ASMG
ADVANCED SYSTEMS
MANAGEMENT GROUP

Information Sharing and Safeguarding

Policy-Driven Data-Centric Solution for Structured Messaging

M.Abramson /E.Penwill

Contents

Introduction

This white paper outlines a Policy-Driven Data-Centric approach to the sharing and safeguarding of information within a structured data and messaging environment such as the National Information Exchange Model (NIEM), Common Alerting Protocol (CAP), and Emergency Data Exchange Language (EDXL). The approach provides a method for translating Information Sharing and Safeguarding (ISS) policies into a set of rules that can be utilized by standard systems and services.

ISS Policies — Principles, rules, and guidelines formulated or adopted by an organization or agency to enable the exchange of information between internal users (individuals or information systems) or external entities (e.g., Allies; Mission, Coalition & Business Partners; Other Government Departments; Non-Governmental Organizations; Private Sector Organizations; and the Public) and typically published as a policy instrument (e.g., legislation, regulation, executive policy or directive, memorandum of understanding, or service level agreement). Policies are designed to influence and determine decisions and actions, and all activities take place within the boundaries of the policy.

ISS Procedures — Specific methods employed to express policies in action in day-to-day operations of the organization. The combination of ISS policies and procedures ensure that information is handled, shared and safeguarded in a manner that conforms to the direction of stakeholders, decision makers and/or governing bodies.

This paper focuses on structured data because of the unique challenges it presents in the operational environment. Whereas unstructured data elements (e.g., documents, reports, email, text messages, image and video files) remain static in structure and format after publication, structured data is repeatedly assembled (aggregated, transformed, marked and filtered), disassembled, processed (parsed, transformed, marshalled) and stored in real-time to address various community and user needs. In addition, structured data often takes different structures and formats when at rest (stored), in transit (being shared) and in use (being processed and presented). The conversion of data from one structure and format to another may require the execution of 10s, 100s, or 100os of rules. These rules in turn, must also ensure that the resulting data conforms to an increasing number of security and privacy rules. The current practice of encoding these rules within the interfaces of information systems, services and applications imposes several challenges:

1. Inability to adapt information systems to rapidly changing operational requirements;
2. Inability to certify information systems;
3. Inability to provision quality (e.g., timely, accurate, complete, relevant, digestible, and concise) information to decision makers;
4. Inability to retain the institutional knowledge needed to develop, deploy and sustain interoperable information systems; and
5. Inability to control the life-cycle costs of interoperability solutions.

In order to overcome these challenges, a new and enhanced set of practices, techniques and tools need to be developed. The Information Exchange Framework (IEF) is an Object Management Group (OMG) initiative to address these challenges. Its goal is to deliver a Policy-Driven Data-Centric solution to the interoperability (e.g., information sharing and safeguarding and responsible information sharing) requirement of communities, including: first responders, Emergency Management, public safety, national security (e.g., Intelligence Sharing, Critical Infrastructure Protection, and Cyber Security SA), and the military.

It is now common practice to add or bind metadata (a form of structured data) to unstructured data. (e.g., record & document management and smart data solutions). This practice is intended to improve the management, discovery, sharing and safeguarding of the data holdings. To be effective, this metadata will require the same application of rules as do other forms of structured data, and thus solutions described in this document will be effective for sharing and safeguarding these forms of structured data as well.

Interoperability Challenges

The malicious or inadvertent disclosure of sensitive (e.g., private, confidential or classified) information can have serious repercussions to the security of a person, organization or nation. In response to this challenge, most organizations has been focused on establishing "need-to-know" or "need-to-protect" frameworks around their sensitive information. That is, sensitive information is only made available to those persons with appropriate clearances and a "need-to-know"; those persons that can demonstrate they need the information to conduct official duties.

Increasingly, the contrary position is strenuously argued: that the failure to disclose information can have serious repercussions to the security of a person, organization or nation. E.g.:

- A Pharmacist's or Doctor's inability to access patient medical history can result in the prescription of a drug, even an over the counter drug, that is contra indicated based on an existing condition or previously prescribed medication.
- A failure to notify first responders to an imminent risk, such as: structural failure of a building or the presence of toxic chemicals, could result in serious injury or death.
- The failure to provide that critical fragment of data, to an Intelligence Analyst that identifies an impending terrorist attack could impede efforts to protect citizens or critical infrastructure.

Numerous reports[i,ii,iii,iv], Strategies[v,vi], directives[vii] and Memorandum[viii] communicate the critical need to enhance information sharing across a broad spectrum of public, private and military domains. There is no more critical a need for information sharing than during national or international emergency or crisis response operations, when coalitions or partnerships of independent agencies must dynamically form. On the onset of an emergency/crisis, these partnerships have an immediate need for information: to provide a shared understanding of the situation and the environment (e.g., situational awareness); to enable collaborative planning; or enable monitoring and coordination (command and control). This need exists whether it is a local emergency (e.g., major fire or traffic accident), weather event, natural disaster, terrorist incident, humanitarian relief, or combat operation. These partnerships or coalitions often form through international or interagency cooperation, rather than well-defined agreements or memoranda of understanding. Each participating agency can offer whatever resources it can muster to support the given emergency/crisis. These situations can occur suddenly, simultaneously, and often with little warning. At times, participants to one partnership/coalition to address one crisis may be adversaries in another, raising difficult security issues with respect to information sharing.

