MTR 05B000095

COI Lessons Learned:
Observations and Recommendations

November 28, 2005

Scott Renner

MITRE

Center for Air Force C2 Systems
Bedford, Massachusetts

Section 1

Introduction

The DoD Net-Centric Data Strategy (NCDS), published in May 2003, describes a vision for a net-centric environment and the data goals for achieving that vision [1]. The NCDS goals for data – make it visible, accessible, understandable, interoperable, and trusted – will lead to the information exchange that is essential to net-centricity. When producers post data to shared spaces, they make it available to unanticipated users and applications, leading to improved flexibility and increased warfighter agility.

Communities of Interest (COIs) are a key element of the NCDS. The COI concept was introduced as an alternative to DoD-wide data element standardization. COIs are defined as "collaborative groups of users who must exchange information in pursuit of their shared goals, interests, missions, or business processes and who therefore must have shared vocabulary for the information they exchange." They promote data visibility through metadata catalogs, which contain discovery metadata enabling consumers to find data assets that meet their needs. They promote data accessibility through shared spaces, which permit consumers to retrieve needed data, while still enforcing policy on access control and priority. They promote data understandability through shared vocabularies, comprised of semantic artifacts (dictionaries, data models, taxonomies, ontologies, etc.), which help the member establish a shared understanding of the data they exchange.

COIs and the NCDS are new ideas – the policy, procedure, and technology aspects are not completely worked out. It is reasonable to anticipate both difficulties in getting it right and risks of getting it wrong. For this reason, the NCDS calls for pilot activities with "trial COIs" to generate experience that will be the basis for refining the COI construct. These pilot activities are under way. In addition, certain COI-like activities, predating the NCDS, have provided useful experience.

This paper summarizes the author's experience with nine such COIs. It presents nine observations generalizing from this experience. Based on those observations, it makes nine recommendations, some of which might be followed by individual COIs, some concerning the COI concept in general.

Section 2

Observations

The observations in this paper have been derived from experience with the following nine COIs. It is limited experience, in some cases, but the author has attended several of the meetings, talked with the COI members and bystanders, listened to the presentations, and read the meeting minutes and briefing charts.[1]

- Air Force Mobility and Transportation (AF Mobility)
- Blue Force Tracking (BFT)
- Command and Control Space Situational Awareness (C2-SSA)
- Global Force Management (GFM) Data Initiative (DI)
- Global Force Management – Air Force (GFM-AF)
- Joint Air and Missile Defense (JAMD)
- Meteorology-Oceanography (METOC)
- Network Operations (NETOPS)
- Time Sensitive Targeting (TST)

Four of these COIs are new efforts sponsored by OASD/NII in order to demonstrate and build experience with the NCDS. These COIs – BFT, C2-SSA, NETOPS, and TST – are here called the "pilot COIs".

The Global Force Management Data Initiative is the very first COI formed in response to the NCDS. While the Air Force GFM effort is a part of GFM-DI, it is also developing additional AF-specific capabilities, and is intended to demonstrate and build experience with the Air Force Information and Data Management Strategy (I&DMS) [2], and so it is worth attention in its own right. The AF Mobility COI was also established to build experience with the AF I&DMS.

JAMD is a "bottom-up" COI. It formed in late 2004 because of perceived need in the developer community, and only then went looking for an official sponsor.

The METOC COI is by far the oldest of the nine. It is one of the success stories to emerge from the (largely failed) 8320.1 data administration program. METOC formed in 1995 and spent the next five years constructing a large[2] conceptual data model [3]. Several systems have derived physical data models from this METOC vocabulary.

---

[1] Some of this material may be located via the COI Directory, at https://gesportal.dod.mil/sites/coidirectory. All nine COIs are listed there.

[2] The Joint Meteorological Conceptual Data Model (JMCDM) defines over 3000 data elements.

Not all of the following nine observations deserve to be called "lessons". They may still be useful for understanding the problems that current and future COIs will face.

1. Each COI must establish its own operating procedures
2. COI efforts face a difficult tradeoff between useful scope and quick results
3. A coalition of the willing does not always marshal enough resources
4. Known consumers with known needs are important for COI success
5. System builders are usually important contributors to COI vocabulary work
6. Most COI efforts do not have time to create or learn large new vocabularies
7. Vocabulary overlap among COIs is inevitable
8. Core enterprise services are not well understood
9. Access control decisions need more attention

1. Each COI must establish its own operating procedures

Anyone searching in FY05 for a COI operator's manual went away disappointed. As of this writing, there is no such document – no instruction, no guidance, and no handbook. OASD/NII has created a process for *establishing* a new COI. However, the people in charge of each new COI effort must choose their own goals and operating procedures.

