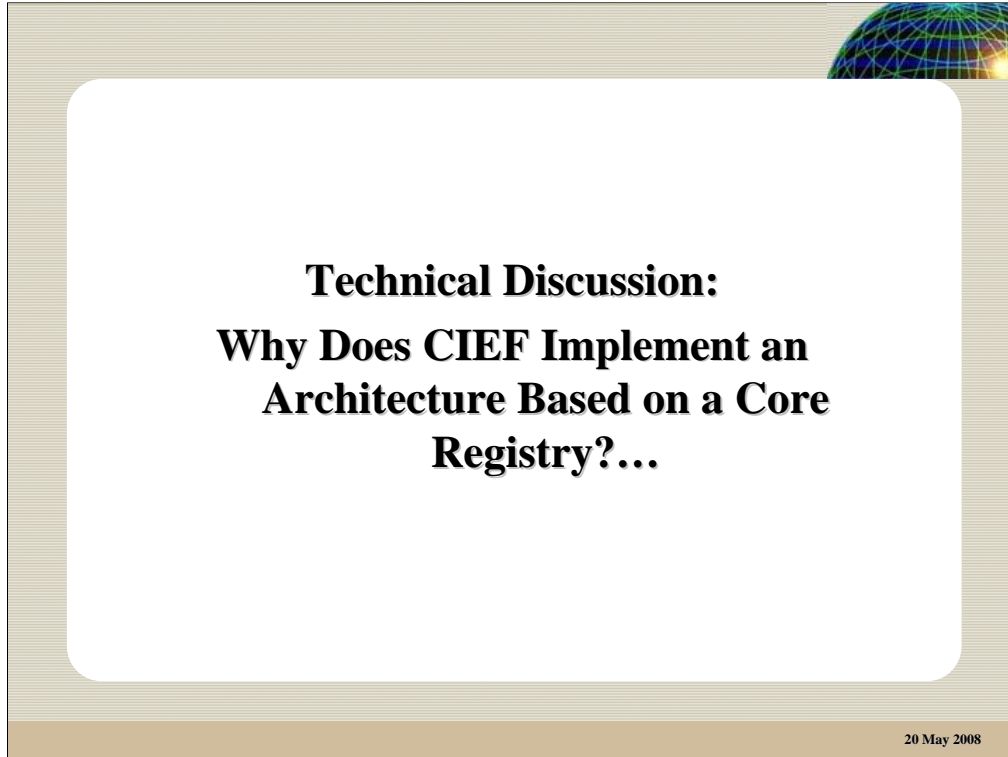
Assumptions: There is no one single system or method for building COP of disparate pieces of information.

CIEF Advantage: CIEF allows interest groups to work together and “publish” Information Streams (IS) or web services. Powerful semantic tools assist in the collection of information.

20 May 2008

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

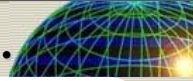
Thus, CIEF’s design was driven by national mission requirements, rather than abstract or general requirements.



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.

Why a Registry Based Architecture?...



Typical IEDM:

- Central core of shared lexicon, schema, and processes
- Pre-negotiation of all exchange formats and processes

Pro's:

- Easy to model to real world business processes

Con's:

- **Scaling difficult because of increased complexity of data objects and record density**
(See next slide)

CIEF:

- Central core of shared registries
- Data is described centrally, but exposed locally in share formats (RDF/RDFS)
- Supports dynamic processes and semantic models*

Pro's:

- Scales to meet real-time QOS
- Complexity is hidden
- Local control of data is maintained

Con's:

- Abstraction and encapsulation of data objects added tasks

*Several biological models are used within CIEF, to include "sleep" functions of value assessments based on usage, error correction, convergence analysis, etc.

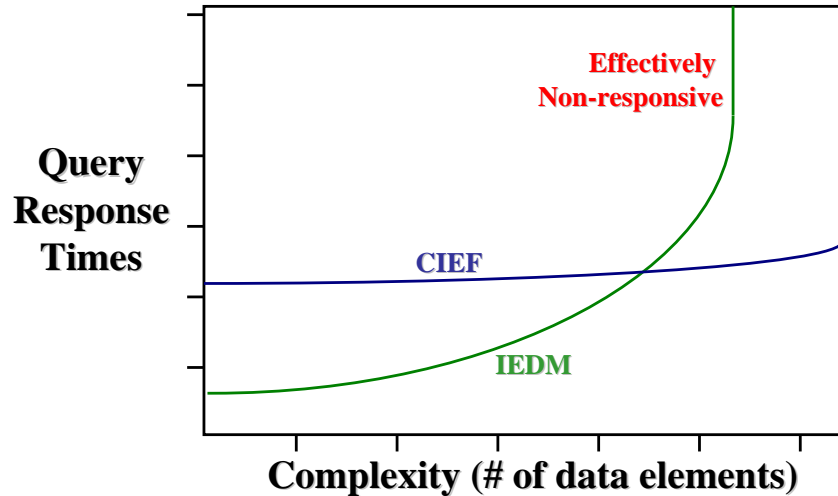
20 May 2008

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).

