Data Standards For Data Interoperability

Dr. Scott Renner
sar@mitre.org

Presented at *Critical Issues In C4I,*
GMU-AFCEA Symposium,
22 May 2013

MITRE

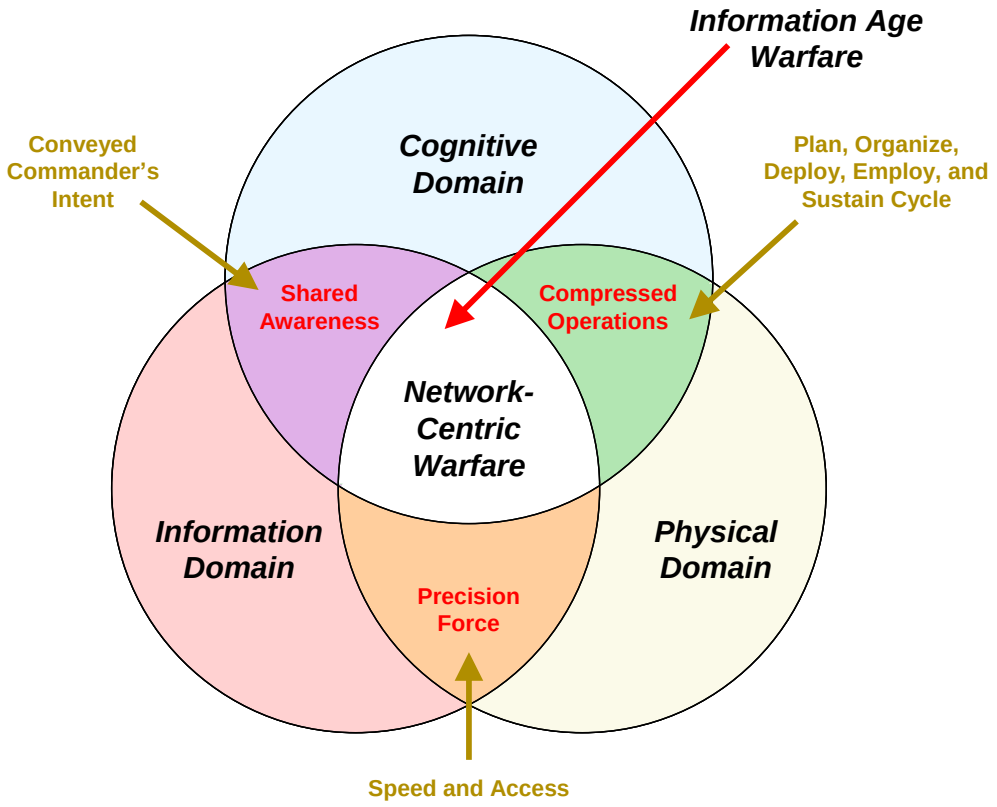
Outline

- Data interoperability and Net-Centric Warfare
- The Universal Core (UCore)
- The National Information Exchange Model (NIEM)

MITRE

Net-Centric Warfare

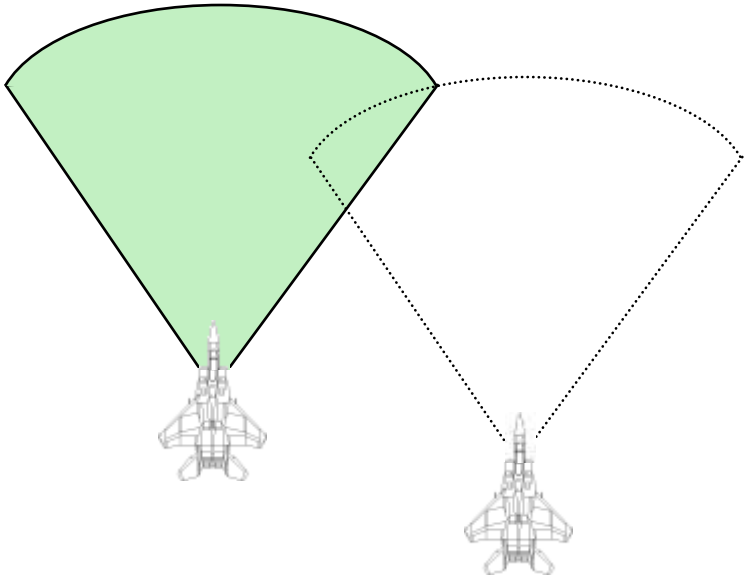
- Seamless interoperability
  - The network is only the beginning!
- Permits sharing of
  - Information
  - Situational awareness
  - Commander's intent
- Leading to
  - Speed of command
  - Self-synchronization
  - Enemy lock-out
- Producing increased combat power



NCW: Creating A Decisive Advantage
ASD/NII, Office of Force Transformation (2005)