In the absence of clearly defined policies, procedures and training, disclosure of mission critical information may or may not occur: placing individuals, organizations or national interests at risk. The core challenges include:

1. How to balance the need to share with the need to protect;
2. How to enable users to make these decisions during the course of complex and rapidly changing situations;
3. How to rapidly deploy and integrate ISS capability;
4. How to rapidly adapt information systems to planned and unplanned real-world events; and
5. How to translate general policies into sets of rules that directly relate:

a. To the data/information environment;
b. To the mission, operation or situation being addressed;
c. To the variation in roles, responsibilities, and levels of trust within a coalition or partnership; and
d. To the need to adapt to dynamic real-world events and operational context (e.g., scope, severity, resources, threats and risks).

Organizations must develop and maintain capabilities that balance the sharing and safeguarding of sensitive information throughout its lifecycle. Information security policy must detail how sensitive information is marked, stored, shared and destroyed in a manner that decisions can be automated and audited. The operational tempo of these environments and complexity of information domains make it virtually impossible for users to make the correct share/safeguard decision in real-time.

Future information security programs and systems must address capabilities that: provide for the rapid deployment of mission ISS capability; enable rapid adaptation of ISS patterns and content to dynamic real-world events; enable the use of information as a resource (/force) multiplier; and enable real-time ISS incident reporting and forensic auditing.

Operational Need

Increasingly, stakeholders seek to deploy ISS capabilities that:

a. Enable responsible information sharing;

b. Employ information as a strategic, operational, and tactical asset in the delivery of information and decision advantage;

c. Provide decision makers with access to quality (e.g., accurate, complete, authoritative, relevant, timely, digestible/usable, concise, trusted, and secure) information;

d. Extend agency boundaries; to enable semantic interoperability, shared situational awareness (SA), collaboration, and interoperation;

e. Establish a common shared infrastructure (to reduce cost) based on open systems and standards (to increase vendor competition and innovation); and

f. Adapt (manually, and/or automatically) to the dynamics of real-world operations (including executives'/commanders' intent, situational awareness, plans, partners / organization / coalition, role, responsibilities, threat, scope /scale, severity, networks and Communications, and Operational and Tactical Status).

Stakeholders recognize that conventional strategies, tools, and technologies struggle with the dynamics and tempo of modern operational environments. Conventional strategies and solutions do not provide the flexibility and agility to define, integrate, and enforce rules governing the packaging and processing of structured data during all phases of an operation. Nor do these solutions provide the ability to rapidly adapt their operation to that of the current operational context (e.g., threat, risk, resource configuration & availability, and organization). The goal of providing '*the right information, to the right place/person, at the right* time' continues to elude many operational communities. In actuality, few

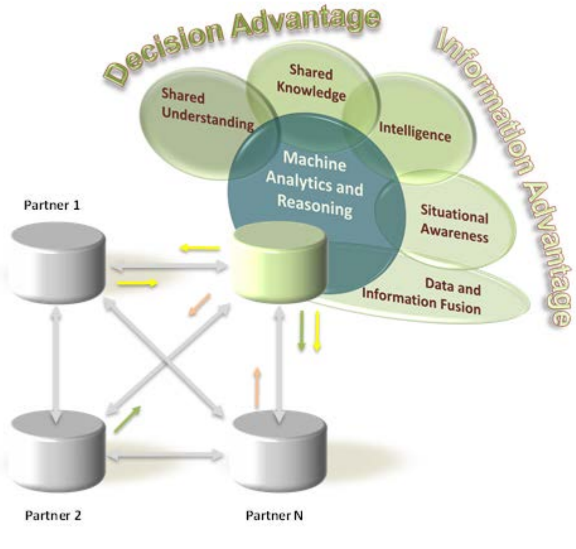


Figure 1 -Developing Information & Decision Advantage

architectural, engineering or software development techniques, and technologies enable agencies/communities to establish an information environment that fully defines:

The Right Information

There is no single correct answer to the meaning of the right information. The '*right information*' is defined by the situation or context (e.g., role, responsibility, threat, risk) in which the decision maker finds himself/herself and the decision at hand. It is dependent on criteria that evolve or rapidly change over the course of the operation: e.g., location, communication capability, and capacity, differing policy rights of the senders and receivers. Most importantly, the propensity of many conventional strategies to equate more information with quality, or the right information – has led to data overload – meaning that it is increasingly difficult for users to identify, and discover, let alone, use the 'right information'

Further complicating this issue, when addressing structured data, there is likely no single definition of what the 'right information' is. Data and information elements can be combined and recombined to serve a variety of operational contexts: e.g., role & responsibility, Quality of Service, degraded modes of operation, security, and privacy.

The Right Person /Place

Again, there is no single correct answer to the meaning of the '*right person/place*'. As a part of different missions, or over the course of an operation, individuals, organizations, agencies, and nations change their roles, responsibilities, and locations. The "*right person / place*" is a dynamic quality that is also a context qualifier on the '*right information*'.

The Right Time

The 'r*ight time*' generally means that the decision maker has the information, with sufficient time to consume the information and make the required decision.

To address stakeholder needs/challenges, the next generation of strategies, tools, and technologies need to:

1. Separate operational concerns from technical implementation;
2. Provide a continuous process that rapidly translates policy to operational capability;
3. Align information sharing and safeguarding concerns; and
4. Provide portability across multiple platforms and technologies.

Background

Structured Data

Structured data is "data that resides in fixed fields within a record or file. Relational databases and spreadsheets are examples of structured data". It simultaneously provides levels of rigidity and levels of flexibility. Its fields, structure and relationships are rigid and brittle, and for practical purposes, unalterable during the course of an operation. However, its data elements/fields/attributes can be combined and recombined to address multiple user needs. The following table provides an example of the representations of structured data.