Why a Registry Based Architecture?...

CIEF query response times will be initially higher because of the negotiation and filtering protocols, but scales to any complexity level



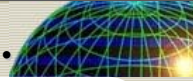
20 May 2008

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.

Why a Registry Based Architecture?...



CIEF registries support the different functionality in the storage and the recovery of information:

- **Storage – Taxonomic classification (e.g., Dewy Decimal System), tagging of all relevant meta data, updating of indices**
- **Recovery – Functional classification (e.g., Card Catalog), key search parameters, task orientation**

The complexity of information exchange is decomposed into a serial process (CIEF scales to any complexity):

- **Initial centralized search**
- **Discrimination of relevancy (topic, time, location)**
- **Connection information to source location provided**
- **Dynamic updates to format and lexicon from central URI**
- **Point to point information exchange... network matrix**

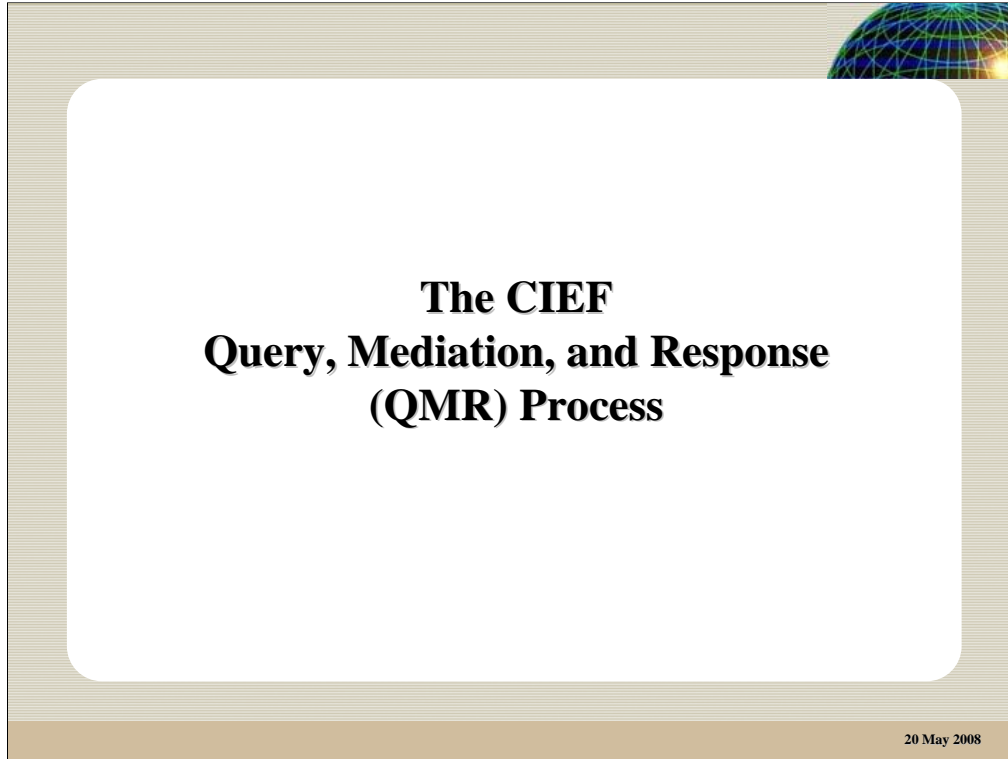
CIEF augments existing information systems... not a replacement!

20 May 2008

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.

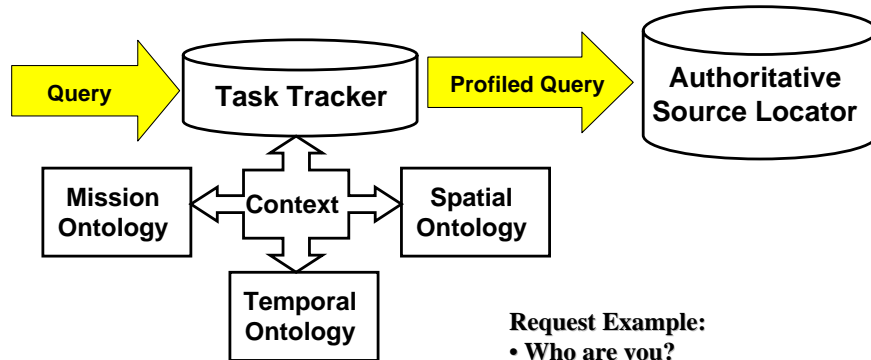
The retrieval of information in CIEF is mission/task based and has uses very different set of attributes to recall information (e.g., Library Card Catalog). This orientation takes into consideration the mission tasking in determining the context of the information request.



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

CIEF QMR Process...