New COIs cannot even rely on a common understanding of what a COI is, and what it does. The "community of interest" term has become heavily overloaded. It is possible to discern three different kinds of COI, based on their goals. A *vocabulary* COI is focused on the development of the community's common vocabulary. A *sharing* COI is made up of the people who are actually sharing information with each other. A *proponent* COI seeks to establish new or improved information sharing, often through their influence over acquisition. The common aspect is a group of people cooperating to solve a data sharing problem.

Vocabulary COIs have the advantage of long experience with data standardization. (For example, the METOC COI has been in existence for ten years.) New vocabulary COIs can draw on this experience. Of course, some aspects of the old 8320.1 data administration program should not be repeated. There are a few bad habits to avoid, including the urge to assimilate neighboring COIs,[3] working only at the implementation level of detail, and insisting on physical internal implementation of the community data standard.

---

[3] Sometimes expressed as "We'll extend our model to accommodate your needs, and then you can use it".

However, for the most part, each new COI is starting from scratch. There are several documents describing the reason for having COIs, and the good things they are supposed to accomplish, but so far only one that contains anything like a specific sequence of tasks, or the roles of the people who perform them. This is the *COI Handbook*, originally a MITRE technical report [4], subsequently released for information by the Air Force CIO [5]. The GFM-AF, TST, and C2 SSA COIs basically follow the Handbook's model, and have adopted some parts of this material in working out their own plans and procedures. However, all three have found it necessary to make significant changes and revisions to the Handbook's suggestions.

2. COI efforts face a difficult tradeoff between useful scope and quick results

One of the first questions faced by most new COIs involves two related variables: how much to accomplish in their first spiral, and in what amount of time.[4] It is widely believed that COI efforts should produce something of value quickly. Large-scale, long-running, "big bang" migration projects often fail, and are rightly considered high-risk. Partly for this reason, the COI pilots are supposed to select a scope permitting them to deliver useful results within nine to twelve months.

However, it is not always possible to partition a problem area into twelve-month chunks. Sometimes, the smallest useful capability takes longer to implement. For example, the GFM COI is working on a schedule with over two years between inception and prototype capability. Smaller, shorter subprojects were not considered responsive to the consumer requirements.   We also have the example of the METOC COI, which spent nearly five years creating its shared logical data model.

Long projects are usually risky. Short projects do not always deliver what the users need. There is usually a noticeable amount of work involved in arriving at the best tradeoff – a simple, rigid rule will not always produce the best results.

3. A coalition of the willing does not always marshal enough resources

There is no central pot of money for COI funding; resources must come from the COI members.   OASD/NII is providing a small amount of funding for some of the pilot COIs, but only as a "lubricant" to the efforts funded primarily by the Programs of Record (PORs). The PORs do have money, and also have a motive for concept demonstrations as a means of risk reduction for future spirals, so they can be persuaded to participate in COIs – but it

---

[4] There is a third variable: available resources. Typically this variable isn't part of a tradeoff; instead, the answer is usually "all we can get", and then the tradeoff occurs between the other two variables.

doesn't always make it above their cut line on the list of good things to do. At least one COI (TST) appears to be limited by resources, not the tradeoff between scope and time.

4. Known consumers with known needs are important for COI success

When it comes to persuading the PORs to participate, few things are as effective as a list of important operational users who would like to be consumers of the COI's new data sharing capability. We have observed this effect with the Air Force Cursor-on-Target [6] prototypes. Concept demonstrations are nice, but useable capabilities are far more persuasive, as satisfied consumers become an advertisement for subsequent spirals. Unanticipated users are always going to be an important part of the NCDS; however, at least for now, resources for the COI effort will be more easily obtained if there is an understood path to a capability used by known consumers.

5. System builders are usually important contributors to COI vocabulary work

A COI vocabulary represents the community's shared understanding of the terms they use to describe information and to define data within their subject area of interest. This shared understanding almost always has to be recorded in some tangible format, which may take the form of a taxonomy, ontology, collection of data models and data elements, or some combination of these [7]. COIs use this documentation to teach new members what they need to know about the common vocabulary. They use it to support tools that help the members do their jobs – which could be describing desired information flows, or discovering new information, or implementing application-level exchanges, or understanding the information they receive each day. However, the fixed form of the vocabulary is not nearly so important as its existence as shared knowledge in the minds of the COI members. What really counts is the community's shared understanding, not the documentation used to write it down. Building this common knowledge is an important part of a COI's "vocabulary work". Without it, all you have is useless documentation.