Simple Structured Data Example:

(Challenge 1) This example provides fixed field options (e.g., flat file (spreadsheet) and Relational tables for the storage of data describing an operational Unit. From this example:

Flat File Version (Spreadsheet)

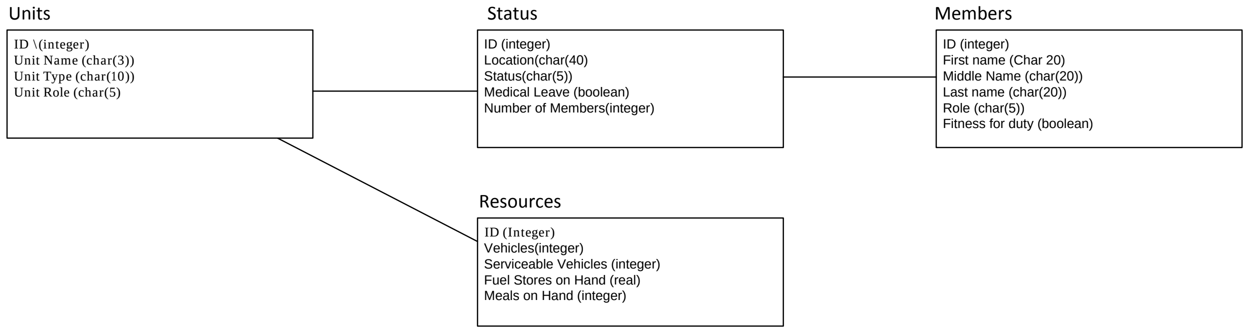
Index (integer)	Unit Name (String)	Unit Type (String)	Unit Role (Sting)	Location (String)	Status (String)	Medical Leave (Boolean)	Number of Members (Integer)	Member Name [0..*] (String)	Vehicles (Integer)	Serviceable Vehicles (Integer)	Fuel Stores on Hand (Real)	Meals on Hand (Integer)	

Table 1 - Spreadsheet

Alternate Flat File Version (Comma Separated File)

Index, Unit Name, Unit Type, Unit Role, Location, Status, Medical Leave, Number of Members, Member Name [0..*],Vehicles, Serviceable Vehicles, Fuel Stores on Hand, Meals on Hand

Relational Database Version



Units
ID \(integer)
Unit Name (char(3))
Unit Type (char(10))
Unit Role (char(5)

Status
ID (integer)
Location(char(40)
Status(char(5))
Medical Leave (boolean)
Number of Members(integer)

Members
ID (integer)
First name (Char 20)
Middle Name (char(20))
Last name (char(20))
Role (char(5))
Fitness for duty (boolean)

Resources
ID (Integer)
Vehicles(integer)
Serviceable Vehicles (integer)
Fuel Stores on Hand (real)
Meals on Hand (integer)

Figure 2 – RDBMS

Effort to Reduce Complexity

It is clear that each of these structures could hold the same data elements describing an Operational-Unit. It is equally clear that the application of multiple sets of business rules (e.g., queries, stores-procedures, and application code) would enable the successful exchange of data between these data stores. However, as the variations in the data structures increase, so would the number of interfaces needed to support the exchange. This form of peer-to-peer interfacing can evolve into N*(N-1) interfaces or (N*(N-1))/2 bi-directional interfaces (as illustrated).

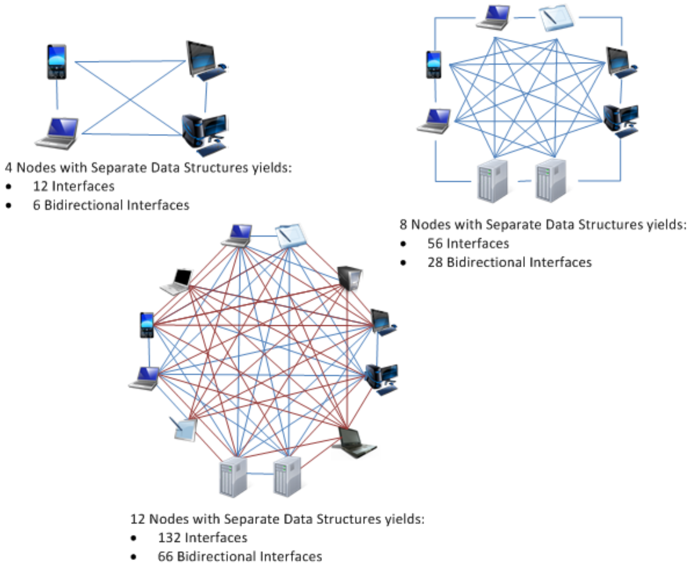

4 Nodes with Separate Data Structures yields:
12 Interfaces
6 Bidirectional Interfaces
8 Nodes with Separate Data Structures yields:
56 Interfaces
28 Bidirectional Interfaces
12 Nodes with Separate Data Structures yields:
132 Interfaces
66 Bidirectional Interfaces

Figure 3 - Peer-to-Peer Interfaces

The initial efforts to address this exponential growth in interfaces focused on the development of common data exchange models (e.g., Joint Consultation Command and Control Information Exchange Data Model (JC3IEDM) and later standard messaging using XML schemas (e.g., NIEM). This approach has the potential to dramatically reduce the number of discrete interfaces.