Query Profiled



Request Example:

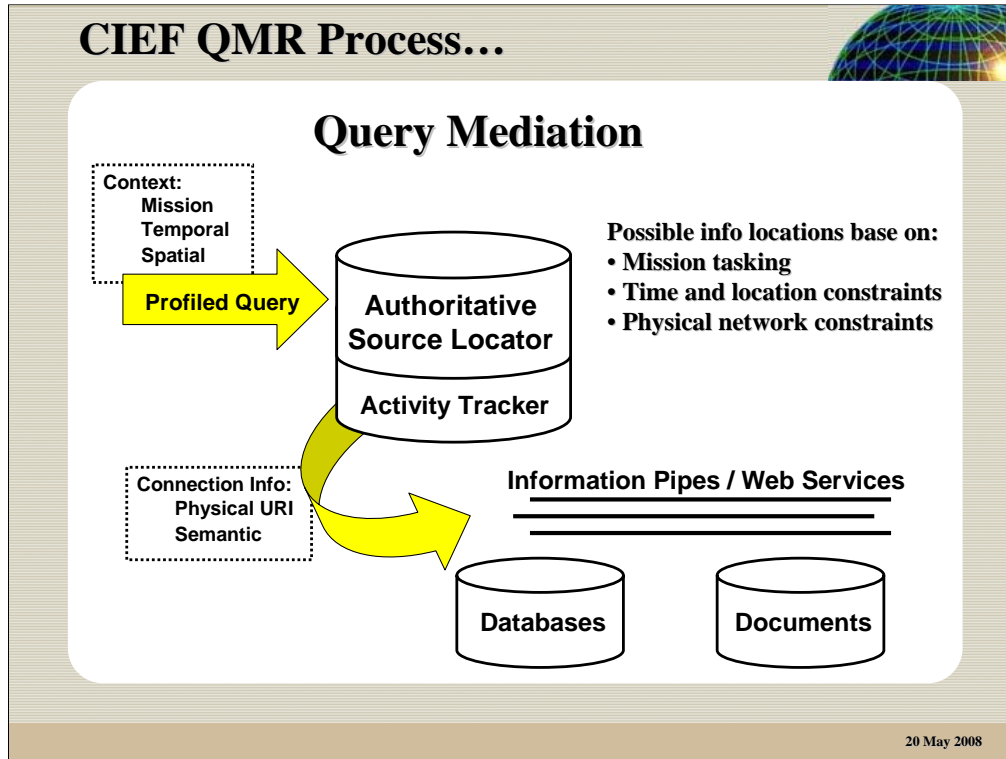
- Who are you?
- What are you doing (mission)?
- When (urgency of information)?
- Where are you (geo-location)?

20 May 2008

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.

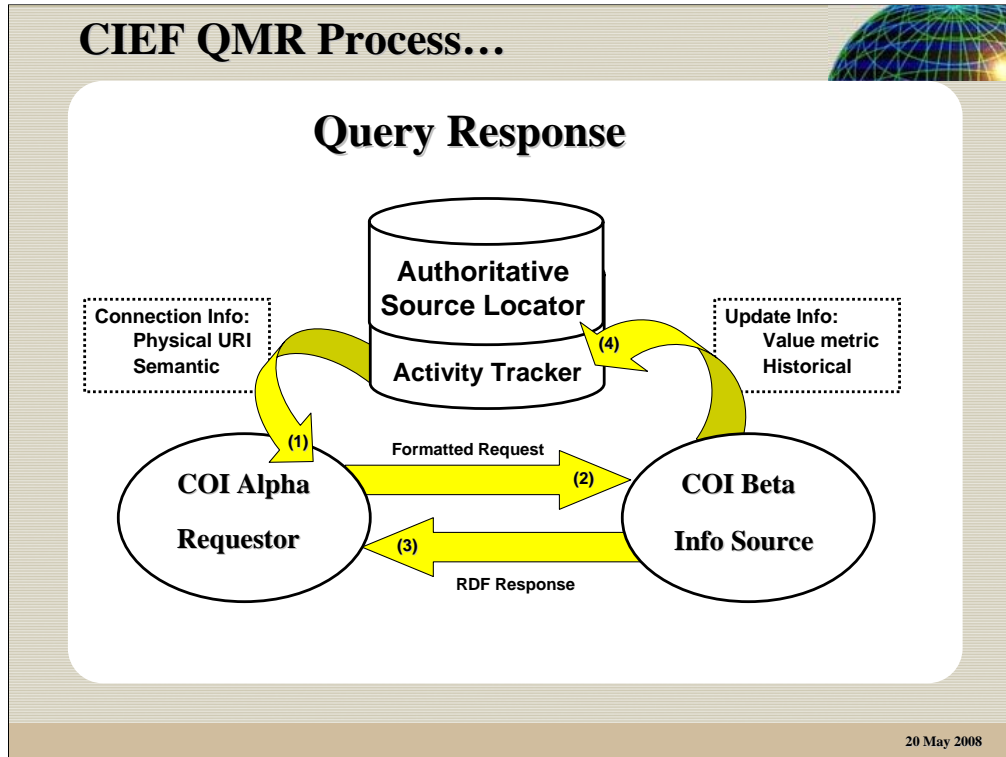
CIEF QMR Process...



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.

CIEF QMR Process...