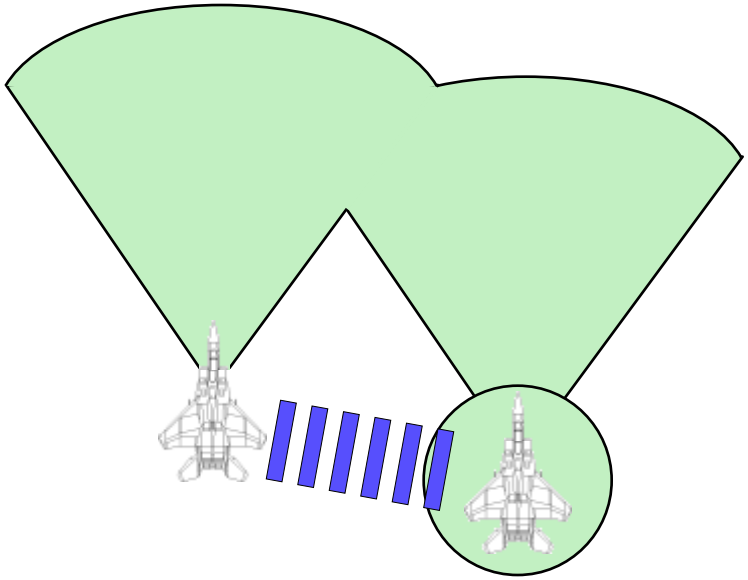
MITRE

Air-To-Air Combat Example Of NCW: Value of Tactical Data Links



Voice only

3-to-1 Kill Ratio

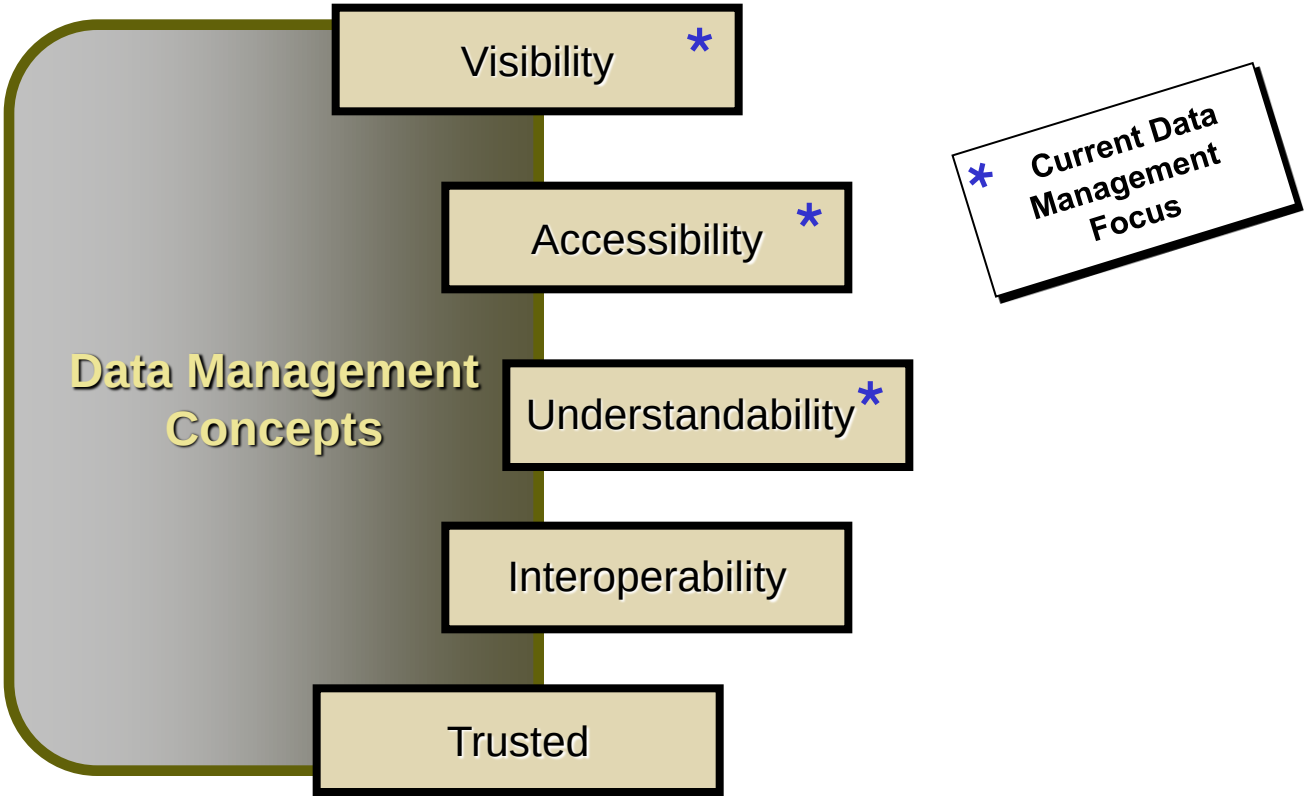


Voice plus data link

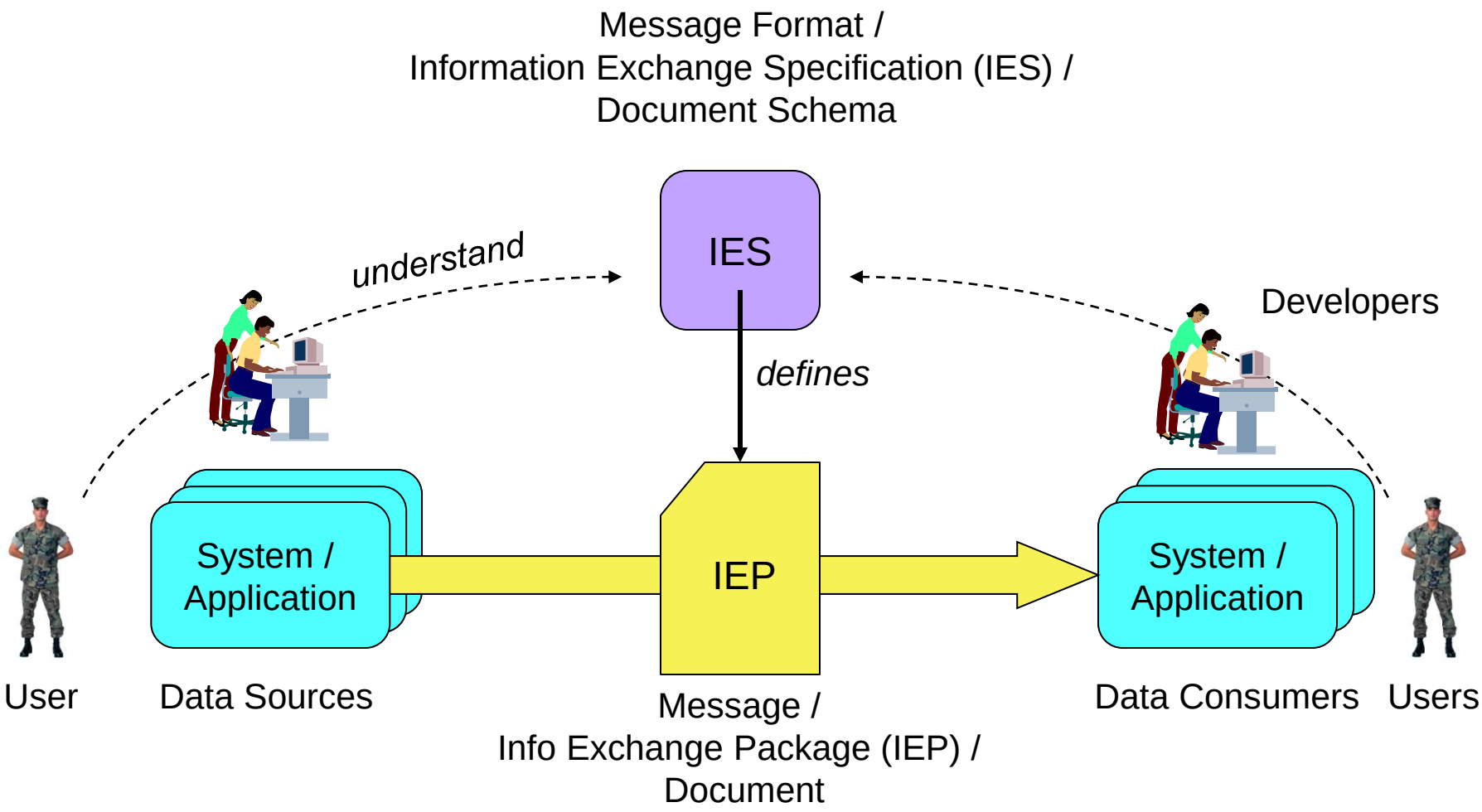
8-to-1 Kill Ratio

DoD Net-Centric Data Strategy: Goals

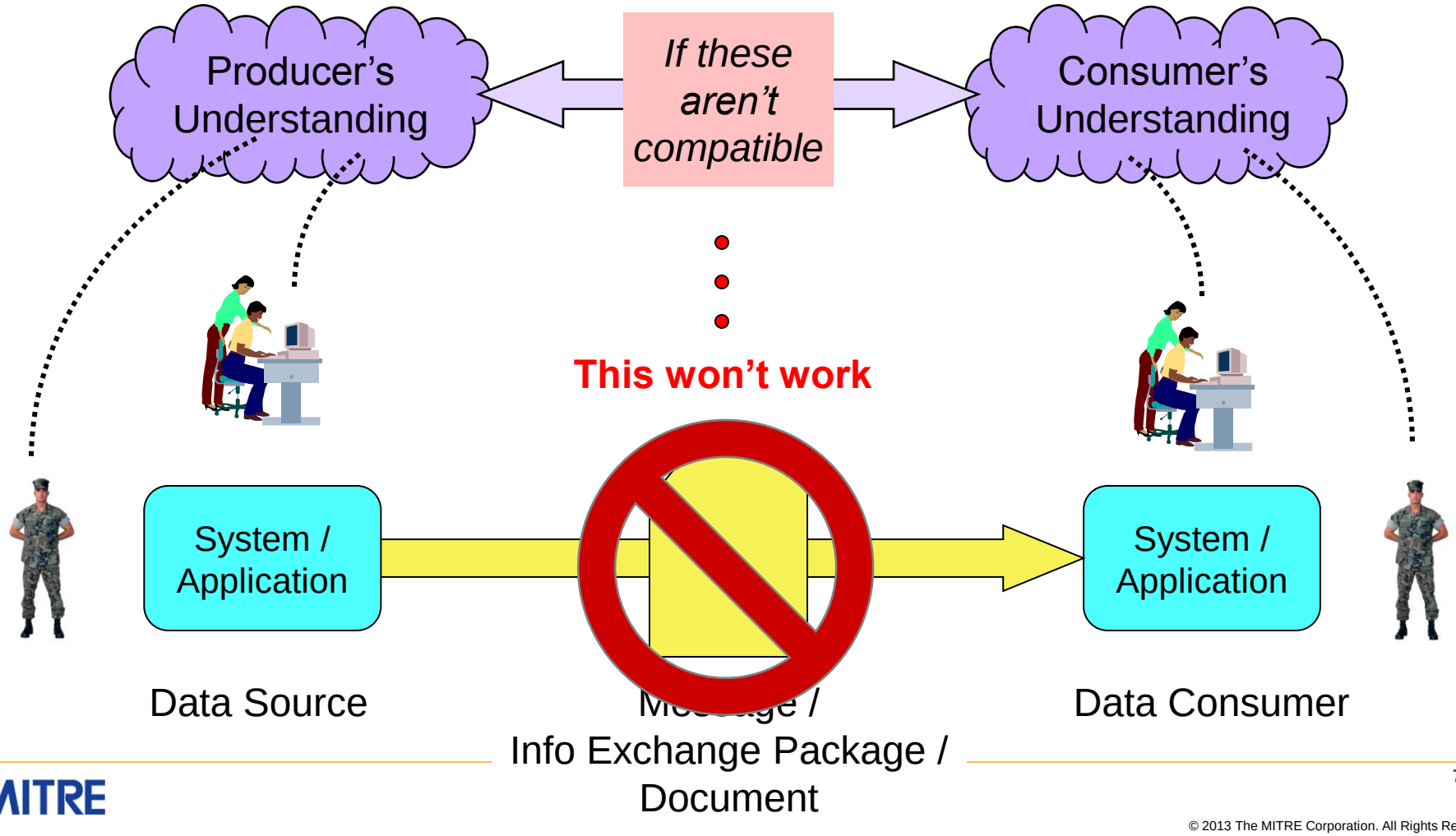
DoD CIO slide, circa 2004



Data Interoperability

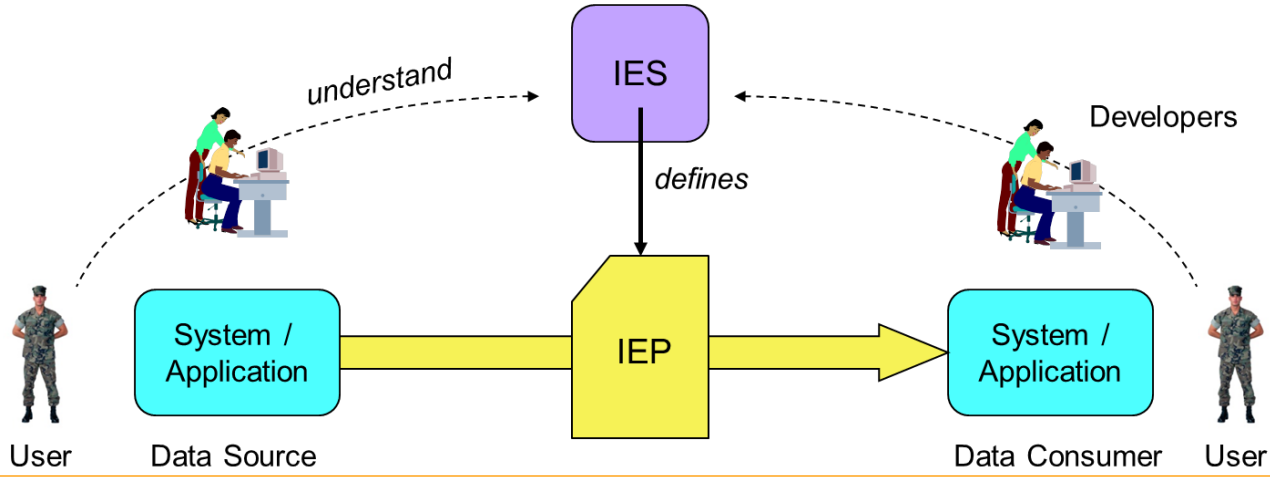


Data Interoperability Problem



Other Data Problems (Out of Scope)

- Is the data accessible?
  - Delivering bits, identity management and access control
- Is the data visible?
  - Enterprise discovery and discovery metadata
- Does the desired information actually exist?
  - System capabilities, producer's operating procedures
  - Data quality concerns