All of the COIs examined for this study have some need for application-level, machine-to-machine data exchange. These application-level exchanges require a shared understanding that extends to the implementation level. That is, the people in the COI have to know enough about the applications to understand precisely what data has to be shared between them. They have to know enough about that data to establish a semantic match, and enough to cope with any mismatch in the way it is represented by the applications. No matter how the data exchanges are implemented – via standard data within the applications, a common interchange format, or on-demand data mediation – somebody has to know how to map the application internal definitions to the community vocabulary. These people may not be the official application developers, but they do have to know a lot about what the application needs to work.

Most of the nine COIs studied have included people with this implementation-level knowledge on their vocabulary teams. The TST COI is an exception: While the TST vocabulary team contains several competent software developers, most of the implementation-level knowledge of the participating applications lies with people in a separate team. This division appears to be impeding progress by creating organizational boundaries between people who should be working together more closely.

6. Most COI efforts do not have time to create or learn large new vocabularies

The important part of a COI vocabulary is the shared knowledge held by the COI members. Learning this knowledge takes time. This is true even if the proposed vocabulary is already developed – unless the COI members already know it, they'll have to learn it, and this takes time and effort. Most COI efforts attempt to produce useful results in a short time, typically nine to twelve months, and so they do not have the time to develop or learn a large, comprehensive vocabulary. Instead, COIs typically focus their vocabulary work on building a shared understanding of a small intersection of the data exchange needs, especially those essential to the goals of the current spiral.

We have observed this pattern in several COIs. For example:

- The BFT COI based its spiral-one vocabulary on the Cursor-on-Target schema, which at its core requires a shared understanding of only fourteen data elements.   With extensions, the BFT vocabulary was between two and three times the size of the CoT core – still much smaller than a vocabulary defining all the information that might be exchanged between participants.
- The GFM COI is adopting a subset of C2IEDM [5] for its shared vocabulary. This subset, named "GFMIEDM", is noticeably larger than the BFT vocabulary. However, it defines only the information included in the current COI spiral, and subsequent extensions are expected.
- The NETOPS COI is considering adoption of a subset of the Distributed Management Task Force (DMTF) Common Information Model (CIM) – the subset required to define information included in their current spiral.

7. Vocabulary overlap among COIs is inevitable

---

[5] Command and Control Information Exchange Data Model (C2IEDM). More information is available at www.mip-site.org

Vocabularies overlap when distinct COIs include the same real-world thing in their subject-area of interest. We see many examples of overlap within the examined COIs. And as the COI efforts expand the scope of their data sharing, those areas of overlap will only increase. There is no reason to believe that a perfectly separation of COIs is even possible. Certainly, any attempt to eliminate vocabulary overlaps by "getting the COI boundaries right" will take a great deal of time. This is not compatible with the demand to quickly produce useful results. We conclude that vocabulary overlap is inevitable.

COIs will naturally choose different names, structures, and representations for the data that describes the things of mutual interest – unless the members make the effort to agree. This effort also requires more time than will fit into the typical twelve-month schedule. Therefore, we should plan to cope with diverging, overlapping COI vocabularies. One consequence is that most PORs should plan on participating in multiple COIs, not just one.

8. Core enterprise services are not well understood

All of the COI pilots have been told to use the GIG core enterprise services in their data sharing prototypes. To do so, they need to know precisely what core services will be available. They need to know about the interfaces and standards that will be supported. They need to understand the deployment schedule. No COI reports this level of understanding. Instead, they report that this knowledge is very hard to discover, and that they must invest scarce resources and time into gathering the information they need.

9. Access control decisions need more attention

The NCDS goal of data accessibility does not mean an end to access control restrictions; instead, it means that access restrictions will be based on deliberate policy decisions (which may be quickly changed), and not on accidents of implementation (which change very slowly). Access control may then be divided into two concerns: the authority of the person who makes the decisions, and the implementation through which those decisions are enforced. We have observed some minor COI interest in access control implementation. We have not observed any COI thinking about who will have the authority to decide. As COIs increase in number and scope, as they make more data available across system and organization boundaries, this will become an important problem.

