Enabling Secure Interoperability Among Federated National Entities: It's a Matter of Trust

C. L. Connors, Dr. M. A. Malloy, E. V. Masek
The MITRE Corporation

Abstract

This paper discusses issues relevant to accomplishing secure data sharing among federated national entities. This is an important, timely problem particularly in the defense sector since multi-national operations have become the norm. Similarly, the commercial marketplace today is a global one, requiring the exchange of information in support of transactions on an international scale. Traditionally, federated secure information sharing has been accomplished through bi- and multi-lateral information exchange agreements that require complex, human-centric and time-intensive processes to stand up or modify. The objective vision is to support such information sharing by establishing domain functional areas, and by creating discovery metadata standards that leverage security information to filter published data sets. A concurrent requirement is that only limited degradation of pertinent information can be tolerated to sustain common understanding. While XML alone is not enough to accomplish the vision, it is a critical enabler. We point out appropriate insertion points for XML technologies in meeting federated, secure interoperability challenges, and note as yet there is no "shrink wrapped" solution.

Introduction

United States' (U.S.) military operations are no longer conducted unilaterally or even bi-laterally with well-established allies like the United Kingdom. It is essential for the U.S. to interoperate with "coalition" forces as well as other international organizations as part of its new normal operational mode. For example, more than 50 North Atlantic Treaty Organization (NATO) and non-NATO nations have contributed military forces directly to the Kosovo-Bosnia peace-keeping function. Information exchanges must be conducted in such a way that sensitive or secure information is not inadvertently revealed to those lacking the authorization to receive it.

Similarly, the commercial marketplace today is a global one, where contracts, goods, services, etc. – and hence information – are transacted on an international scale. Trading partners must ensure their exchanges do not inadvertently reveal proprietary information that might compromise matters such as intellectual property or competitive market position, as two examples. In this paper, we propose ways that XML technologies can help facilitate exchanging critical information among multiple international communities where each employs heterogeneous information systems with widely different capabilities. But first, we must understand more details about the challenges and implications that this "new norm" imposes on information exchange.

Terms and Definitions

A simple definition of ***interoperability*** is "the meaningful exchange of data or information." It is not enough just to get information from point A to point B. The exchange must be conducted with sufficient context so that the purpose to which the recipient applies the received information is consistent with its use as intended by the originator. Security becomes a concern when exchanges cross the "trust boundary;" beyond this logical line of demarcation it is rarely possible for a producing entity to assume that all potential recipients are authorized to access all information they are capable of discovering and consuming. At a minimum, a successful ***secure*** information exchange implies that the exchange's content has been protected from unauthorized users, yet is sufficiently "complete" in that it still has meaning to the recipient to the extent it is possible to ensure this is so.

An ***entity*** is a generic term that is used to refer to an independent organization or business component. A ***federation*** consists of multiple autonomous entities that have a trust relationship with common governance over <u>some</u> aspects of their mutual functions or operations. For example, the Federal Reserve System (FRS) is composed of a small number of autonomous banks that agree on and then implement the FRS's interest rate changes. Federations also may be "tightly" or "loosely" coupled; the interoperability requirements in these arrangements often differ significantly. In general, tight coupling implies that interoperability solutions have been pre-engineered by requiring exchanges in an agreed format and conducted via rigorously defined interfaces; whereas loose coupling implies that *ad hoc* exchanges can be handled "on the fly" in an agile, responsive manner.

Contrast the meaning of "federation" with that of "enterprise." An ***enterprise*** is comprised of all the components that operate under the ownership or control of a single organization. An enterprise exercises governance over <u>all</u> aspects of its components' mutual functions and within the scope of a particular purpose or focus area. An enterprise may be a business, service, or membership organization; it may consist of one or several components and may operate at one or several locations. Examples of enterprises include operations in the U.S. Department of Defense (DoD) and Chase Manhattan Bank.

The term "federated enterprise" is sometimes used when discussing the concept of secure interoperability (data exchange) among federated national entities. In light of the above definitions, in particular the different governance scopes the terms respectively imply, we feel the term "federated enterprise" is contradictory and therefore we will neither define nor use it in this paper.