20 May 2008

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”).**

Query By Example (QBE) GUI...

Name (ID)	<input type="text"/>	
Information Level	<input type="text" value="▼"/>	(Raw, Working, Historical) *
Domain Category	<input type="text" value="▼"/>	(Mission Ontology)
Functional Category	<input type="text" value="▼"/>	(Task Ontology)
Temporal Constraints	<input type="text" value="▼"/>	(Temporal Ontology)
Spatial Constraints	<input type="text" value="▼"/>	(Spatial Ontology)
Key Indicators	<input type="text" value="<More>"/>	* Detail dynamically provided as higher level selections are made.
	<input type="text" value="<More>"/>	
	<input type="text" value="<More>"/>	
Posting Date	<input type="text" value="System provided"/>	
COI	<input type="text" value="Default provided"/>	
Owner	<input type="text" value="Default provided"/>	
Distrib Access Code	<input type="text" value="Default provided"/>	
Information Location	<input type="text"/>	<Browse>



20 May 2008

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.

Query Example...

Name (ID)	<input type="text" value="*"/>	<p>Note: This is a request for Working information. Further Level One, Raw information, or Level Three, Historical information could also be requested.</p>
Information Level	<input type="text" value="Working"/>	
Domain Category	<input type="text" value="*"/>	
Functional Category	<input type="text" value="Possible Threat"/>	
Temporal Constraints	<input type="text" value="Immediate"/>	
Spatial Constraints	<input type="text" value="Local, OKC"/>	
Key Indicators	<input type="text"/>	<More>
	<input type="text"/>	<More>
	<input type="text"/>	<More>
Posting Date	<input type="text" value="=< 22 August 2007"/>	
COI	<input type="text"/>	
Owner	<input type="text"/>	
Distrib Access Code	<input type="text"/>	
Information Location	<input type="text"/>	<Browse>

20 May 2008

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

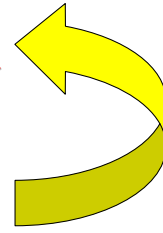
The query can be refined or broadened per the client's search criteria.

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.

Response Example...

Name (ID)	Mary A. Smith
Information Level	Working ▼
Domain Category	Monitor ▼
Functional Category	Possible Threat ▼
Temporal Constraints	Immediate ▼
Spatial Constraints	OK, OKC ▼
Key Indicators	Hazmat driver license <More>
	Terrorist affiliation <More>
	Medical degree <More>

Three information sources found



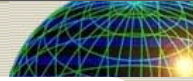
20 May 2008

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.

Response Example (continued)...

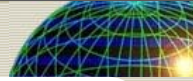


Name (ID)	Mary A. Smith
Information Level	Working ▼
Domain Category	Monitor ▼
Functional Category	Possible Threat ▼
Temporal Constraints	Immediate ▼
Spatial Constraints	OK, OKC ▼
Key Indicators	Hazmat driver license <More>
	Terrorist affiliation <More>
	Medical degree <More>
Posting Date	20 August 2007
COI	National Watchlist
Owner	Frank Jones, Spec Agent
Distrib Access Code	DTIC B, Fed only
Information Location	POI Report_7145_OKC <Browse>

20 May 2008

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

Response Example (continued)...



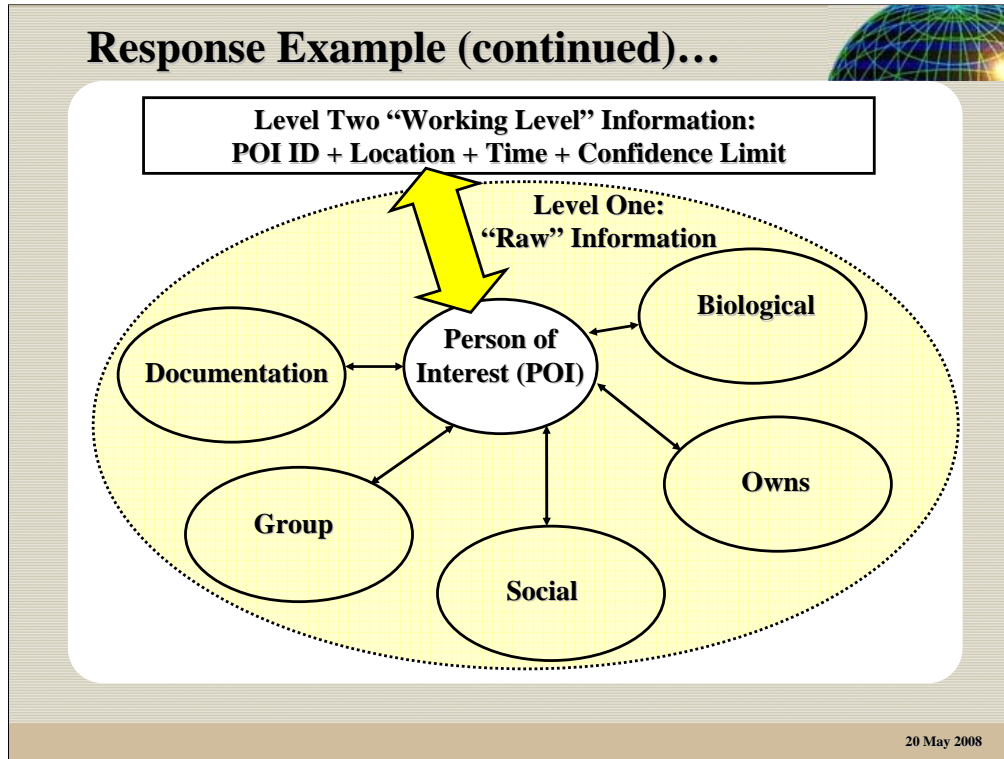
Name (ID)	<input type="text" value="Mary A. Smith"/>	
Information Level	<input type="text" value="Raw"/>	
Domain Category	<input type="text" value="*"/>	▼
Functional Category	<input type="text" value="*"/>	▼
Temporal Constraints	<input type="text" value="*"/>	▼
Spatial Constraints	<input type="text" value="*"/>	▼
Key Indicators	<input type="text"/>	<More>
	<input type="text"/>	<More>
	<input type="text"/>	<More>
Posting Date	<input type="text" value="*"/>	
COI	<input type="text"/>	
Owner	<input type="text"/>	
Distrib Access Code	<input type="text"/>	
Information Location	<input type="text"/>	<Browse>