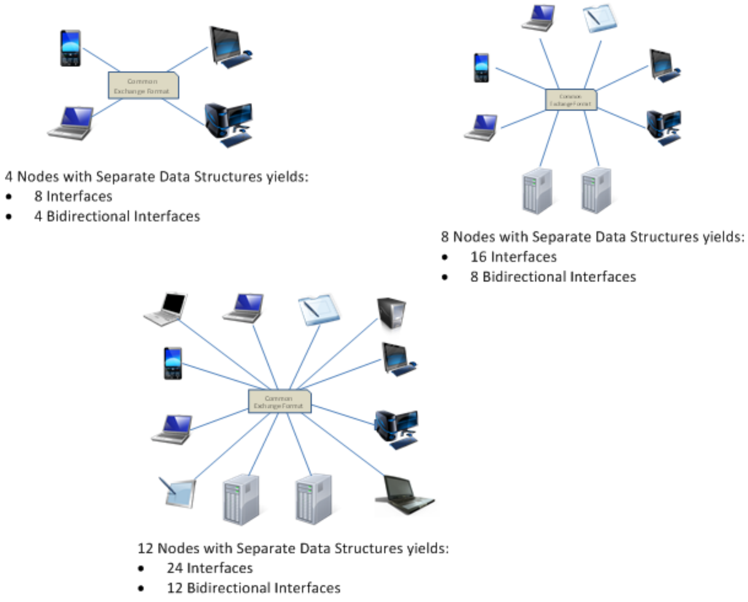

Common Exchange Format
4 Nodes with Separate Data Structures yields:
8 Interfaces
4 Bidirectional Interfaces
Common Exchange Format
8 Nodes with Separate Data Structures yields:
16 Interfaces
8 Bidirectional Interfaces
Common Exchange Format
12 Nodes with Separate Data Structures yields:
24 Interfaces
12 Bidirectional Interfaces

Figure 4 - Common Message Format

In both of these environments the participants to these data exchanges are responsible for developing and maintaining the bidirectional logic of the interface to ensure:

- The translation of their local data structure to the common structure; and
- The translation of the common structure to their local data structure.

Theses translations must also:

- Transform data values (integer ←→ string; [Male, Female] ←→ [M, F]; etc.);
- Provision data attributes that are not supported by a single local data structure, requiring the aggregation of multiple local data structures;
- Process attributes in the common structure not supported by the local data store structure; and
- Provision attributes required by the local data store but not supported by the common structure.

Reintroducing Complexity

In many instances, information exchanges must be tailored for each operational role or recipient in order to address privacy, confidentiality or security concerns and/or policy. Information is often aggregated to address operational needs and then must be filtered to address: Communication challenges (e.g., bandwidth constrains, quality of service), or information quality (e.g., filter for clarity, relevance). Most information technologies are unable to accommodate the often contradictory rules that must be applied and the operational decisions required to reduce or eliminate restrictions; resulting in silo'ed information environments that preclude the delivery of shared situational awareness, interagency collaboration, and information/decision advantage.

The following table provides an example where a local data structure (Organization Unit) is used to issue several data subsets: Situational Awareness, Logistics, Supply, and Personnel. The objective is to provide each operational role with exactly the information their data store can ingest and store.

Organization Unit Structure	Situational Awareness	Logistics	Supply	Personnel
• Unit Name • Location • Status, • Medical Leave • Number of Members • Member Name [0..*] • Vehicles • Serviceable Vehicles • Fuel Stores • Meals	• Unit Name • Location • Status • Medical Leave • Number of Members • Member Name [0..*] • Vehicles • Serviceable Vehicles • Fuel Stores • Meals	• Unit Name • Location • Vehicles • Serviceable Vehicles	• Unit Name • Location • Number of Members • Fuel Stores on Hand • Meals on Hand	• Unit Name • Location • Medical Leave • Number of Members • Member Name [0..*]

Table 2 -Data Segmentation

The interface must also be designed with the ability to redact the data as it is assembled based on privacy and security policy. In most instances, this requires the capacity to interrogate the data values and ascertain if that specific set of data values violates some aspect of policy contained in policy instruments that may include:

- Legislation;
- Regulation;
- International Treaties and Agreements;
- Organizational Policy;
- Memorandum of Understanding;
- Service Level Agreements; and/or
- Standard Operating Procedures.

As illustrated in the following example, a single data structure may hold information that is unclassified; unclassified but sensitive (e.g., private or confidential), restricted and classified (e.g., secret, top secret, and top secret special access). The data may be marked with its particular sensitivity, or not (as illustrated). The data may assume a sensitivity level of the application, system, file system or network within which it resides. It may not be sensitive on its own, but increases in sensitivity when it is aggregated with other data elements.

What further complicates issues surrounding the release of information is the requirement to protect the data in one context and share it in another. Some data must be protected from one participant in a community and not another based on established levels of trust or agreement. However a change context (e.g., role & responsibility, threat level, and severity of the incident) may require the lessening or removal of the restriction.

Based on the unmarked data below, few would argue that 'Value Set 1' has little sensitivity if inadvertently released, as it contains little in the way of information that may be on a corporate or government web-site. On the other hand, publishing information on a Special Forces unit that is currently operational, is likely classified, as its release could cause serious harm to members of the unit, operational outcomes, or national interests. This not so subtle example demonstrates that data structure often has little relevance to the sensitivity of its content, often, it is not until runtime and the actual data values are entered that we must view the data through a security or privacy lens.