Assumptions

We want to stay focused on security issues so we will only mention other topics to the extent needed to propose appropriate ways ahead. For example, politics factors hugely in the discussion of information exchange across national boundaries. We assume that the legalities and formal agreements for information exchanges among national entities will be accomplished separately since the focus of this paper is on technical solutions.

State-of-the-Practice

Coalition interoperability presently is conducted through a complex collection of bi- and multi-lateral agreements that require frequent human intervention to negotiate and maintain. Secure federated exchanges at present are characterized by the following qualities:

- based on a "push" or "pull" architecture;
- pre-engineered at face-to-face meetings; and
- documented in various forms such as interface exchange documents (ICDs).

Such exchange agreements are not agile and often cannot respond in a timely fashion to unanticipated changes in the operational environment. In the defense sector, "timely" means "inside the enemy's decision loop." Analogously, in the commercial sector, "timely" typically means "soon enough that market advantage is not lost." In either context, the present pre-engineered approach to secure federated exchanges is a brittle solution where adaptation occurs very slowly, with minimal automation and maximal gray-matter involvement.

Successful information exchanges among international communities are not simply technological matters, but also for many years have relied on the creation and maintenance of supporting artifacts, including but not limited to common vocabularies that are the basis for meaningful information exchanges among participants. While the concept is simple – and many exchanging partners have developed shared vocabularies – practitioners continue to struggle with determining the proper scope and context of their data sets.

In the U.S. DoD, the prevailing construct for accomplishing harmonized vocabularies is the Community of Interest (COI); international defense partners are using a similar approach and are producing horizontal standards and vocabularies (e.g., NATO's Joint Consultation, Command and Control Information Exchange Data Model). In the commercial sector, international standards have been developed in key areas such as health care (e.g., Health Level 7) and meteorology. Vocabulary development nonetheless is fraught with issues because the breath of communities involved and the overlap of functional domains subsumed by them results in vocabulary terms that have a number of associated and equally valid meanings.

For example, the word "fire" is a valid term to a fireman, a soldier and a human resource specialist. Yet the meaning – and consequences – of "fire" to these respective individual roles are vastly different. Further, a person may have multiple roles. This is one simple illustration of how difficult it can be to establish a consistent and meaningful context for any given information exchange between partners just in the aspect of agreed vocabulary.

Discovery is another important aspect for advancing meaningful sharing among diverse functional areas like military operations and homeland security, as well as eCommerce where the potential pool of producers and consumers is global. Presently, *a priori* agreements establish exchanges that "push" or "pull" certain information to or from its source when crossing trust boundaries. ***Discovery*** generically refers to finding and retrieving actionable, decision-quality information "on-the-fly" as opposed to such pre-engineered approaches. A discovery service becomes significantly more complex when the best information must be derived from multi-

dimensional federated environments, each composed of a family of heterogeneous systems. A well-engineered, agile solution must address a number of complex issues, including how the security aspects of the information impacts whether it can be shared with a given partner.

Core Security Issues

Three fundamental security issues that impact federated sharing are:

- Authentication for user interfaces (i.e., viewing and accessing data) across multiple information sources;
- Federated secure discovery capability across multiple information sources; and
- Automated distribution of metadata about information sources and services to support discovery.

These in turn must be supported by enabling technologies and security processes (e.g., digital certificates, electronic signature verification, biometric validation, secure routing). In the next few paragraphs, we will discuss how decisions regarding these core issues are becoming key discriminators for how secure interoperability among federated national entities is likely to unfold.

The Roles of "Role" and "Rules"

User authorization is a core security issue relevant to this discussion. Within the U.S., there is much concern over accomplishing the goal of "single log-on" so that participants can log on once and have their credentials propagate along with them throughout the information infrastructure. Decisions about authorization to share with respect to information and service accesses can be made based on the role of each participant, established at the time of single log-on. For example, due to their role, Portuguese air traffic controllers may be authorized to discover that a U.S. aircraft will be in Portugal's airspace during a particular timeframe, but not authorized to know that it's a stealth aircraft.