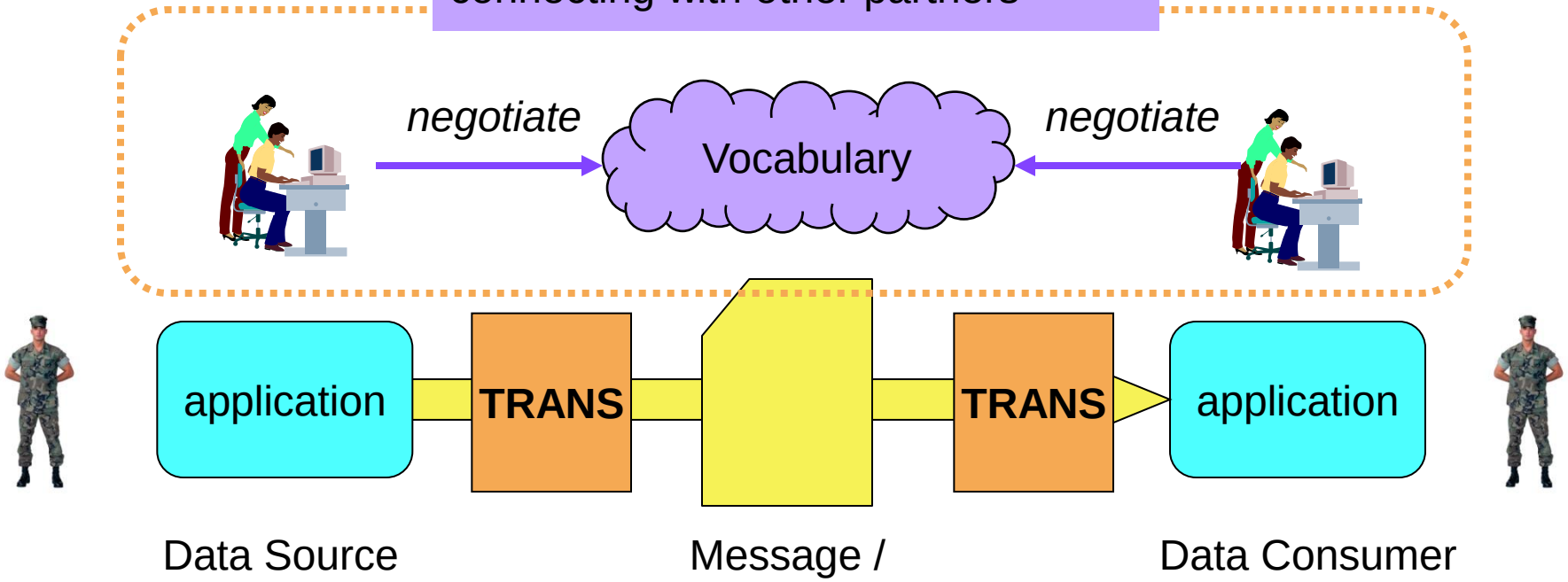


MITRE

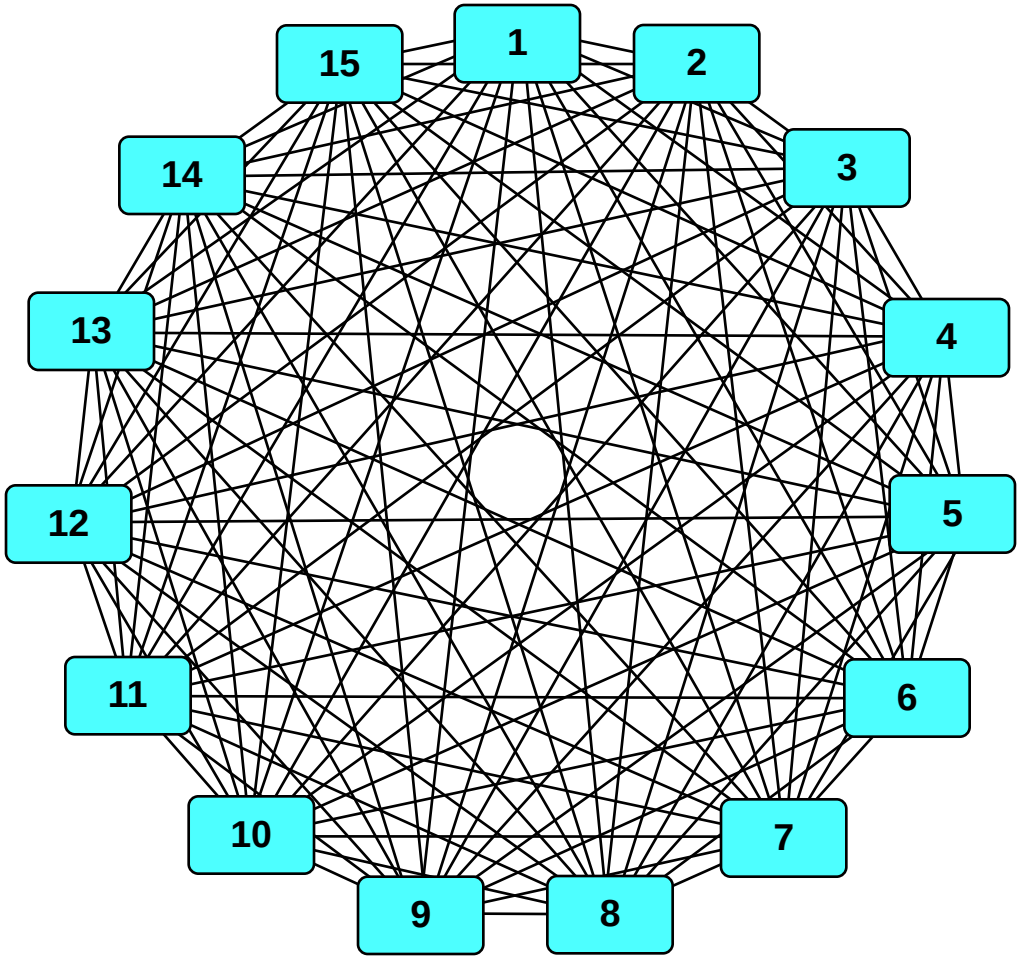
Pairwise Approach Is Simplest …

Shared knowledge is embedded in translation software… very hard to get the knowledge back out for software maintenance, or for connecting with other partners

Community of Two



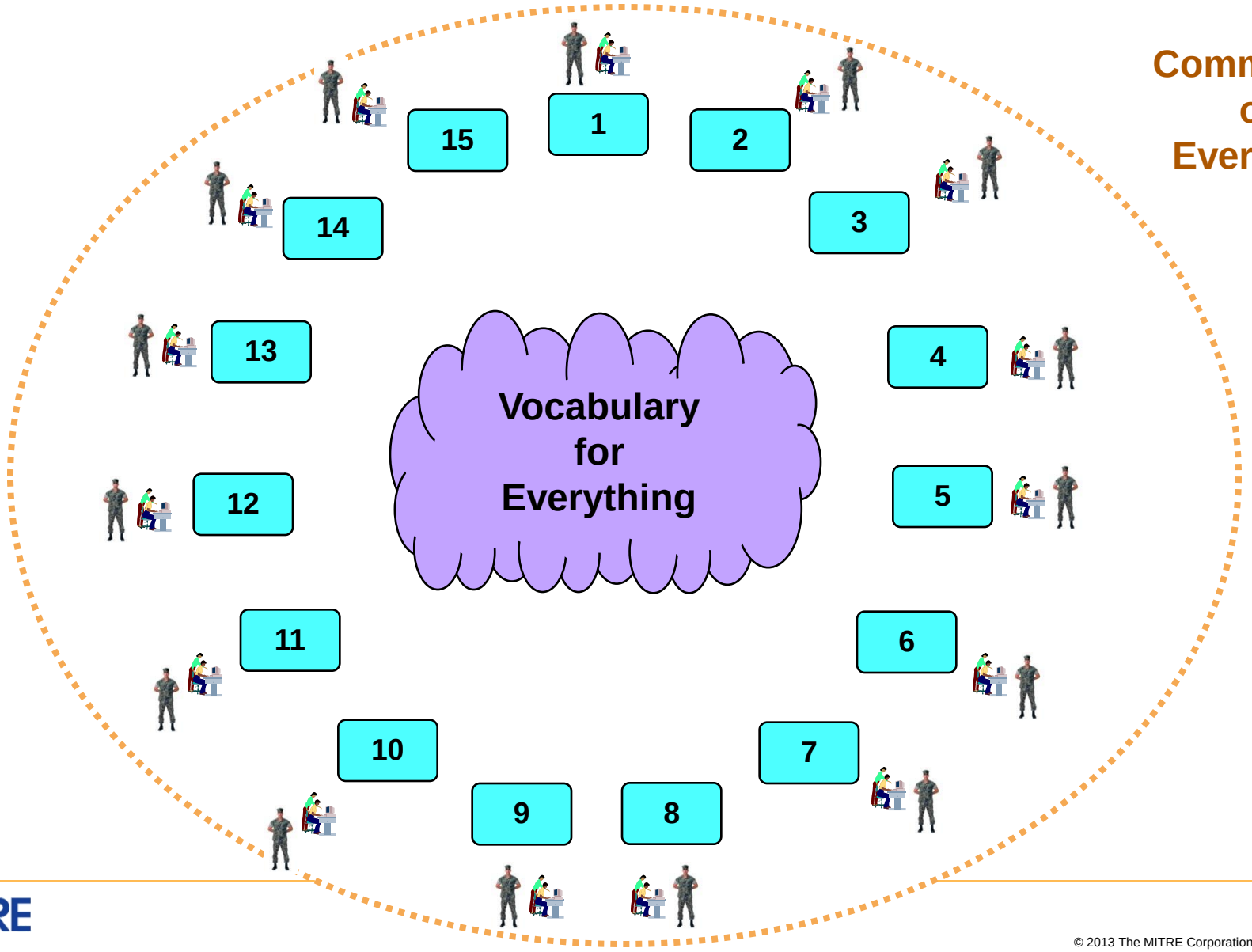
… But There Are Too Many Pairs



The N^2 problem
15 applications
210 negotiations
Total level of effort is $O(n^2)$