Presentation of Data Entities

Flat File Record Structure (Organization Unit)	Value set 1	Value Set 2
• Unit Name	AP	SP-1
• Unit Type	Administration	Operations
• Unit Role	Accounts Payable	Special Operations
• Location	1010 10th street, smithfalls, MN	22.828846: 110.925286
• Status	Active	Active
• Number of Members	5	6
• Medical Leave	1	2
• Member Name [0..*]	fred,nancy,stan,bill,judy	fred,bill,stan,bob,steve,daryl
• Vehicles	0	3
• Serviceable Vehicles	0	1
• Fuel Stores	0	100
• Meals	0	24

Table 3 - Data Sensitivity

What if, the actual dataset was 50,000 records, and the data on the Special Force unit was inadvertently entered into a gateway service that publishes data to less protected information systems. What if the record was aggregated from multiples sources (Mosaic Effect) that were less protected because their data holdings were considered less sensitive – and operational integration with the other information systems was not contemplated. Could this data issue be identified? At Design? At runtime? During a forensic audit?

Can operators/users be expected to identify and address these complexities during operations, in time-sensitive real-world environments, as proposed by many organizations who integrate air-gaps (physical separation of information systems and networks)? Without support tools, can an operator/user be expected to determine the sensitivity of a single data record, or several 1000s of records, governed by multiple policy instruments in response to real-world events. Add the complication that no two events will likely involve the same participants, roles and responsibilities, locations, threat, severity, and operational objectives – and we can state almost categorically that:

- Though an excellent start, users need more than a common data structure;
- Users need a new approach and tools to translate ISS policies into a machine readable and executable form;
- Users need information systems that can be trusted to automate machine executable ISS policies they define; and
- Users need analytic tools to assess and audit the designs and enforcement of ISS policy.

Classifying Information

A key component to all discussion about broad-based information is the ability and capacity to identify the sensitivity of particular information elements and mark it so it will be handled (stored, exchanged and used) in an appropriate manner. For unstructured information elements, the author must assign (or bind) the classification and warning terms, or “caveat” markings to the element as a whole or selected sub-elements as specified by policy. These markings direct how an information element should be handled (e.g., stored, used, shared and destroyed) whether the handling is conducted through a manual or automated process. Without markings, individuals and/or systems cannot reliably understand the sensitivity or handling of the information element. Handling is fairly straight forward for unstructured information. The author marks the information element, typically a file, and the entire information element is handled in accordance with that single marking or set of file markings directly related to handling of that element. Most information exchange services (e.g., middleware), do not interrogate the internals of the file, and rely of the file level marking to determine if the file is sharable.

However, for structured data, the process is not as straight forward.

1. Who specifically is the author:
   a. The operator at the console;
   b. The information system or application that packaged the information element (message or document);
   c. The programmer that wrote the interface code; or
   d. The business analyst or system architect that wrote the specifications?
2. Can the process be automated:
   a. Can the operator be expected to read, assess and mark every generated message, or a system be trusted to assess and mark individual messages? and
   b. What are the rules to be applied to individual assemblies of data and information elements and how do they relate to the applicable policy instruments?
3. Can the process increase flexibility, adaptability and agility?
4. Can the standard structured data definitions (e.g., NIEM XSDs) be reused and still have privacy and security issues addressed? and
5. Will it cost more?

Who is the Author?

For the purposes of this document, the author is the policy decision/enforcement point responsible for packaging the data elements for the exchange. This means that we need to provide stakeholders and data stewards with a level of assurance that their policies (rules) are being enforced for every release of data. For this to occur, we have to assure the following conditions are met:

1. That the machine can effectively assemble and mark (/label/tag) aggregated sets of data in an effective manner;
2. That a machine can effectively enforce the policies associated within data marked with every combination of classifications, restrictions and warnings; and
3. That they are the final arbiter of whether or not their data is shared.

Unstructured documents are authored in human-time, typically by an individual or team of individuals with the aid of a software application (e.g., spreadsheet, word processor, or collaboration tool). The author(s) can be compelled by policy and process to conduct an assessment of the content and apply the appropriate markings at the document or sub-element (e.g., paragraph, or row) level. For images or video files, the marking would be applied at the file level- based on the content of the images. These markings would be bound to the document by the author prior to its publication or implied, based on the classification and protection level of the information environment upon which it was prepared and published.

However, with structured data, content is assembled in real-time and often in large quantities. It is understood that the sensitivity of data often increases as it is collated, aggregated, integrated or fused. As illustrated earlier in this paper, the pliability of structured data makes it extremely difficult to govern the content of an assembly of structured data. If the content is difficult to know for a human operator – an assessment of its sensitivity would be virtually impossible. This conundrum has many communities isolating data and not sharing the data needed to achieve mission objectives and outcomes.

(**Condition 1**) Machines are exceedingly good at processing patterns and elements within those patterns. Given that structured data is based on fixed patterns (e.g., schemas) and increasing the use of common or shared types and values (e.g., NIEM Schemas and Core, and the meta-data structure used to enable mediation, discovery or decision support), machines are likely better able to assess the content and then the sensitivity of a structured document (e.g., XML) in real-time than is a human operator. What is needed is a common language to convey, to a machine, how to assess the patterns and data values, and determine the markings that apply.

(**Condition 2**) This condition is not the focus of this paper. But assuming Condition 1 is met, the systems and services currently enabling the sharing of unstructured files would be applied. The resulting document (from the aggregation or assembly of the data elements) would be marked in the same manner as one containing unstructured data. Note: IEF Reference Architecture (IEF RA) will define a policy-driven data-centric approach to managing access and releasability to several forms of data (e.g., email, instant messaging /chat, file share, web services and structured messaging). The IEF RA will describe how this condition is met.