While the heuristics or "business rules" for many such exchanges have been figured out and documented ahead of time, the extent to which unanticipated exchanges or accesses can be handled on-the-fly across national boundaries is unclear at this time. One challenge is that many business rules are buried in the implementation logic of existing applications that have been built based on pre-engineered exchange agreements. Still others are written down in documents and manuals, or even "stored" in people's heads so that applying them during processing requires human intervention.

Methodologies must be employed in which all these rules – both the explicit and the implicit ones – are "mined" and exposed for management separately from the processes that use them and the data they impact. This sort of arrangement is pictorially illustrated in Figure 1. Only if the rules are poised for automation will it be possible to decide "on-the-fly" whether a never-before-requested information exchange with a partner having a given role is permissible from a security viewpoint. One way to accomplish this is by establishing a commonly agreed set of metadata that describes the data and associated business practices as a frame of reference for

expressing the rules. For example, the NATO community is moving towards an ebXML model. Additional details and approaches are discussed below.

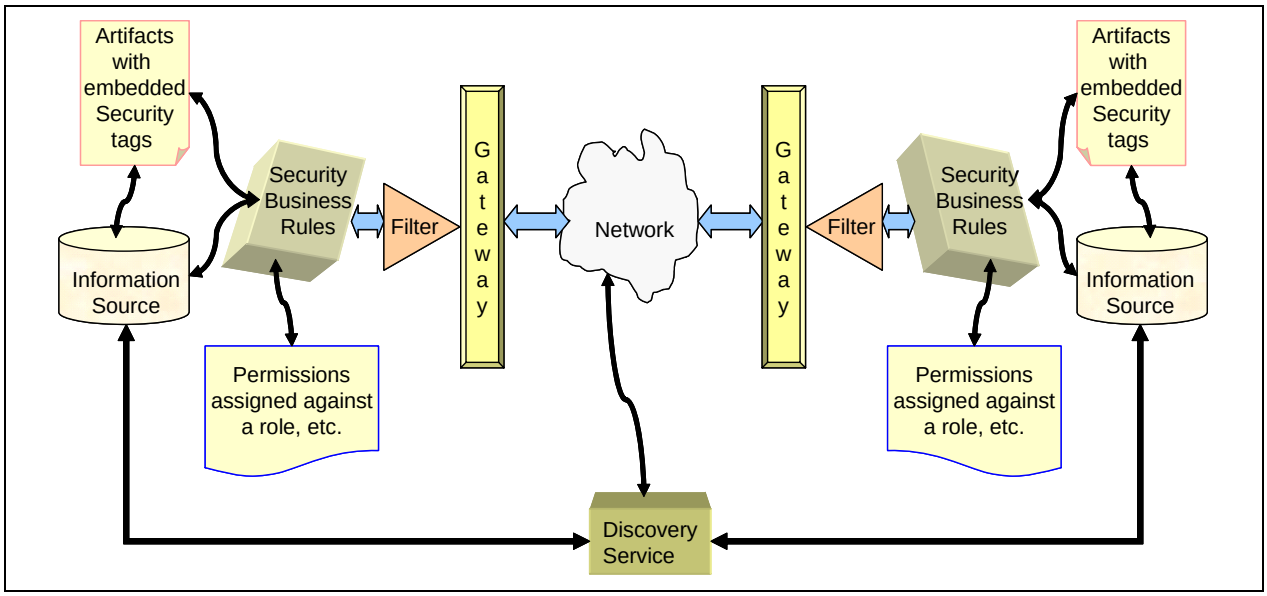

Figure 1. Rules and embedded security tags support automated secure sharing

Packaging Data

As mentioned earlier, security concerns impact the ability to exchange meaningful data when the exchange crosses a "trust" boundary. The current security attribution process commonly in use provides a single classification attribute (e.g., "Secret," "Top Secret") to a pre-engineered aggregation of data. At the one extreme, after filtering out the data to meet security requirements – whether based on the potential recipient's role or some other set of criteria, such as international laws – a recipient may get all the data requested; at the other extreme, they may get none at all.