MITRE

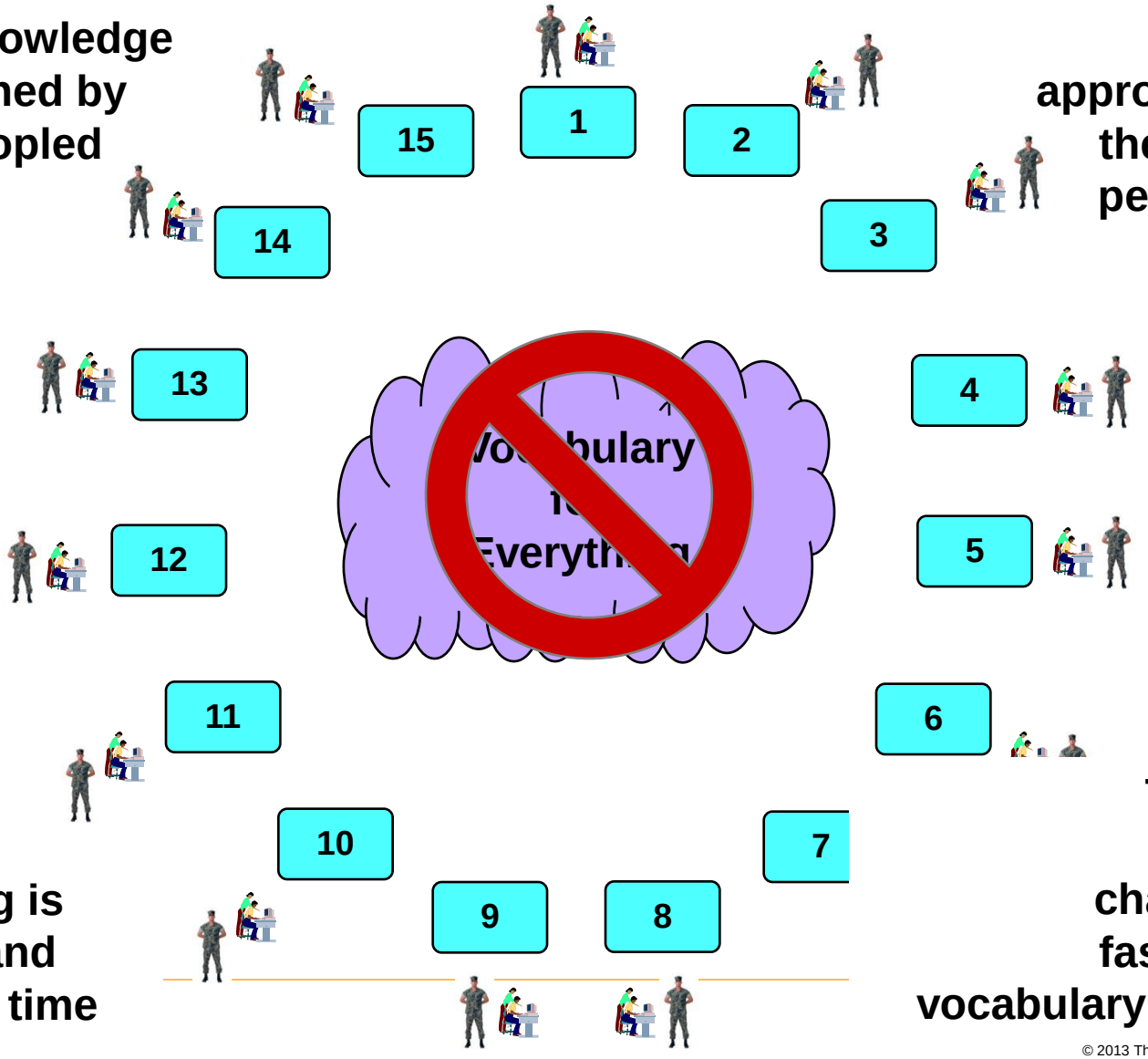
A Single Vocabulary Would Be Nice…



… But Won't Work At Large Scale

Too much knowledge must be learned by too many peopled

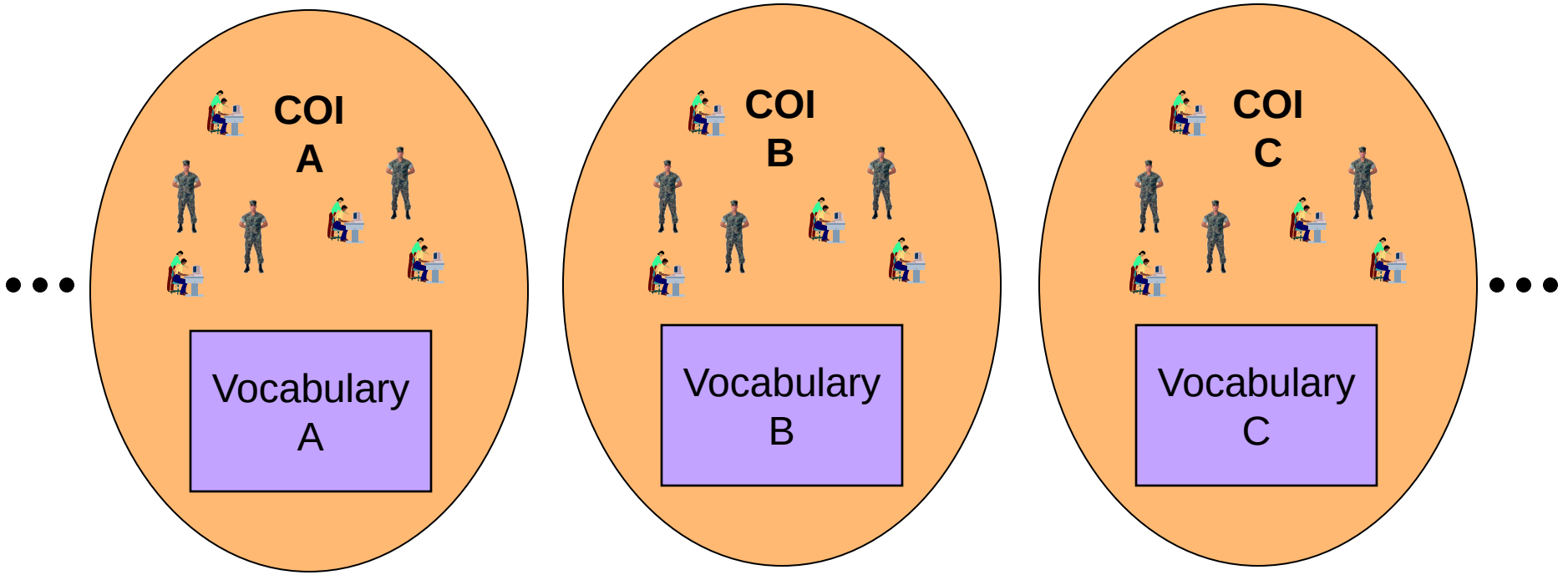
A top-down approach ignores the knowledge people already have



This learning is expensive, and takes a long time

The mission of everyone changes much faster than the vocabulary can change

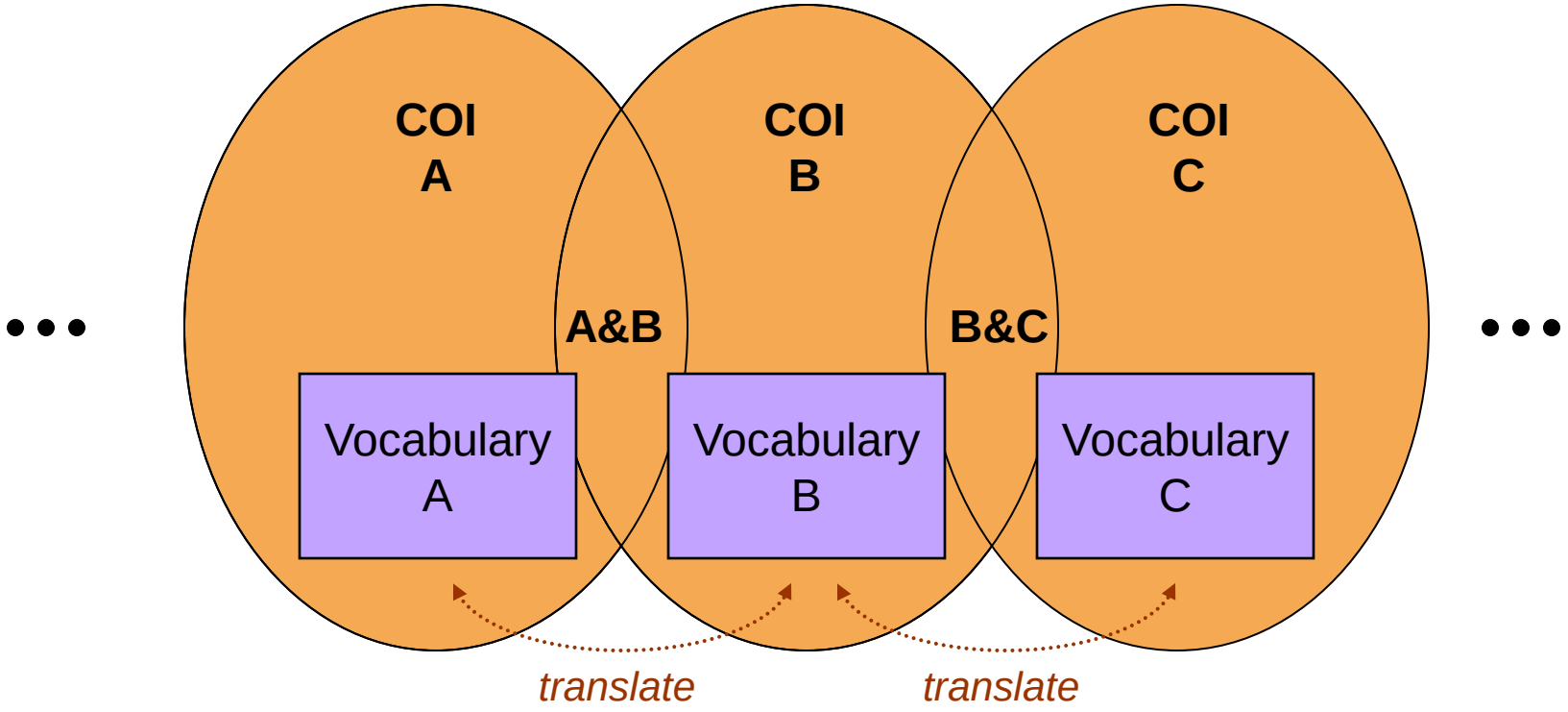
… Which Is Why We Have COIs



If you can't have one universal vocabulary understood by everyone, then you must have several vocabularies, each understood by some group

MITRE

Alas, COIs Will Always Overlap

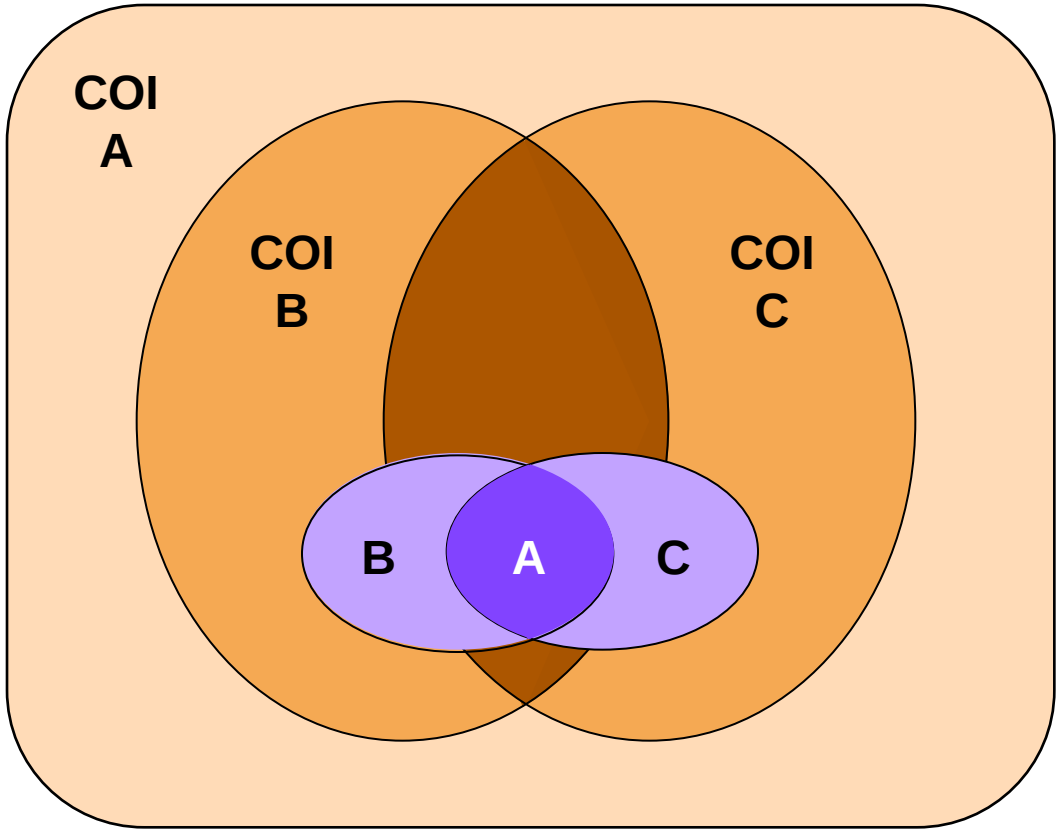


When COIs must exchange data, we must translate between vocabularies
(This by itself is an improvement: N^2 is not so bad when N is small)