A Back Sweep of all "Raw" information on Mary A. Smith can be requested. Note "*"

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

20 May 2008

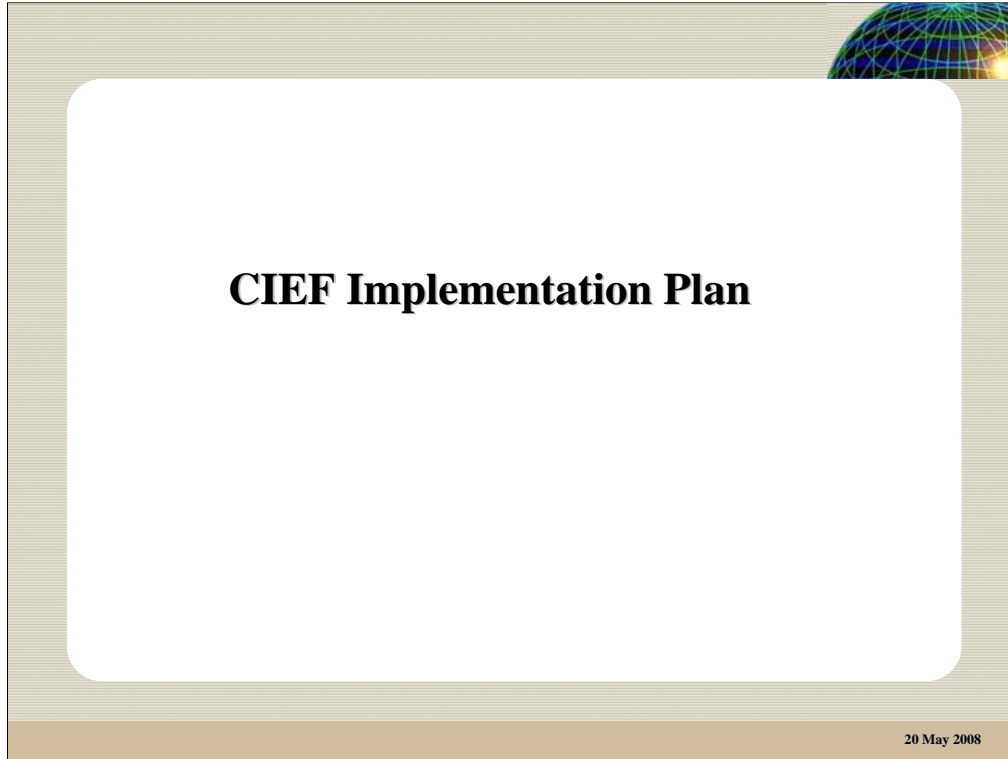
Response Example (continued)...



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.

CIEF Implementation Plan...

Step #1. Start with standard set of tools to setup and manage Communities of Interest (COI)

Step #2. Build a centralized method to register (publish) and search for information

- Profile registry (provides context of the request)
- Content registry (ontological categorization)
- Admin registry (usage stats, queuing, routing)

Step #3. Develop tools to assist in the publication and subscription to information

- “Wrap” and augment existing information systems
- Based on mission threads/processes
- Integrate or modify existing tools

20 May 2008

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.

CIEF Implementation Plan (Detail)...



Step #1. Standard set of tools to setup and manage a COI / Domain:

- a. Support ANY collaboration suite... SharePoint, CollabNet, Oracle Collab Suite (vendor agnostic)
- b. Create an initial version of the central information locator registries (system high security)
 - Universal temporal lexicon/format
 - Universal spatial lexicon/format
 - Universal mission / task ontologies
- c. Create web appliqué (secure web service) to share information (assists in publishing information metadata in the Universal Core lexicon and format)

20 May 2008

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.