To support secure, federated information sharing, a methodology must be developed that allows more granular, and more unanticipated, data views to be shared. This will progress us away from pre-engineered information packages and the restrictive "all" or "none" response. Optimally, then, each most granular data element must have its own associated security attributions. The methodology must also address the following concerns:

- When creating *ad hoc* data aggregations, what is the minimal subset of data that still conveys meaning in the context of a discovery event?
- *Ad hoc* aggregations of individual data elements may accrue a new meaning, as well as a different security attribution. Therefore, these new aggregations must have their own security attribution documented and perhaps even further attributed than its constituent parts. Such supplemental attribution, including pedigree metadata, can factor into deciding at some future time whether an *ad hoc* aggregation and its security attribution has become "stale" or no longer invalid.

Packaging Capabilities

Similar considerations apply to packaging services to support federated sharing. Obviously this is true when the services in question are "data services." But the new operational norm also means that more sharing of other types of services may be required across national boundaries as well.

One approach being pursued at present is to develop Service Level Agreements (SLAs) to document and support pre-engineered exchanges between Programs of Record (PoRs). Alternatively, "service wrappers" are used to deploy existing systems as "black box" products through a well-defined interface that hides any underlying implementation details. As indicated for data aggregation, services also must include security attributions. Similarly, as services are re-purposed beyond their original intended use (i.e., reused and/or employed in new aggregations), new security characterizations may be needed, and the security attributions of the new orchestrations documented.

Infrastructure

A number of key infrastructure components already in various stages of deployment and maturity must be factored into secure federated interoperability approaches due to the substantial investment stakeholders have already made in them. For example, international organizations such as the Red Cross and NATO are required to rely on an existing international federated infrastructure to support their networking requirements.

The U.S. DoD is implementing a major change to its existing information-sharing infrastructure, including the hardware, software, services and personnel that currently support exchanges. This is the DoD's Global Information Grid (GIG), which will increase the capabilities of the existing DoD infrastructure while retaining a decentralized network management model as it crosses enterprise domains. The GIG is intended to improve DoD's ability to provide rapid, accurate computer-based information exchanges across the spectrum of users, including disadvantaged users in the field. The DoD model includes the idea of registries where relevant metadata for the discovery and employment of data and services are posted for discovery and access by authorized participants. A related though broader effort is the Federal Enterprise Architecture Data and Information Reference Model, which promotes the common identification, use, and appropriate sharing of information across the Federal government.

What is the Vision?

To respond to the core security issues cited above, and to evolve from the baseline of state-of-the-practice approaches, the vision for secure federated information sharing must minimally constitute an equivalent capability as present. In other words, the migration infrastructure must support the documentation and fulfillment of appropriate data and services attributions that satisfy pre-engineered information exchange requirements, including a synchronous, *a priori* vetting of security concerns when it is practical and expedient to do so. Logically, we can think of this "as is" approach as a kind of smoothly, well-defined puzzle that supports thinking about exchanges in the context of fixed, pre-negotiated boundaries, as shown on the left in Figure 2.

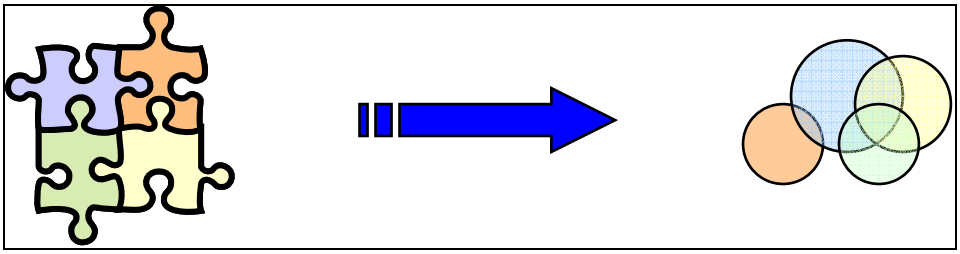

Figure 2. The "As-Is" versus the "To-Be" Information Sharing Landscape

The migration path to the vision should additionally leverage available technology so that solutions are poised to go beyond simply re-packaging the current level of capability (using XML or any other technology). To automate the security negotiation – thereby minimizing the need for human involvement – we must drive the security attribution down to the lowest level of granularity supported in the exchange model, so that filtering who-can-have-what can be done "on-the-fly." Logically, we can think of this "to be" approach in terms of a Venn diagram, as shown on the right in Figure 2. Each circle represents an entity's information assets; where two or more circles overlap represents where mutual sharing is permissible. In actuality, the relative positions and overlaps of these circles are almost constantly changing.