However, we can do even better than this…

MITRE

Hierarchy of COIs and Vocabularies

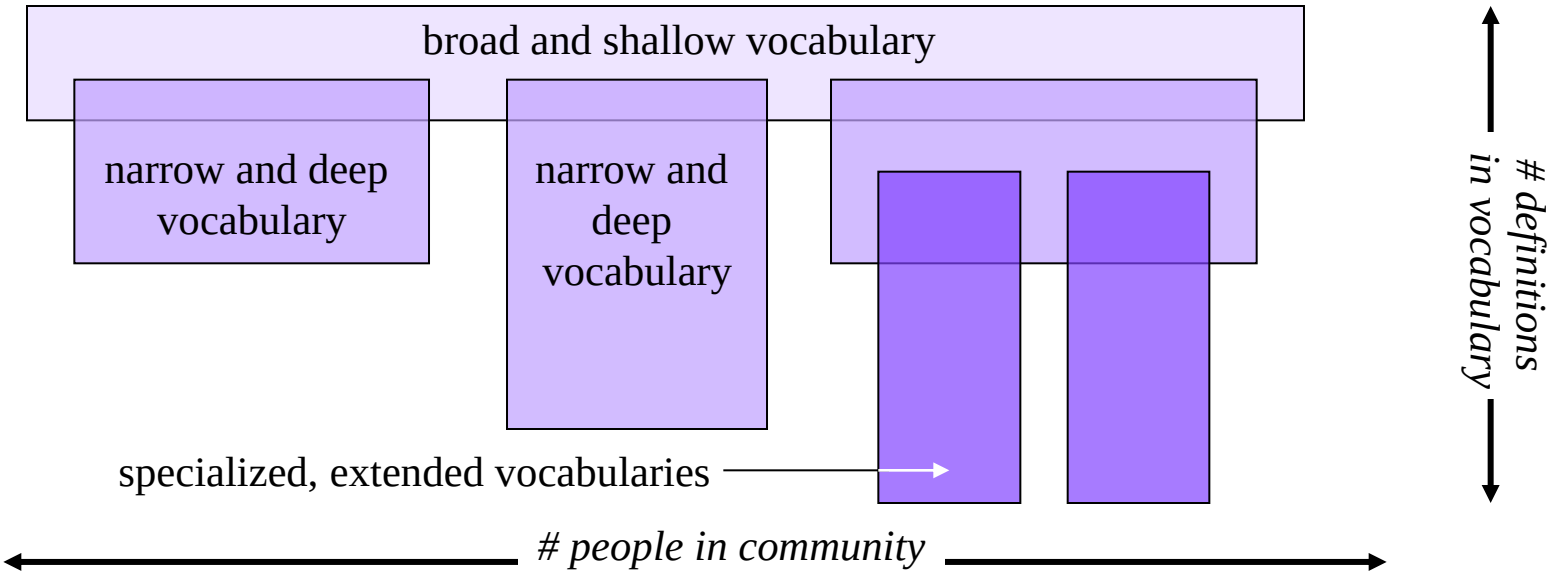


Find the larger community (union) with a smaller vocabulary (intersection)
B and C vocabularies then extend the common A

MITRE

Optimum COI / Vocabulary Approach

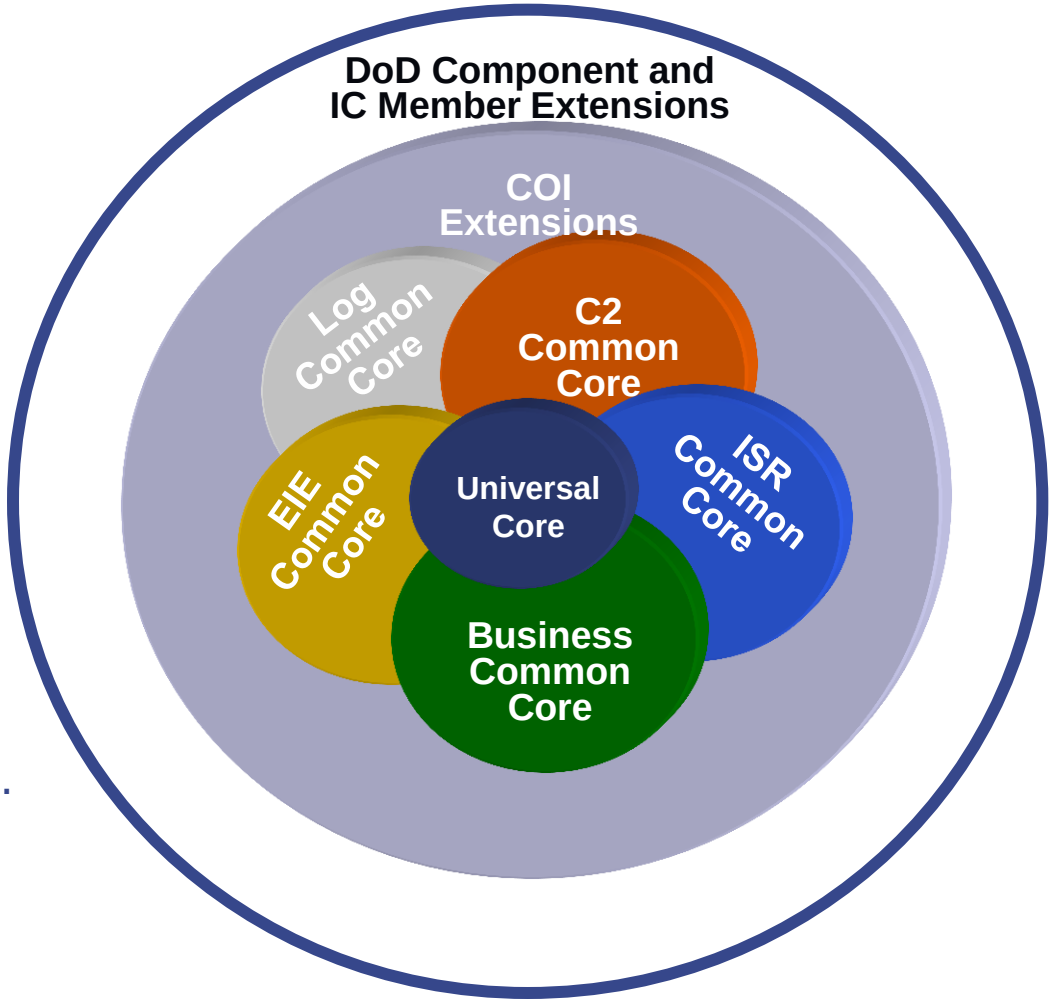
- **The best possible arrangement of COIs and vocabularies:**
  - **A small number of broad, shallow vocabularies: few definitions, understood by many people**
  - **A larger number of narrow, deep vocabularies: many definitions, understood by few people**
  - **Arranged in a rough hierarchy; deep vocabularies adopt and extend the shallow**



Premise for a Universal Core

DoD CIO slide, circa 2006

- Information sharing requires agreement with partners on semantics and syntax.
- Anticipating your partners will be difficult.
- A valuable approach is the promotion of common syntax and semantics for the few things we can all agree on (i.e. a universal core).



Increased Data Interoperability

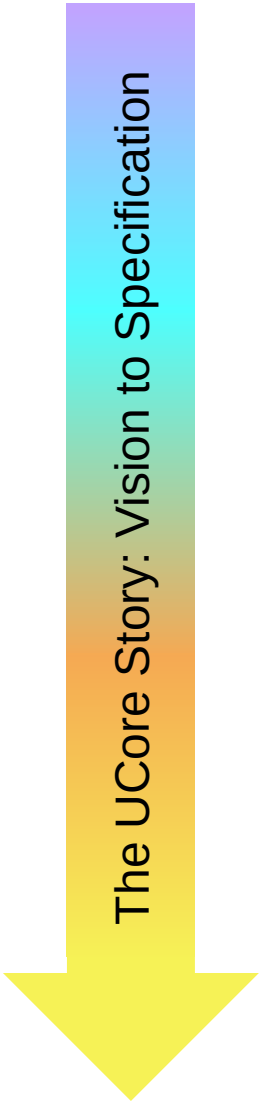
Outline

- Data interoperability and Net-Centric Warfare
- The Universal Core (UCore)
- The National Information Exchange Model (NIEM)

The Universal Core (UCore): 2007-2012

- **Agreed representation for the most commonly shared and universally understood concepts**
  - **Who, What, When Where**
- **Three UCore versions**
  - **UCore 1: Technology refresh for Cursor-on-Target**
  - **UCore 2: Message framework shared with DoJ, DHS, DNI**
  - **UCore 3: Data components for reuse in many exchanges**
- **No UCore development after April, 2012**

The Ideal Story of UCore



UCore Vision	A persuasive “UCore story” would begin with a single vision, expanded to a goal that makes sense to developers, connected to engineering specifications through worked examples
UCore implementation goal	What does UCore accomplish? What do developers do with it? When do they use it? When do they use something else?
UCore implementation details	Worked examples which follow the specifications, and which, if widely imitated, would accomplish the goal and satisfy the vision
UCore specifications	Precise instructions for developers, plus conformance rules

The Actual Story of UCore

UCore Vision	“A framework for sharing the most commonly used data concepts of Who, What, When, and Where [4W]... to improve information sharing by defining and exchanging a small number of important, universally understandable concepts between data-sharing communities without requiring complex mediations” [UCore 3 Way Ahead]

Consensus at a very high level

The Actual Story of UCore

UCore Vision	"A framework for sharing the most commonly used data concepts of Who, What, When, and Where [4W]... to improve information sharing by defining and exchanging a small number of important, universally understandable concepts between data-sharing communities without requiring complex mediations" [UCore 3 Way Ahead]			
Multiple conflicting UCore implementation goals	A "universal" profile of standards necessary in any data exchange	A simple information exchange specification for situation awareness	A 4W discovery metadata summary of a resource	A small collection of 4W data components, for every data exchange

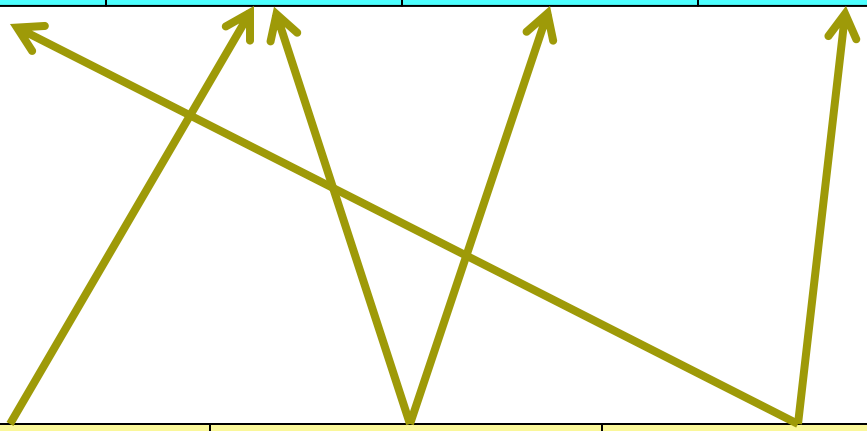
Consensus at a very high level

Unacknowledged and unrecognized conflict

Documented in
The Purpose of UCore: A Conflict of Visions (2010)