CIEF Implementation Plan (Detail)...



Step #2. Extend the centralized method to register (publish) and search for information:

a. Task Tracker... Contextual Constraints

- Authenticated Name
- Access Rights (COI, SIG, Individual)
- Mission (Type, Temporal, Spatial)

b. Authoritative Source Locator... Metadata

Descriptors

- Access Level, Functional Level, Mission Elements (context),
- Resource Location and Format

c. Activity Tracker... Operational Constraints

- Resource value (authoritative source)
- Availability (responsiveness)
- Reliability (historical)

20 May 2008

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**.

CIEF Implementation Plan (Detail)...



Step #3. Develop tools to assist in the publication and subscription to information:

a. “Wrap” or augment existing information systems

- Tool to describe local information with published taxonomies in Content Registry (common meta data)
- Register and publish web services of local information
- Register and publish Information Pipes (modified news feeds)

b. Based on mission tasking / processes

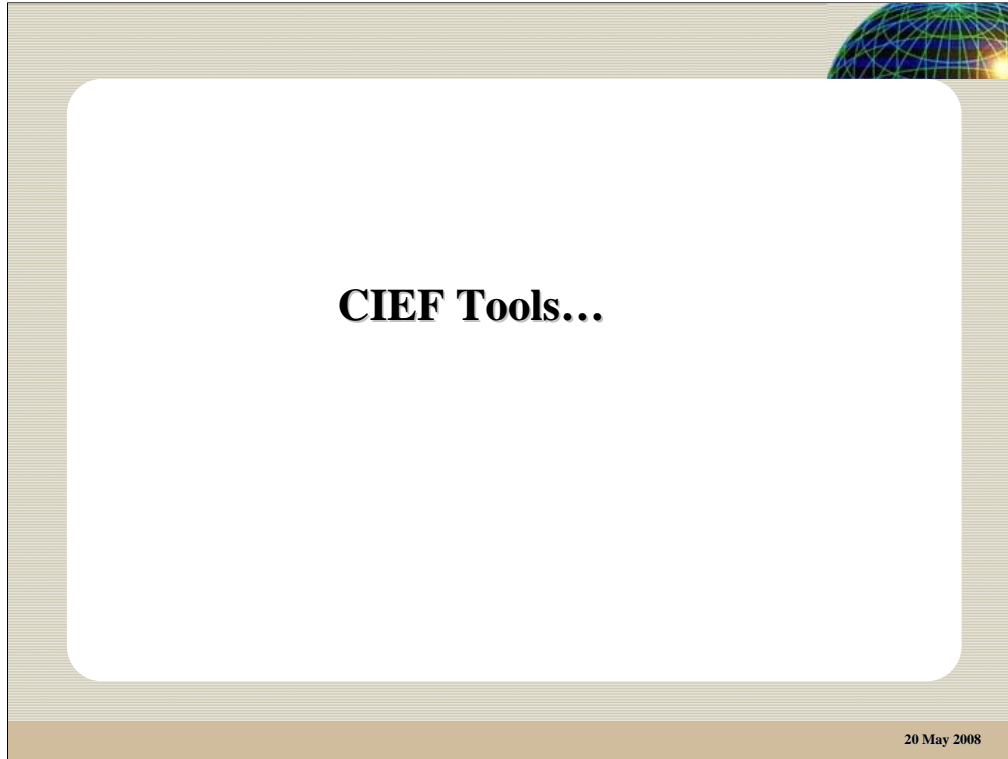
- Subject Matter Experts (SME) determine key mission elements and processes
- Interfaces that “hide” complexity, and assist or automate mission tasking

c. Integrate existing tools... wikis, visualizations, etc.

20 May 2008

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...

CIEF Tools...

Example of the adaptation of the RSS 2.0 news feed format for use in CIEF:

Channel elements:

Language	DoD Encryption Method / Level
Managing Editor	Registered CIEF COI Manager
Web Master	Registered CIEF Domain Manager
Category	Universal Core Information Level ontology
Docs	DTIC DAC (Distrib Accessibility Code)
Cloud	Universal Core Domain (mission) ontology

Items:

Category	Universal Core Capability (task) ontology
Guid	Universal Core unique identifier

RSS Extensions:

Priority	Universal Core Temporal ontology
Geo-location	Universal Core Spatial ontology

20 May 2008

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)**

CIEF Tools...

RSS 2.0 modification (re-use) example:

	RSS 2.0	Info Stream 1.0 (CIEF adapted)
XML/RDF Based...	Yes	Yes
Semantic Orientation...	Yes	Yes
Centralized Registry...	Partial	Yes
Public Format...	Yes	Partial
Commercial Reader...	Yes	Yes
Analysis Tools...	Partial	Yes
Re-publication Tools...	Yes	Yes
DoD Oriented Schema...	No	Yes
- Mission Context...	No	Yes
- Time Context...	No	Yes
- Geo-location Context...	No	Yes
DoD Authentication...	No	Yes
DoD Encryption...	No	Yes
DoD Access Control...	No	Yes

20 May 2008

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 management systems, databases, etc.) was demonstrated in SPAWARSCOM’s Virtual Knowledge Repository (VKR) in 2003. Note: VKR was built on semantic technology.