Section 3

Recommendations

Based on the previous observations, we offer the following recommendations concerning COIs in the NCDS. Some of these are process suggestions directed to existing and new COIs. Some concern the general role of COIs in the data strategy, and are directed to OASD/NII.

1. Gather more experience before defining a fixed COI process.
2. New COI pilots should build deployable capabilities for known consumers.
3. Make sure COI efforts have a powerful champion.
4. Start the COI clock after the champion is in place and the scope established.
5. Always consider the "loose-coupler" approach when developing COI vocabulary.
6. Don't try to define implementation-level vocabulary without implementers.
7. Don't partition a COI implementation effort into separate teams
8. Teach COIs about core enterprise services; don't make them learn on their own.
9. Think about information management governance, not just COI cooperation.

1. Gather more experience before defining a fixed COI process.

Eventually we will want to write an instruction and set of manuals for the 8320.2 directive. However, it is too early in the NCDS acceptance path to define a fixed COI process. Instead, we need to experiment and gather experience from the early adopters. We also need guidance handbooks that describe what it means to comply with 8320.2 and supply suggestions and recommendations for doing so.[6]

2. New COI pilots should build deployable capabilities for known consumers.

We have reached the point in the NCDS acceptance path where we need consumer-driven COI efforts, targeted at providing valuable information to known operational users. We are past the point where mere concept demonstrations are needed. We have not yet reached the point where producer-driven, "just build it and they will come" efforts will produce an acceptable return on investment. It is time to show tangible, deployable benefits, and learn the lessons which will come with that experience.

---

[6] The new 8320.2G, *Guidance to COIs for Implementing Net-Centric Information Sharing*, is in draft.

3. Make sure COI efforts have a powerful champion.

At present, COI efforts need a powerful champion to succeed. These champions might have authority over the known information consumers, thus representing them directly. Or they might take responsibility for providing a new information capability to the consumers. They must have effective control or influence over resources for implementation. Instead of "build it and they will come", the slogan should be "build this for me, and then others can use it".

4. Start the COI clock after the champion is in place and the scope established.

As a risk-reduction measure, COI efforts should usually be expected to produce results in a short time. Twelve months is a good standard for pilot efforts (though longer timelines may sometimes be necessary and worth the increased risk). However, the time clock should start running *after* the champion is in place, the consumers identified, and the operational threads understood. Those steps can take a great deal of time, cannot be bounded as easily as development, and do not create as much risk of failure as do long development efforts.

5. Always consider the "loose-coupler" approach when developing COI vocabulary.

The Cursor-on-Target (CoT) prototypes have popularized the "loose-coupler" approach to vocabulary development. Following this approach does *not* necessarily mean adopting the specific "what/where/when" CoT vocabulary. The approach calls for establishing a small shared vocabulary that covers only the information to be exchanged in the present spiral, thereby minimizing the total amount of vocabulary learning required. This shared vocabulary does *not* have to be adopted internally by the participating applications; it is instead used as an intermediate exchange format, or as a reference schema for data mediation.

Nothing in the loose-coupler approach requires developing a totally new vocabulary from scratch. If the small shared vocabulary can be extracted from one or more larger, established standards, so much the better. In this way, the loose-coupler approach serves to gradually introduce the larger vocabularies into new applications. Gradual progress is important, because COI efforts usually do not have time to accomplish large changes in vocabulary knowledge (involving many facts and many people). When large changes are desired, they should be approached as a series of small steps, each expected to produce some imperfections to be cleaned up later.

6. Don't try to define implementation-level vocabulary without implementers.

Any COI attempting to implement a machine-to-machine data exchange must include some people with implementation-level knowledge of the applications and the data to be

exchanged. System builders working for PORs have parts of that knowledge. Also, system builders will inevitably be involved in the implementation of the data sharing interfaces. System builders should be included in the COI vocabulary work. While you probably don't want a COI vocabulary team drawn entirely from the developers of the existing applications, it's also a mistake to keep them completely off the team.

7. Don't partition COI implementation efforts into separate teams

The *COI Handbook* divides COI activities into a cycle of three stages: one concerned with shared vocabulary, another with the information owners/producers, the third with the shared information space and services. Partly as a result, both the TST and GFM-AF COIs divided their efforts into three similar stages.

The TST COI established three working groups, to work in parallel: shared vocabulary, shared information space and services, and information owners/producers. The people defining the vocabulary are in the first group, those implementing the COI data sharing capability in the second. There is no common membership, and the working groups do not meet together.