The Actual Story of UCore

UCore Vision	"A framework for sharing the most commonly used data concepts of Who, What, When, and Where [4W]... to improve information sharing by defining and exchanging a small number of important, universally understandable concepts between data-sharing communities without requiring complex mediations" [UCore 3 Way Ahead]			
Multiple conflicting UCore implementation goals	A "universal" profile of standards necessary in any data exchange	A simple information exchange specification for situation awareness	A 4W discovery metadata summary of a resource	A small collection of 4W data components, for every data exchange



UCore specifications	UCore 1.0 (2007)	UCore 2.0 (2009)	UCore 3.0 (2012)

Consensus at a very high level

Unacknowledged and unrecognized conflict

Actual engineering specs of varying quality and maturity

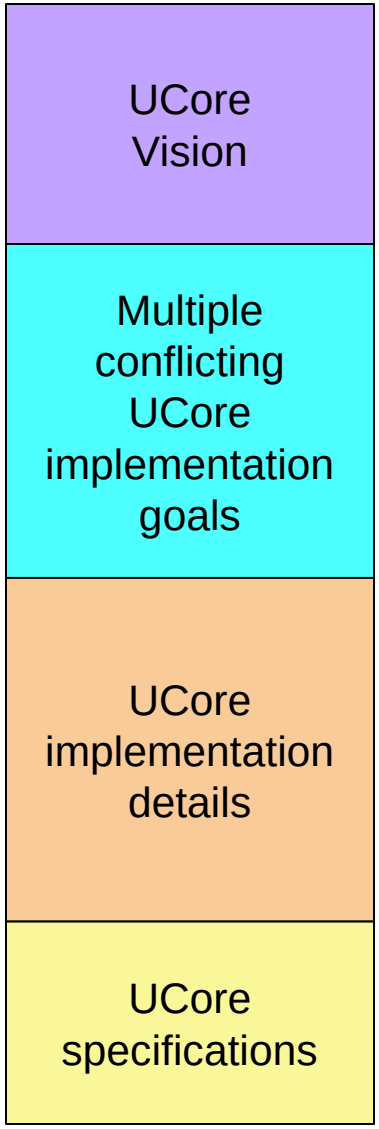
MITRE

The Actual Story of UCore

UCore Vision	"A framework for sharing the most commonly used data concepts of Who, What, When, and Where [4W]... to improve information sharing by defining and exchanging a small number of important, universally understandable concepts between data-sharing communities without requiring complex mediations" [UCore 3 Way Ahead]				Consensus at a very high level
Multiple conflicting UCore implementation goals	A "universal" profile of standards necessary in any data exchange	A simple information exchange specification for situation awareness	A 4W discovery metadata summary of a resource	A small collection of 4W data components, for every data exchange	Unacknowledged and unrecognized conflict
UCore implementation details	**"... and then a miracle occurs"** Efforts that don't follow the specifications Efforts that don't satisfy the goals No solid story or example connecting vision to specifications				Some good work, also plenty of misinformation, disinformation, and slideware
UCore specifications	UCore 1.0 (2007)	UCore 2.0 (2009)	UCore 3.0 (2012)		Actual engineering specs of varying quality and maturity

MITRE

Result = Confusion



Consensus at a very high level

Unacknowledged and unrecognized conflict

Some good work, also plenty of misinformation, disinformation, and slideware

Actual engineering specs of varying quality and maturity

- **We can't tell a good story using worked examples to connect the vision, goals that make sense to developers, and the UCore specification**
- **We can't tell developers when they should use UCore, for what, and why**
- **But there's talk of mandating UCore for all DoD data exchange**
  - **CJCSI 6212.02**
  - **DIEA 1.0**
  - **DoDD 8320.02 (draft)**
- **Result: Confusion and Uncertainty**

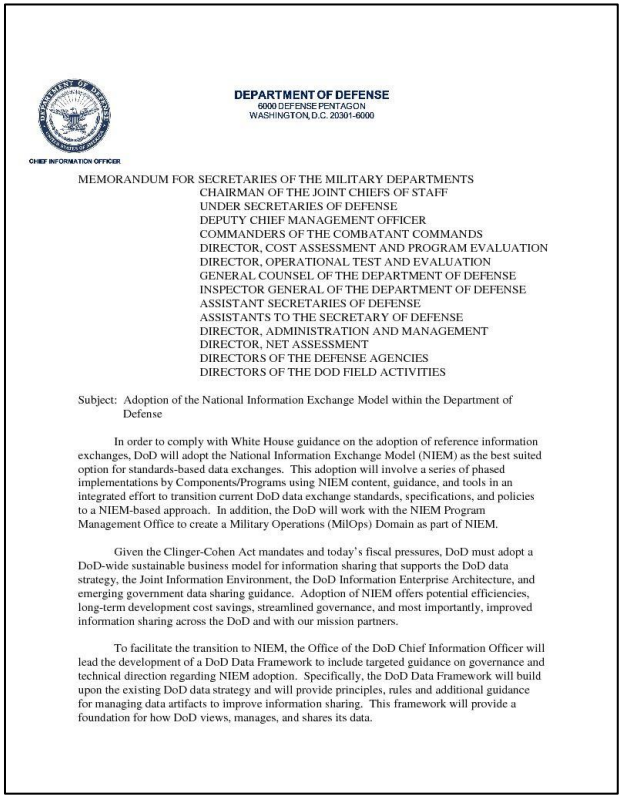
Outline

- Data interoperability and Net-Centric Warfare
- The Universal Core (UCore)
- The National Information Exchange Model (NIEM)

MITRE

DoD Adopts NIEM

- DoD CIO memo (March 2013) declares intent to adopt NIEM
- Finds that NIEM is the best suited option for standards-based data exchange
- DoD organizations shall first consider NIEM for information sharing solutions
  - Exceptions to NIEM expected
  - Will be approved when warranted
- UCore efforts are subsumed into NIEM transition



DEPARTMENT OF DEFENSE
6000 DEFENSE PENTAGON
WASHINGTON, D.C. 20301-6000

CHIEF INFORMATION OFFICER

MEMORANDUM FOR SECRETARIES OF THE MILITARY DEPARTMENTS
CHAIRMAN OF THE JOINT CHIEFS OF STAFF
UNDER SECRETARIES OF DEFENSE
DEPUTY CHIEF MANAGEMENT OFFICER
COMMANDERS OF THE COMBATANT COMMANDS
DIRECTOR, COST ASSESSMENT AND PROGRAM EVALUATION
DIRECTOR, OPERATIONAL TEST AND EVALUATION
GENERAL COUNSEL OF THE DEPARTMENT OF DEFENSE
INSPECTOR GENERAL OF THE DEPARTMENT OF DEFENSE
ASSISTANT SECRETARIES OF DEFENSE
ASSISTANTS TO THE SECRETARY OF DEFENSE
DIRECTOR, ADMINISTRATION AND MANAGEMENT
DIRECTOR, NET ASSESSMENT
DIRECTORS OF THE DEFENSE AGENCIES
DIRECTORS OF THE DOD FIELD ACTIVITIES

Subject: Adoption of the National Information Exchange Model within the Department of Defense

In order to comply with White House guidance on the adoption of reference information exchanges, DoD will adopt the National Information Exchange Model (NIEM) as the best suited option for standards-based data exchanges. This adoption will involve a series of phased implementations by Components/Programs using NIEM content, guidance, and tools in an integrated effort to transition current DoD data exchange standards, specifications, and policies to a NIEM-based approach. In addition, the DoD will work with the NIEM Program Management Office to create a Military Operations (MilOps) Domain as part of NIEM.

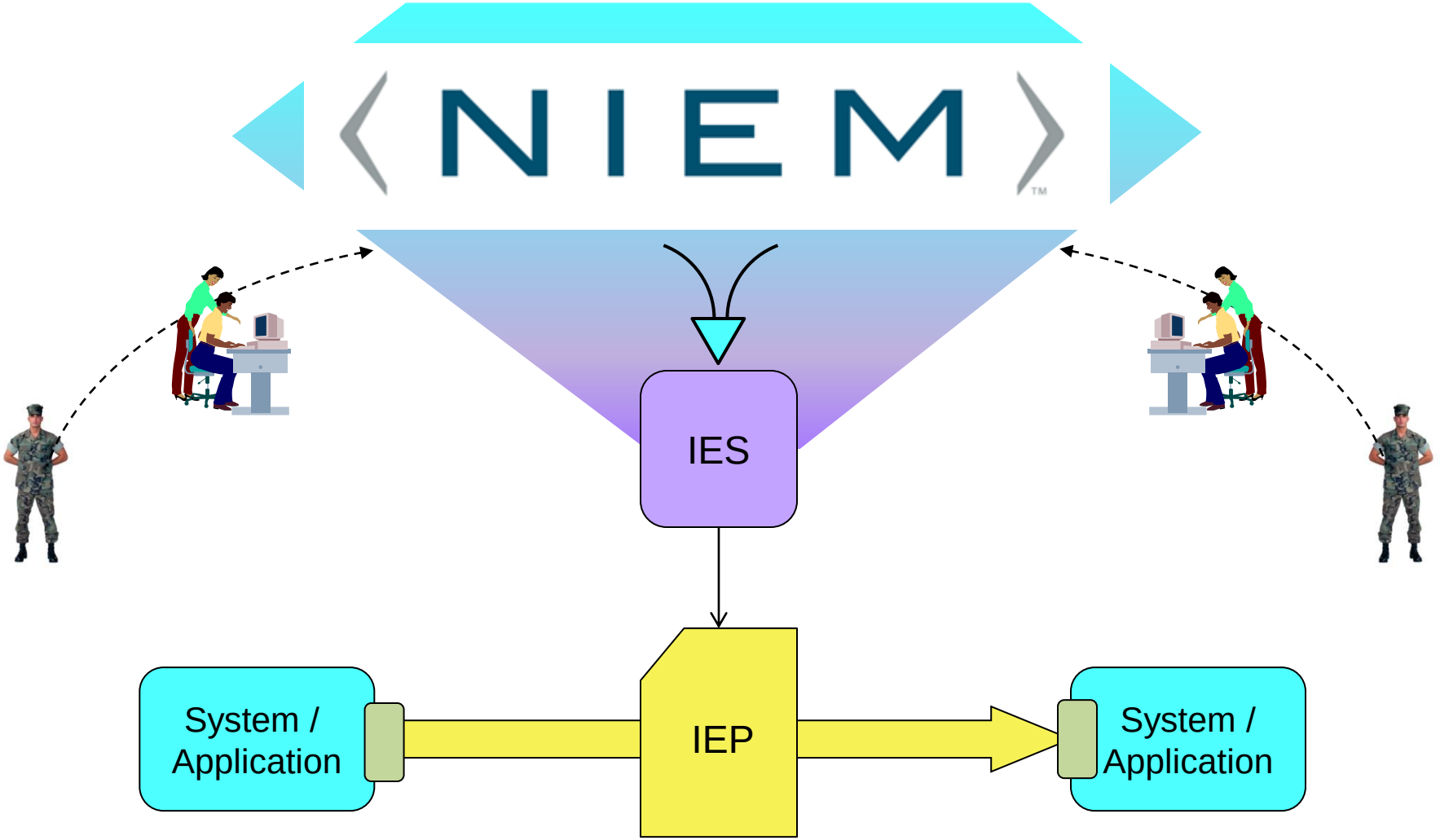
Given the Clinger-Cohen Act mandates and today's fiscal pressures, DoD must adopt a DoD-wide sustainable business model for information sharing that supports the DoD data strategy, the Joint Information Environment, the DoD Information Enterprise Architecture, and emerging government data sharing guidance. Adoption of NIEM offers potential efficiencies, long-term development cost savings, streamlined governance, and most importantly, improved information sharing across the DoD and with our mission partners.