CIEF Tools...

With minor modifications, information sources for CIEF can include:

- **Technology eXchange Clearinghouse (TXC) – Technology assessment application that offers a web service. Uses DON ontologies for the classification of information.**
- **Virtual Knowledge Repository (VKR) – Uses semantic technology to reprocess information from hundreds of sources into topic and key word clusters.**
- **Or any web service enabled application, such as the Information Factory (following slide)...**

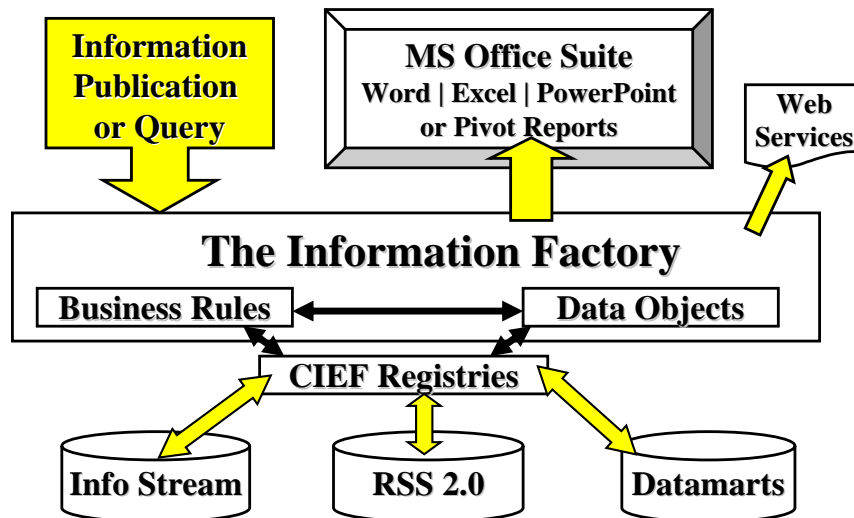
20 May 2008

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).

CIEF Tools...

Example of a publication and subscription tool



20 May 2008

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.

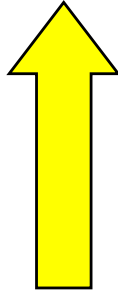
With the additional step of registering the Factory’s web services and describing metadata content, the Factory could be used as a shared information publication tools.

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

CIEF Tools Support...

Shared Understanding

Published Context



No Shared Context

Poor Understanding

CIEF tool assistance – Abstract metadata and context to publish locally and register in a central registry

News Feed RSS 2.0 – Structured content in public information domains

Web Services/UDDI – Negotiated data and methods

Reports – Structured domain specific content publication

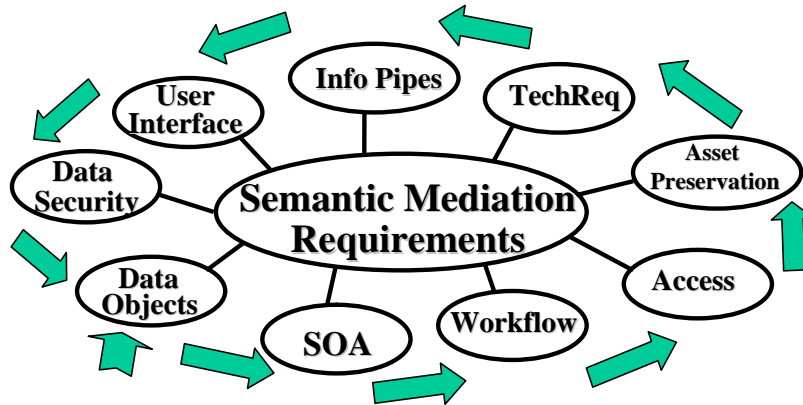
Documents – Free form or unstructured content

20 May 2008

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.

Other CIEF Focus Areas...



A successful data exchange strategy is built around a core of many elements and processes.

20 May 2008

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.

Summary...

CIEF is an operational framework that will allow:

- **Multiple information exchange domains to be developed by specialized teams but still “plug into” an operational framework using Universal Core lexicons, ontologies, formats, and processes.**
- **Direct participation by mission subject matter experts in the definition of information patterns, data elements, and information workflow.**
- **Metrics collection that address increases in mission efficiency (net-centricity), valued information, and related Return on Investment (ROI).**

20 May 2008

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.

Points of Contact...