(**Condition 3**) To enable stakeholders and data stewards to trust that they are the final arbiter of the release of any data and that it is performed in a secure manner, the solution must support:

1. Policy (rules) must be:
    a. Traceable to the originating policy instrument;
    b. Align with the platforms, systems, services, applications, interfaces, and decision & enforcement points where each of their policies is enforced;
    c. Align with the specific information and ISS (e.g., missions, operations, communities of interest) domains to which they are applied.
2. Decision and Enforcement points must be demonstrated to perform only what is allowed under the policy set applied:
    a. What conditions trigger the activation of each ISS policy;
    b. What data is assembled for each exchange; and
    c. Which communication channel is used for each exchange.

3. Operational changes to the policies (rules) are recorded/logged and audited; and
4. Policies can be adapted, by designated/authorized users, to changes in the operational context (e.g., Strategic/Tactical intent, roles & responsibilities, severity, and threat).

The solution must demonstrate that the stakeholders and data stewards are in control of the environment.

** A government Employee may share a segment of private data with another government employee, for a period of ninety days, if the process is audited.

Can the Process be Automated?

This paper would end here if we say “NO”. However automation needs to be discussed within several different contexts.

First is the translation of policy instruments into machine executable instructions. This is the process of applying generalized policy instruments written in natural or legal language** to a specific data domain (e.g., situational awareness, personnel, finance, logistics, supply, and planning), operational domain (e.g., healthcare, emergency management, finance, critical infrastructure, cyber security national security, and military) and operational context.

The policy statement in the previous text box refers to “a segment of private data”. To determine what constitutes “a segment of private data” depends on the data domain and the assemblies of data that need to be shared. In many cases is also depends on the run-time values of the data elements in the assembly. This mapping of policies to data domain is currently a software design activity. The proposed approach separates the transformation of policies into machine enforceable rules from the software life-cycle and services infrastructure used to enforce them. The development of policy models (mappings) is integrated into the enterprise architecture. The Policy models define the data patterns governing the assembly, labeling and packaging of the data to be shared and which patterns generate sensitive (private, confidential, classified and legally significant) information.

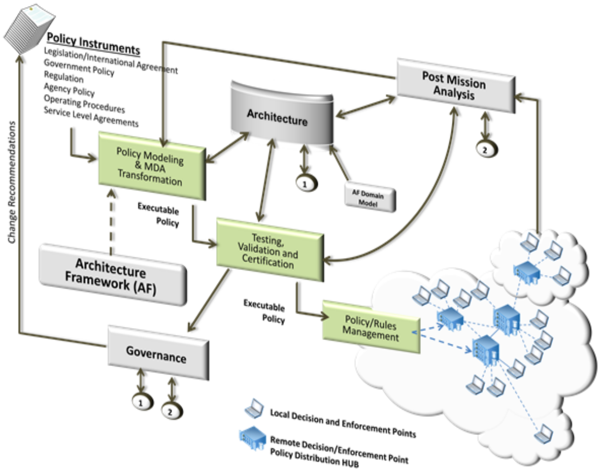


Figure 5 - Policy Life-cycle

The development of these patterns falls naturally into the domain of information architecture (data and information views in DODAF/UPDM). For this process the IEF defines a UML profile based on the Information Exchange Packaging Policy Vocabulary (IEPPV[vi]).

The architecture based approach to the development of policy models results in the alignment of policy application to architectural views and viewpoints depicting: Missions; Operations; Communities of Interest; Platforms; Systems; Services (decision & enforcement points); Applications; Communication channels; and Data interfaces. Model Driven Architecture (MDA) tools can then be used to serialize machine readable policy sets and the deployment patterns for the policies based on mission and operational views. The architectural metadata can be used to validate and verify policy models, and enable (facilitate) ISS modeling and simulation (M&S), testing, auditing, and certification.

Separating policy from its decision and enforcement points allows users to treat policy as data. As data, it can be stored, shared and safeguarded using the same approaches as any other sensitive data elements. It means that the process of disseminating policy can also be automated. The IEF has demonstrated exchange (/dissemination) of policy to update operational PDPs at runtime, using both XML and Binary message formats.

Logging and Auditing Change

The IEF proposes:

1. A formalized policy life-cycle, tightly coupled to architecture framework (e.g., DODAF), profiles (e.g., UPDM/IEPPV) and tools to fully document the application of policy to a specific domain and Model Driven Architecture (MDA) into machine readable and enforceable form. This process will produce the metadata needed to:
    a. Record/log the translation of policy instruments to machine executable rules for a specific data domain and mission/operation;
    b. Trace the application policy instruments to decision and enforcement points in the operational environments;
    c. Audit the implementation of the policy environment for conflicts and/or inconsistencies; and
    d. Establish a baseline for form modeling and simulation and post mission analysis.
2. A tamper resistant log of policy changes, policy decisions and transactions across the operational environment.

Through these capabilities, the IEF captures and maintains a full life-cycle history for ISS policies.

Increase Flexibility, Adaptability and Agility?

The keys to the IEF approach are:

Separation of Concerns — The IEF separates the business/operational concerns of stakeholders from many technology decisions:

1. The Reference Architecture defines an information sharing and safeguarding layer that is designed to integrate with the stakeholders infrastructure;
2. The translation of policy instruments into machine enforceable rules has been decoupled from development of the services and infrastructure that enforces it; and
3. IEF services:
    a. Ingest a machine readable set of user defined policies/rules at runtime;
    b. Enforce user defined Policies/rules;
    c. Interoperate with user defined IM, IT and IA services through standard interface specifications;
    d. Record each transaction in a tamper resistant log; and
    e. Enable run-time administration of policies/rules.