The GFM-AF COI charter also calls for three separate COI teams. However, what has actually happened is better understood as a sequence of activities: first to develop a common vocabulary, then to identify the sources of the needed data, and finally to design and implement the data sharing capability. Many key participants have been involved in all three phases.

We believe that the GFM sequence of activities is better than the parallel, separate TST working groups. Close cooperation between the people involved appears to be important to the formation of shared knowledge.

8. Teach COIs about core enterprise services; don't make them learn on their own.

COI efforts should be actively taught how to use core enterprise services. They do not have time to research and assemble this knowledge for themselves. Instead, the learning material should be assembled once, and deliberately taught to all of the COI teams. This material should describe the core services that will be available during the current COI effort, plus interface details. Any available details of the future core services are also essential; these will help COIs choose temporary expedients that can be easily replaced when the real services arrive.

9. Think about information management governance, not just COI cooperation.

Consensus and cooperation are an important part of the NCDS. However, some information management decisions must be made by assigned authority. Will system A be capable of collecting facts X, Y, and Z, or not? Will the users of system A actually collect those facts, or not? Can people in organization B have access to those facts, or not? These decisions must be made even though there may be no choice that satisfies everyone.

It is time to think about those information management decisions, the roles of the people who will make them, and how those roles should be assigned within the enterprise. We recommend the following separation of roles for IM governance [8]:

- *Information owners:* organizations that exercise authority over the creation and maintenance of data. They are responsible for the production of data in the enterprise.

- *Shared information spaces* (or *infospaces*): collections of data intended to suit the needs of different groups of consumers of data. Each has a controlling authority, who represents and often has authority over the consumers.

- *Semantic communities*: groups of people who establish a shared understanding of terms and definitions in some subject-area domain; these are captured in a *common vocabulary*.

None of these are the same as a "community of interest".[7] COIs have never been held to overtly control resources or direct activities, and this is something we would not change. We believe that COIs should continue to act exclusively through the cooperation of their members and their respective organizations.

---

[7] A "semantic community" is the same thing as a "vocabulary COI", as described in section 2.

Section 4

Summary

COIs and the NCDS are comparatively new ideas, introduced in 2003. The DoD is still developing the policy, procedures, and technology details. In this paper we have summarized and generalized our experience with nine of the pathfinder COIs. Some of our recommendations are directed at existing and COIs:

- Always consider the "loose-coupler" approach when developing COI vocabulary.
- Don't try to define implementation-level vocabulary without implementers.
- Don't partition a COI implementation effort into separate teams

Other recommendations concern criteria for choosing new pilot COIs, or the role of COIs in the NCDS, and are intended for OASD/NII:

- Gather more experience before defining a fixed COI process.
- New COI pilots should build deployable capabilities for known consumers.
- Make sure COI efforts have a powerful champion.
- Start the COI clock after the champion is in place and the scope established.
- Teach COIs about core enterprise services; don't make them learn on their own.
- Think about information management governance, not just COI cooperation.

List of References

1. DoD Chief Information Officer, *DoD Net-Centric Data Strategy*, March 2003. http://www.defenselink.mil/nii/org/cio/doc/Net-Centric-Data-Strategy-2003-05-092.pdf

2. U.S. Air Force, *Air Force Information and Data Management Strategy Policy*, March 2004. https://www.my.af.mil/gcss-af/USAF/cms/AFMC/files/R2L-mZ0_0+63+-y-rOJ_AF_Information_Data_Management_Policy.pdf

3. T. Nabors, "Joint Meteorology and Oceanography (METOC) Data Standards", *Federal Database Colloquium*, San Diego, September 2000.

4. S. Renner, D. Hebert, S. Rainer, and J. Wilson, *COI Handbook: Practical Guidance for Communities of Interest Implementing the DoD Net-Centric Data Strategy,* MITRE Technical Report, December 2004.

5. U.S. Air Force, *A Guide for Communities of Interest Implementing the DoD Net-Centric Data Strategy and the Air Force Information and Data Management Strategy,* April 2005.

6. R. Byrne, "Getting DoD Linked", The MITRE Corporation, Jan. 2005. http://www.mitre.org/work/tech_papers/tech_papers_05/04_1260/

7. U.S. Air Force, *A Style Guide for Common Vocabularies,* April 2005.

8. S. Renner, "Net-Centric Information Management", in *Proc. 10th Int. C2 Research and Technology Symposium,* McLean, VA, June 2005. http://www.dodccrp.org/events/2005/10th/CD/papers/348.pdf