Program Manager:

Paul Shaw, Paul.Shaw@navy.mil

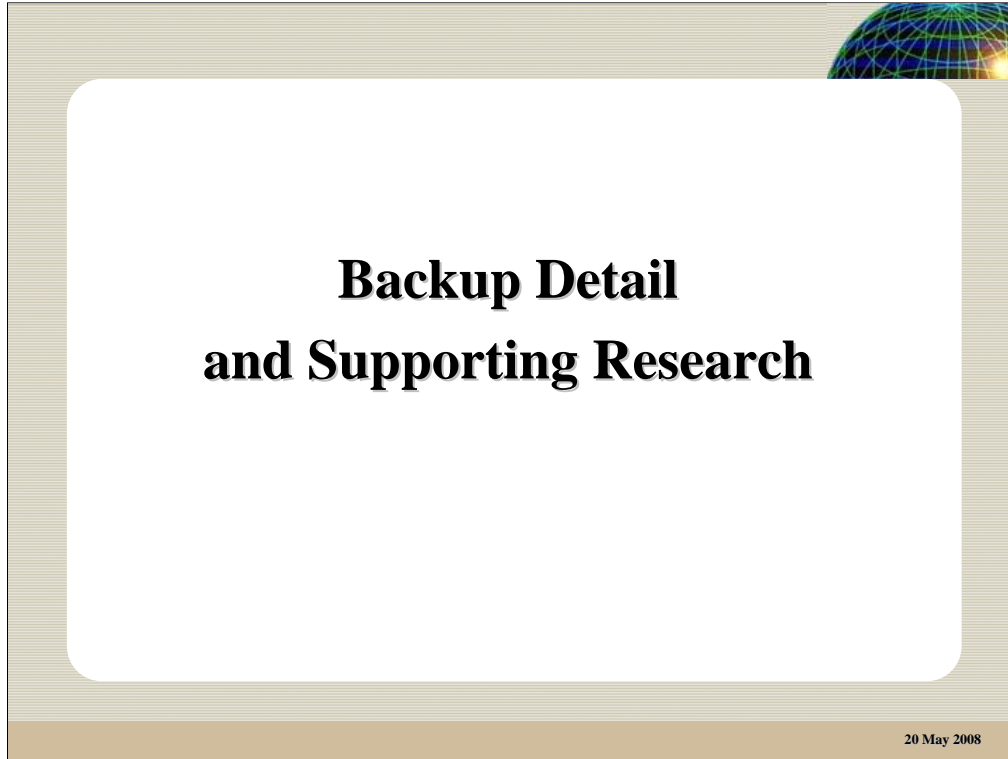
Chief Scientist:

Dr. David J. Roberts, [droberts@ibaset.com](mailto:drobot@ibaset.com)

20 May 2008

**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.

Information Abstraction Levels...



Level One... Assimilation (fuse, correlate, pattern recognition)

- **Sensor Information**
- **Primary information sources**

Level Two... Application (plan, execution, assessment, and adjustment)

- **Resource / threat assessment**
- **Decision making... applied strategies**
- **Initiate action plans**
- **Monitor for effect**
- **Modifications based on effect**

Level Three... Historical (reports, strategies, methods, guidance, policy)

20 May 2008

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.

Information Abstraction Levels Syntax..

Level One... Assimilation

Object (type, context)
Context (temporal, spatial)

Level Two... Application

Threat (type, temporal, spatial)
Resource (type, temporal, spatial)
Plan (type, temporal, spatial)
Value (type, temporal, spatial)
Intent (type, temporal, spatial)

Level Three... Historical

Report (type, temporal, spatial)
Strategy (type, temporal, spatial)
Policy (type, temporal, spatial)

20 May 2008

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.

CIEF Implementation Schedule...



CIEF Initial Design... (done)

Requirements Gathering...

System level use cases (30 days)

Tasking and assignments (30 days)

Build One... COI Management Tools (120 days)

Build Two... Content Registry (180 days)

Build Three... Profile Registry (180 days)

Build Four... NIEM Integration (parallel effort)

Build Five... Admin Registry (parallel effort)

Build Six... Analysis Tools (parallel effort)

Duration of 540 work days...

20 May 2008

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.

The CIEF Team...

Professional team members

- Oracle certification
- Authorities in the field of Semantic Technology
- Proven track records

Team builds since 2000:

- SPAWAR Virtual Program Office (VPO)... 22,000 users
- SPAWAR Business Opportunity Page... 2,000 vendors registered in the bid and solicitation system
- SPAWAR Knowledge Center (SKC)*... 2,500 users
- Technology eXchange Clearinghouse (TXC)... Technology assessment tool for DON CIO's eBusiness effort
- Virtual Knowledge Repository (VKR)... Semantic integration tool for DON CIO's eBusiness effort
- The Information Factory... Publication and analysis tool developed for NAVSEA Logistics Center, Fairfield

* Prime contractor on 2006 update

20 May 2008

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.

Other CIEF Research Areas...

Reduction of Source and Translation Errors

- Perception and Categorization
- Translation boundaries (e.g., House vs. Home)
- Losses in compression / decompression

Determination of Reliability

- Historical basis (Admin Registry)
- Quantum Error Correction (redundancy of primary information versus clones)

Quantum Information Processing

- Rich Descriptive Format (RDF) packet refinement for Level One through Three
- Process refined schemas
- “At level” processing (i.e., duplicate elimination, arbitration)
- Recombinant processing to higher abstraction level

20 May 2008

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.