This combination of requirements allows a user, agency, or community to develop and deploy an IEF environment and then configure that environment for the specific mission or operational requirement when needed. The same requirements enable the user(s) to tailor and adapt its operation to changing mission needs.

This separation of concerns also:

1. Allows for the evolution of capability though the implementation of policy (data sets) rather than the enhancement, certification and deployment of software applications;

2. Eliminates the need to define and document all mission and operational requirements as part of the request for proposal (RFP) or system specification;
3. Allows stakeholders to repurpose information systems and infrastructure – rather than replace; and
4. Allows for the enhancement of individual components without the need for major development, integration and certification efforts.

Policy-Driven

The IEF defines a process (life-cycle) through which user defined policy instruments are translated into machine readable and enforced rules. This process results in full traceability from policy instrument to operation.

Data-Centric:

The IEF specifies policy vocabularies and services that enforce policies/rules against individual data elements. This provides the ability to develop policies to uniquely protect data and information elements (the assets we are seeking to share and safeguard). This will enable users to selectively expose and share content with individual users or communities from a single information store, maximizing the content that is available to share. Each data and information element is individually protected at rest, in transit and in use.

Will the Approach work with NIEM

The NIEM is an XML-based information exchange framework. The NIEM framework provides:

1. A standardized process for the development and harmonization of a set of common, well-defined data elements that enhance interagency interoperability;
2. A standard set of definitions and design artifacts known as the Information Exchange Package Documentation (IEPD);
3. A common XML-based data model called *NIEM core* that provides data components for describing universal objects such as people, locations, activities, and organizations;
4. Repository of IEPDs and support tools to develop, validate, document, and share NIEM artifacts; and
5. A governance model.

The IEF extends the NIEM framework in four (4) key areas. First, the IEF defines a common vocabulary for the specification of policies (set of rules) governing the packaging and processing of a XML document from local structured data store(s). The IEPPV[xviii] extends the documentation captured in the IEPD (e.g., Community Agreed Exchange Standard/Specification) by defining:

1. **Assembly** (aggregation, transformation, marking, and redaction)
   a. Structure and data transformations;
   b. Data and information element marking (Ownership, Security, Privacy, QoS, ...);
   c. Static and Dynamic Filters (Security, Privacy, QoS, ...); and
   d. Retrieval of data from local data stores.
2. **Processing** (parsing, validation, transformation and marshalling)
   a. Message and data disassembly;
   b. Domain and semantic validation;
   c. Data and structure transformation; and
   d. Entry of data into local data stores.

The *IEPPV*[xviii] is accompanied by a UML profile that provides users with the ability to develop policy models that can be transformed into a machine readable form and then enforced (automatically) by machine level decision and enforcement points. The policy models also expose the procedural logic typically concealed in the interface code of software applications. Focusing on code based solutions often results in the implementation of peer-to-peer interfaces that prove difficult to change in response to operational needs and are costly to maintain.

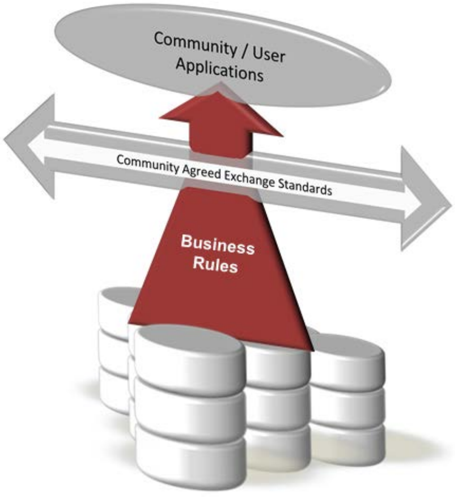

Community / User Applications
Community Agreed Exchange Standards
Business Rules

Figure 6 - IEPPV Focus

Second, the integration of the IEPPV[xviii] into the Unified Profile for DODAF and MODAF (UPDM version 3[ix]) will align ISS policy to interconnected architectural elements, e.g., organizations (e.g., producers, receivers, communities of interest), platforms, systems, services/applications, and interfaces. The IEPPV[xviii] will replace the SOPES Profile[x] elements in the UPDM version 2.x specifications and extent. This integration into a standardized architecture framework provides the opportunity for the development of practices and tools that enable certification, forensic auditing and retention of institutional memory.

Third, the integration of data/information protection rules assure a recipient is only receiving the information they are permitted to access. By adding the ability in overlay redaction filters, selective inclusion of data elements, and the auto-integration of privacy and security marking into the policy models (data patterns), the IEF enhances the definition and design of safeguarding.

Fourth, The IEF is targeting automation of ISS policy. Automation will occur in three (3) areas:

1. The serialization of policy models into a machine consumable and enforceable form;
2. The enforcement of policy through open standards based on ISS policy decision and enforcement points; and
3. The generation/reporting of data needed to:
   a. Implement defense in depth strategies to the data elements being shared;
   b. Identify and issue runtime alerts and warnings;
   c. Certify the ISS sharing and safeguarding patterns and services; and
   d. Conduct forensic audits for ISS activities.

This paper is specifically targeting the specification of the policy; serialization, automation and generation will be detailed in separate white papers.