Based on the foregoing discussion, the future infrastructure must support posting appropriately attributed data and services in shared spaces. This must enable potential partners to asynchronously identify and agree to their reuse for exchanges in a way that is consistent with their mutual needs and security concerns, with minimal human intervention. In addition, discovery and access must be able to proceed in an agile fashion that can respond to unanticipated needs and changes in the operational environment. In other words, it must be possible to build new aggregations of data and services with enriched, extended and/or alternative meanings as new needs emerge.

A Solution Framework

Prior attempts to accomplish secure interoperability across traditional boundaries have included:

- One system to rule the world
- One database to rule the world
- One schema to rule the world
- One model to rule the world
- One data capture process to rule the world

These previous attempts have all gone awry! While each may have been sound in theory, in every case it has proven impractical, if not impossible, to gain consensus to a singular solution across even one organization (for example, DoD), let alone multiple international participants where trust boundaries must be crossed. Changes pose additional challenges as the solution must continuously adapt to new contingencies, capabilities and information exchange requirements.

Based on our analysis, we believe a framework to support a secure federated information-sharing environment includes the following key characteristics:

- It harvests and leverages general concepts from ebXML principles to ensure compatibility with international partners.
- It captures and manages data, services and business rules as separate but equally important parts of the whole solution space.
- It tolerates diversity in the data and metadata representations of the participants.
- It embodies a semantic layer (e.g., interrelated vocabularies, taxonomies, ontologies) that describes component vocabularies and provides for mappings between them.
- It drives metadata attribution down to a sufficiently granular level to support assessing whether specific exchanges are lossless from a contextual viewpoint and permissible from a security viewpoint, utilizing the participants' "roles" to guide these decisions.
- It provides participants the means to discover and reuse metadata artifacts in new and different ways from the data and service aggregations they were originally designed to support, with minimal human intervention.
- It enables participants to pre-engineer specific exchanges in terms of data and service aggregations and the business rules that constrain them when it is expedient to do so, and to create other as-yet-unanticipated *ad hoc* exchanges as needed "on-the-fly."

How XML Can Help

At this point, it should be clear that the insertion of XML or any other technology is but a small part of the solution to the federated secure interoperability problem. However, we do view XML as a key enabling technology that can help stand up many facets of the solution framework we outlined above. The key insertion points we see at present are the following:

- In the U.S. (as well as planned for NATO), registries are being employed to store metadata about significant entities' vocabularies or artifacts (e.g., messages, documents). Without exception, all of these registries are using XML artifacts (e.g., Schema, XSLT) to represent the metadata.

- Business rules can be used to handle security information for aggregation of data and metadata, and to determine if an aggregation is meaningful. (NOTE: Some aggregations of data are not meaningful). XML technologies like RuleML, SWRL and the Rule Interchange Format (RIF) will be helpful for formalizing, managing and automating rules.

- Ontologies can be built and interrelated using technologies like RDF, OWL and DAML. In addition, these implementations can include security and other assertions about each data element or aggregation.

Additional Issues

Suppose the underlying technology of the vision is in place. This would mean we have successfully used XML to capture security information at the most granular data element level, and we have stood up the capabilities to enable both pre-engineered and *ad hoc* information

exchanges to take place in an agile, timely manner even for participants with no prior history of exchanges based on their respectively roles. There are still more issues that must be addressed.

The application of rule sets to support security related decision-making is one of the most critical – and most difficult – critical path items in attaining the vision. For example, registry business rules and operations are diverse, and whether and how security information can be propagated among them must be considered carefully. Another challenge is that horizontal standards are not yet mature enough to completely address every facet of the secure federated information sharing problem. The security markup standards themselves are not-yet-ready-for-prime-time, nor are standard rule languages except in fairly restricted contexts.