To facilitate the transition to NIEM, the Office of the DoD Chief Information Officer will lead the development of a DoD Data Framework to include targeted guidance on governance and technical direction regarding NIEM adoption. Specifically, the DoD Data Framework will build upon the existing DoD data strategy and will provide principles, rules and additional guidance for managing data artifacts to improve information sharing. This framework will provide a foundation for how DoD views, manages, and shares its data.

Subject: Adoption of NIEM within the DoD


NIEM
TM

A Standards-Based Approach For IES Design



MITRE

THE NIEM FRAMEWORK

NIEM connects communities of people who share a common need to exchange information in order to advance their missions, and provides a foundation for seamless information exchange between federal, state, local, and tribal agencies. Much more than a data model, NIEM offers an active user community as well as a technical and support framework.

Community

- Formal Governance Processes
- Online Repositories
- Mission-Oriented Domains
- Self-Managing Domain Stewards

Technical Framework

- Data Model
- XML Design Rules
- Development Methodology
- Predefined Deliverables (IEPD)

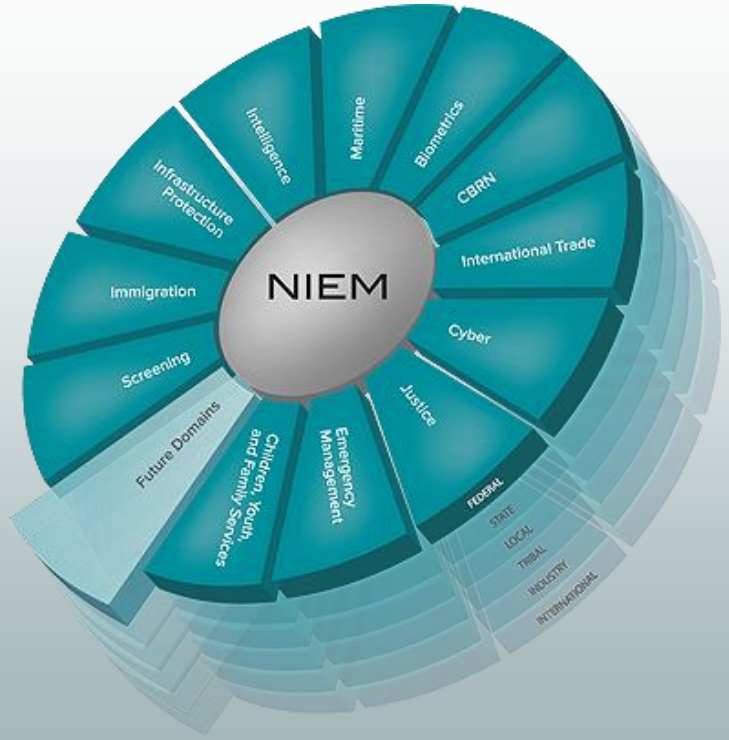
Support Framework

- Tools for Development and Discovery
- Established Training Program
- Implementation Support
- Help Desk & Knowledge Center

THE NIEM LIFECYCLES

Common Language

(Data Model Lifecycle)



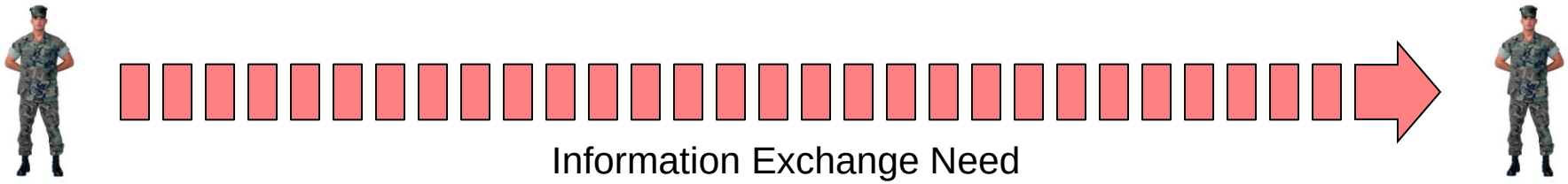
Built and governed by the business users at Federal, State, Local, Tribal and Private Sectors

Repeatable, Reusable Process

(Exchange Specification Lifecycle)

Using the NIEM Approach To Satisfy An Information Exchange Need

Start with an information exchange need. You don't have to know all of the participants in advance. It's enough to know that the producer has data that needs to be shared with someone



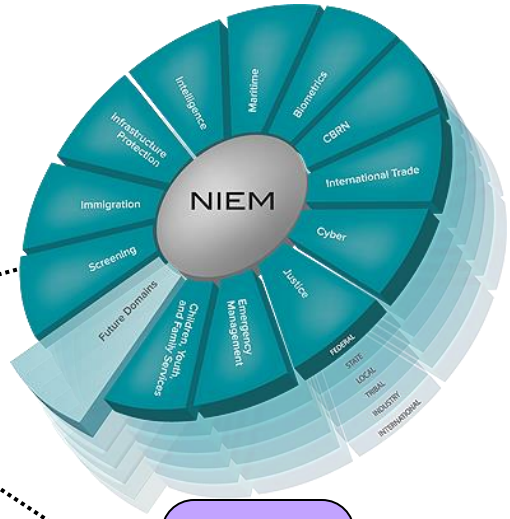
Producers

Consumers

MITRE

Satisfying An Information Exchange Need

Developers can use NIEM rules and NIEM **data components** to define the **exchange** and implement a **service interface** for the data source



IES

It's helpful if the needed concepts are already defined in NIEM Core or a domain, but developers can always create any definitions they need

System / Application

Information Exchange Need

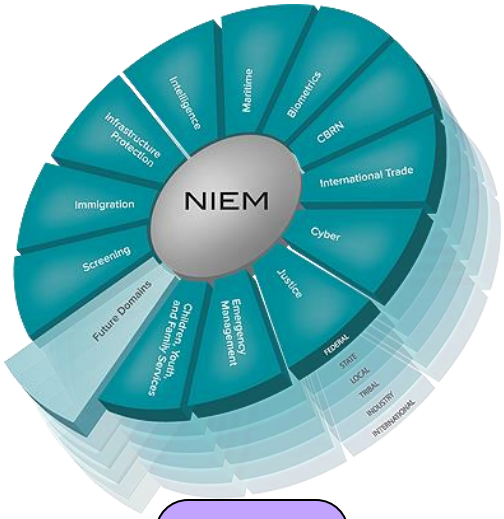
User

Data Source

User

MITRE

Satisfying An Information Exchange Need



IES

Consumer developers write **code to process messages** which follow the IES. Part of their work is already done because they understand the NIEM Core and domain data components

System / Application

Information Exchange Need

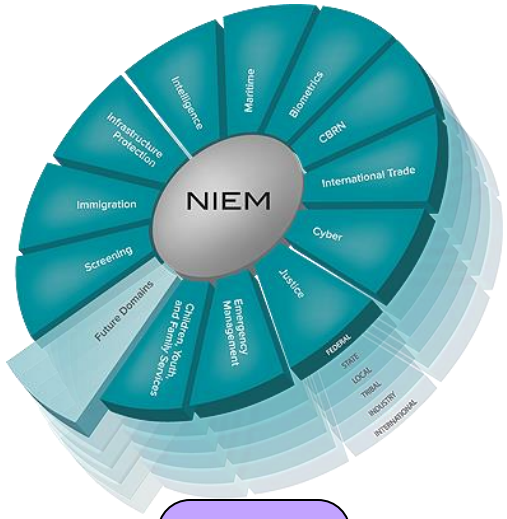
System / Application

User | Data Source | Data Consumer | User

MITRE

Satisfying An Information Exchange Need

By following the NIEM standards-based approach, this machine-to-machine data exchange can be implemented faster and at lower cost



Other producers and consumers may join at any time, in any order, by also implementing the IES.

IES

defines

System / Application

IEP

System / Application

User | Data Source | Data Consumer | User

MITRE

NIEM Is Not Needed For Every Exchange

- Some data exchanges have no need for NIEM; for example:
  - Email is a kind of data exchange, one that is already specified by Internet standards (RFC 2822, etc.)
  - No need to use NIEM for image data; we already have suitable data standards for images
- Reasons to not use NIEM
  - Not building a machine-to-machine exchange
  - Already have a satisfactory exchange specification and installed software base

Data Exchanges Provide Data Interoperability

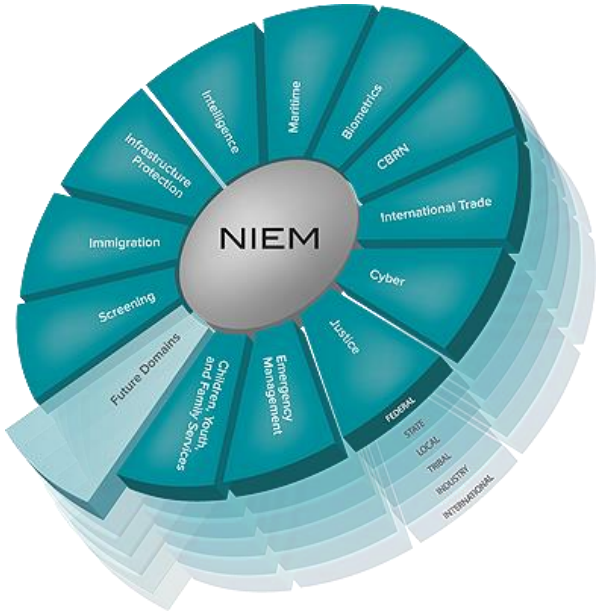
- NIEM is emphatically not a single comprehensive data model for all data exchanges
- A system does not simply “implement NIEM” and thereby become completely interoperable with every other system “implementing NIEM”
- A system can implement a particular NIEM-conforming information exchange specification
- All systems implementing a particular IES are interoperable with each other, for that exchange

In NIEM, interoperability is defined at the IES level

MITRE

NIEM Is About Cooperation, Not Control

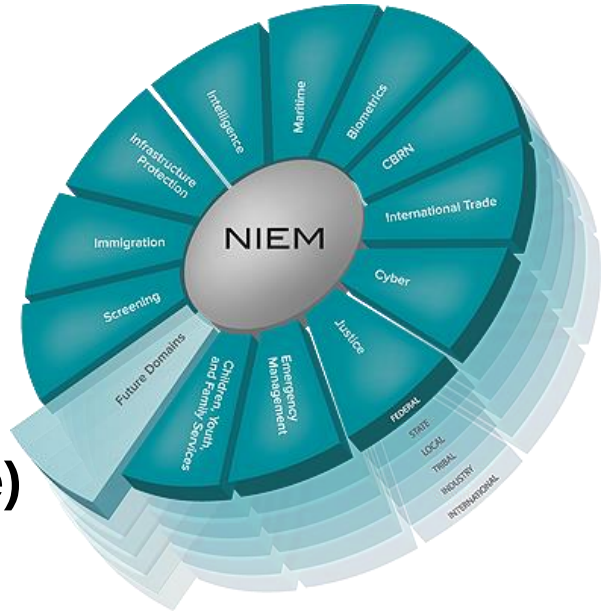
- NIEM Core and NIEM domains create data components by consensus among data exchange designers
- Components are established when participants believe that a common definition will make their exchanges easier to create and implement
- Each domain changes on its own schedule, under its own control
- Changes in the core or in a domain do not force changes in other domains or in any data exchange
- No one is ever required to use a component that does not satisfy the data exchange needs, so there is no leverage for controlling the participants



MITRE

“The Data Exchange Designer Is King”

- Nothing in NIEM or any NIEM domain can ever dictate the information content of any data exchange
- NIEM rules for extension and reuse guarantee that every exchange designer will always be able to specify the exact information needed by his exchange participants
- If you want more control over your enterprise or community, you must layer that governance on top of NIEM
- NIEM will work with
  - The agreement you can achieve
  - The flexibility you need to have
  - The control you choose to impose (within scale limits on vocabulary size)



MITRE

Summary

- Improved data interoperability leads to increased combat power
- Community agreement on semantics and syntax leads to data interoperability
- DoD has adopted NIEM as the best-suited option for creating and managing that agreement
- NIEM is not the complete answer to data interoperability
  - Doesn't attempt to address all the interoperability issues
  - Not the best approach for every single data exchange
  - Consider NIEM approach first, use it where you can

Resources

- **Web site: niem.gov**
- **Papers and briefings**
  - **A Comparison of Cursor-on-Target, UCore, C2 Core, and NIEM (2013)**
  - **UCore Visions, UCore Reality (2010)**
  - **The Purpose of UCore: A Conflict of Visions (2010)**
  - **Core Vocabularies and Information Sharing (2008)**
  - **Universal and Common Cores for DoD and DNI (2007)**
  - **The Enterprise Core Schema Must Be Very Small (2007)**
  - **My Two Cats Are a Community of Interest (2006)**
  - **Net-Centric Information Management (2005)**
  - **Building Information Systems for Network-Centric Warfare (2003)**
  - **A "Community of Interest" Approach to Data Interoperability (2002)**

Scott Renner
sar@mitre.org

BACKUP

NIEM Is A Way To Build Data Exchanges

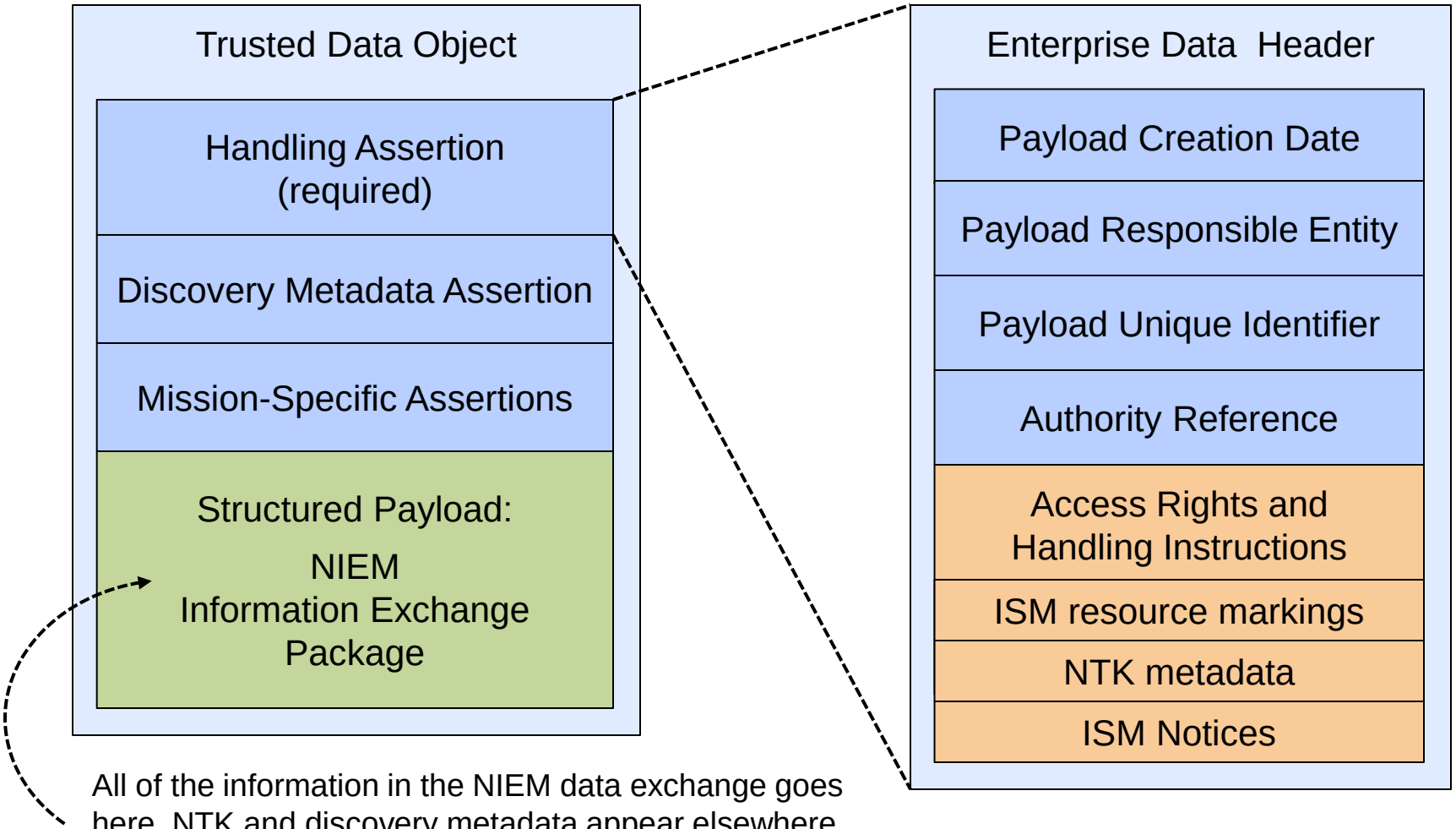
- NIEM specifications describe how to consistently apply industry standards (XML, XML Schema, etc.) to improve reusability and interoperability
- Emphasis on machine-to-machine data exchange
- Each information exchange specification (IES) is designed, then implemented in software
  - The design work could be done by POR developers, COI members, an authoritative data source (ADS), a message standards body, etc.
  - This work (and its governance) is separate from NIEM
  - NIEM only provides a framework: data components and assembly rules

NIEM Works In Many Situations

- Does the NIEM "adapter pattern" cause problems in a NIEM exchange?
- Why doesn't NIEM use xs:import? Why doesn't it do "native reuse“?
- Do the "wildcards" in NIEM keep NIEM messages from passing through a cross-domain solution (CDS) guard?
- Does NIEM conform to ISM, NTK, MAT, etc.?
- Does anything in NIEM make ISM conformance difficult or impossible?
- Does anything in NIEM make NTK conformance difficult or impossible?
- Does NIEM implement tearlines?
- Does NIEM conflict with "safety of navigation"?
- Does NIEM conflict with GFM DI?
- Does NIEM work with XML compression?
- Does NIEM support OWL and SKOS?
- Does NIEM properly define ID attributes for XLink and IDREFs?

These and other technical concerns have been resolved
(Source: *C2 Core Technical FAQ*)

NIEM Works With IC Trusted Data Framework



All of the information in the NIEM data exchange goes here. NTK and discovery metadata appear elsewhere, and need not be part of the NIEM IES. ISM markings are only needed for a multi-portion IEP.

MITRE

The Geospatial Data Challenge

- Geospatial community is not adopting NIEM
  - National System for Geospatial Intelligence (NSG)
  - Settled on OGC specifications (GML, etc.)
  - Large installed software base (ESRI ArcGIS, etc.)
- Rules for GML application schema and GML document are not completely compatible with NIEM
  - Data exchanges which must be specified by a GML application schema should be considered out of scope for complete NIEM conformance
- Feasible and useful to use GML and NIEM together
  - Easy to embed GML data in NIEM message
  - Easy to embed NIEM data in GML document
  - Important to show developers the best way to do each

NIEM Works In Low-Bandwidth Environment

- Efficient XML Interchange (EXI) is the W3C standard for compressed binary XML data
- For a properly designed schema, EXI produces messages very near the information-theoretic minimum size
  - Often smaller than hand-crafted binary message formats
  - Plenty of experimental evidence
- Properly-designed schemas are always possible with NIEM
  - Nothing in NIEM ensures a good low-bandwidth design
  - Nothing in NIEM makes a good design impossible
  - In practice, good design with NIEM is not difficult

**People will always have bandwidth problems,
but these problems will not be caused by NIEM**

NIEM Works In Low-Bandwidth Environment

- Efficient XML Interchange (EXI) is the W3C standard for compressed binary XML data
- For a properly designed schema, EXI produces messages very near the information-theoretic minimum size
  - Often smaller than hand-crafted binary message formats
  - Plenty of experimental evidence
- Properly-designed schemas are always possible with NIEM
  - Nothing in NIEM ensures a good low-bandwidth design
  - Nothing in NIEM makes a good design impossible
  - In practice, good design with NIEM is not difficult

**People will always have bandwidth problems,
but these problems will not be caused by NIEM**