While achieving the ability to document security attributes down to the most granular data element level will greatly increase the likelihood we can achieve the vision, additional thought must go into determining how this technology can be effectively applied. As shown in Figure 3, the current thinking is that at least three pieces of security metadata are needed per data element:

- *Security policy*: rules for protecting information against unauthorized disclosure while maintaining authorized access. (NOTE: This will require establishing a common understanding of handling requirements so policy mappings can be created across security domains.
- *Security classification*: markings that indicate the sensitivity level of the information (e.g., UNMARKED, UNCLASSIFIED, RESTRICTED, CONFIDENTIAL, SECRET, TOP SECRET).
- *Security category*: non-automated, specific sensitivity, dissemination, or an informational markings (e.g., administrative markings [MANAGEMENT, STAFF], release categories [RELEASABLE FOR INTERNET TRANSMISSION]).

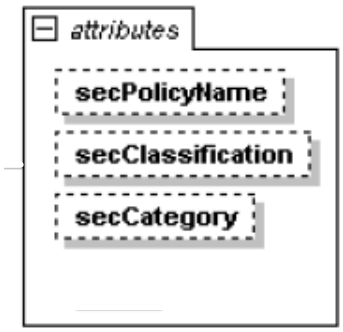

Figure 3. Notional Security Attributions

The amount of metadata required to attribute all assets down to this most granular level is, in a word, scary! A recent Joint Automated Metadata Tagging Pathfinder [1] demonstrated that the quality of manually inserting tags is low and this approach is infeasible given the huge volume of sharable assets that potentially needs to be processed. Automated tagging methodologies are therefore essential to the successful realization of security attribution requirements, as is the development of the elusive "universal metadata layer" that will enable meaningful cross-domain tagging relationships to be established.

Finally, it should be noted that security attributes can change at any time: an ally may become an enemy, or vice versa. This will either require document tagging to be revised, or new policy rules to be developed and deployed. In other words, it's not necessarily a "mark-up once" proposition, nor can sharing policies be expected to be articulated definitively once and never revisited.

Summary and Way-Ahead

In this paper, we explained what we mean by federated secure information sharing and why it's a timely and relevant problem particularly in defense contexts. We pointed out this is a relevant problem in the commercial sector too due to the effects of a global marketplace and eCommerce. We outlined traditional and state-of-the-practice approaches to such exchanges. Based on our analysis, we proposed key characteristics of an XML-enabled solution framework, particularly in terms of security attribution and the ability to support *ad hoc*, unanticipated data and service aggregations beyond the pre-engineered ones possible today. We did not offer a complete solution, however. Additional work is needed to address associated issues and challenges.

For the technical challenges, while XML is a critical enabler, horizontal standards are as yet immature in several areas, including security markup and rule languages. Still other challenges are non-technical. Human elements are possibly the greatest unknowns that will continue to impede secure federated data sharing. Participating entities must be prepared to embrace not only technology innovation, but also to venture beyond their comfort zones through process, organizational and even cultural evolution. Bureaucratic turf wars are likely to continue for some time: information technologists versus operational and business practitioners; those who want to maintain control of information and those who think interoperability means they should have free reign to discover and access it. From a U.S. perspective, we must ensure that security decisions we make do not create new interoperability challenges with our international partners who appear to be committed to ebXML-based solutions.

We will anticipate with interest the alternative potential outcomes for how secure, federated information sharing may impact data and metadata management life cycles in the complex international community. Areas worth watching include: how appropriate roles for standards and metrics will help or impede progressing the vision; the feasibility of building and effectively leveraging federated metadata repositories; and fostering information flow as state-of-the-practice evolves from serving the needs of "pre-engineered, need to know" exchanges to supporting the emerging "ad hoc, need to share" paradigm.

Disclaimer

The views, opinions, and conclusions expressed in this paper are those of the authors and should not be construed as an official position of the United States Department of Defense nor of The MITRE Corporation. All information presented here is unclassified, technically accurate, contains no critical military technology and is not subject to export controls.

References

[1] Assistant Secretary of Defense (Networks and Information Integration) / DoD Chief Information Officer, *Implementing the Net-Centric Data Strategy Progress and Compliance Report*, July 2006.