Will the Approach cost more?

One could argue that adding formalism, standard practices and governance always increase the cost of developing and delivering information systems. However, delivering effective and trusted information sharing and safeguarding solutions that permit users to responsibly share information critical to decision makers will inherently require formalism (e.g., IEPD), standard practices, and governance (e.g., NIEM governance). The IEF is seeking to mitigate the costs of this formality by:

1. Developing open standards that simplify the development of ISS specifications and RFPs (*call up an open standard developed by Subject Matter Experts versus independently developing a complex set of requirements*);
2. Promoting competition during acquisition, development and maintenance (*promoting competition in products, tools, services and SMEs*);

3. Enhancing off-the-shelf capability (*promoting industry developed innovation*); and
4. Improving the retention of institutional knowledge and memory (*use architecture to document policy models and patterns and standard architecture frameworks and tools to capture and maintain this information*).

These elements will combine to control the lifecycle costs associated with current structured messaging solutions.

Are There Other Areas That Could Benefit From IEF?

The IEF approach to structured information is agnostic to the business or information domain. The IEPPV allows users to describe policy models that align policy to the rules governing the packaging (e.g., get, aggregate, transform, mark and filter); and processing (e.g., or parse, transform, marshal and put) of data and information elements. The policy models expose the procedures for publishing and receiving specified messages where structured data, including metadata, is used. Unstructured data elements are treated as opaque elements and treated as attachments. Business, information or domains that could benefit from the IEF approach to integration of sharing and safeguarding rules include:

- Healthcare;
- Justice / Corrections;
- Border Security, Customs and Immigration;
- Emergency & Incident Management; and
- Operational (shared) Situational Awareness:
  - Military,
  - National Security,
  - Public Security / Safety,
  - Crisis Management, and
  - Humanitarian Assistance.

The IEF Approach to Sharing Structured Data

The following sections outline the IEF approach to the development of policy models for the exchange of structured data. The approach is based on three (3) foundational principles.

1. **Separation of Concerns**: separate the development and maintenance of policy from the services and infrastructure used to implement the policy decision and enforcement points.
2. **Policy-Driven**: provide solutions that specifically align policy instruments (see above) to a specific messaging domain (e.g., NIEM[xi], EDXL[xii], CAP[xiii], UML[xiv], and MIP[xv]) and user data domain (e.g., JC3IEDM). Using UML to develop policy models affords the opportunity to integrate the models into standard architecture frameworks and provide full traceability between the policy instruments and the physical implementations. It also affords the opportunity to use model driven architecture to automate key elements of the process.
3. **Data-Centric**: enforces ISS policies (/rules) against the values of the data and metadata at runtime.
4. **Open Standards**: all elements of the IEF will be documented as a set of open standards in order to establish a robust support infrastructure of vendors, solutions, integrators, tools and technology. Where possible, reuse and integrate existing standards (e.g., UML, UPDM, XACML, and DDS).
5. **Operational flexibility, agility and adaptability**: provide the ability to adapt and extend the policy environment to accommodate changes in operational context.

From Model to Operations

Once a policy model is defined, the metadata underpinning the model is mined and serialized in a form that can be read, ingested and executed by a policy (/rules) Service. This policy service was first used to test the SOPES IEDM against the Multilateral Interoperability Programme's test suite. The current serialization supports an XML and proprietary binary form.

The serialized policies are sent to the appropriate nodes in the environment where they are ingested and enforced by the local policy based packaging service. A policy service may participate in any combination or number of Information Exchange Agreements.

The separation of the policies from the services that enforce them provides the flexibility, adaptability, and agility needed by many communities. It enables users:

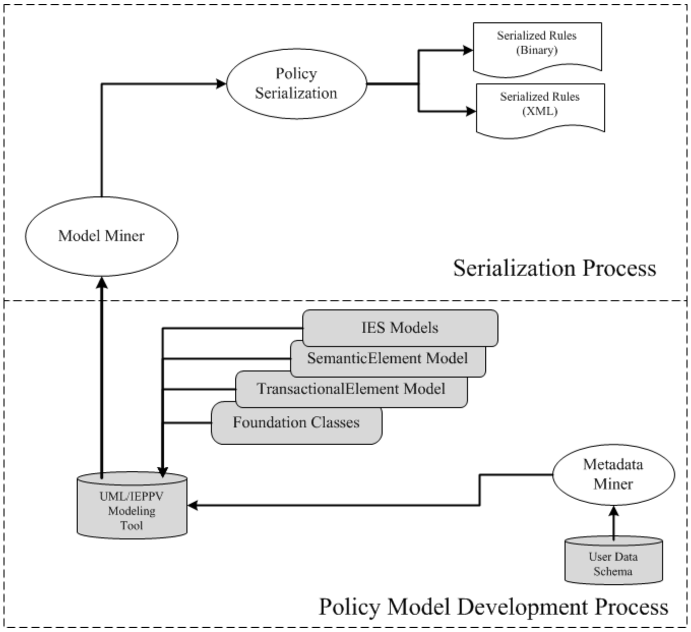


Figure 7 - Translating Policy into Executable Rules

- To deploy an infrastructure that can adapt to mission requirements when and where needed.
- To acquire and integrate capacity when and where needed.
- Adapt policies and reconfigure nodes during the course of an operation in response to changes in operational contexts (command intent, plans and orders, threat environment, scope and severity) without the deployment of new software.
- Rapidly develop, test and deploy executable policy (rule sets).
- Retain institutional knowledge.
- Model and simulate operational environments.

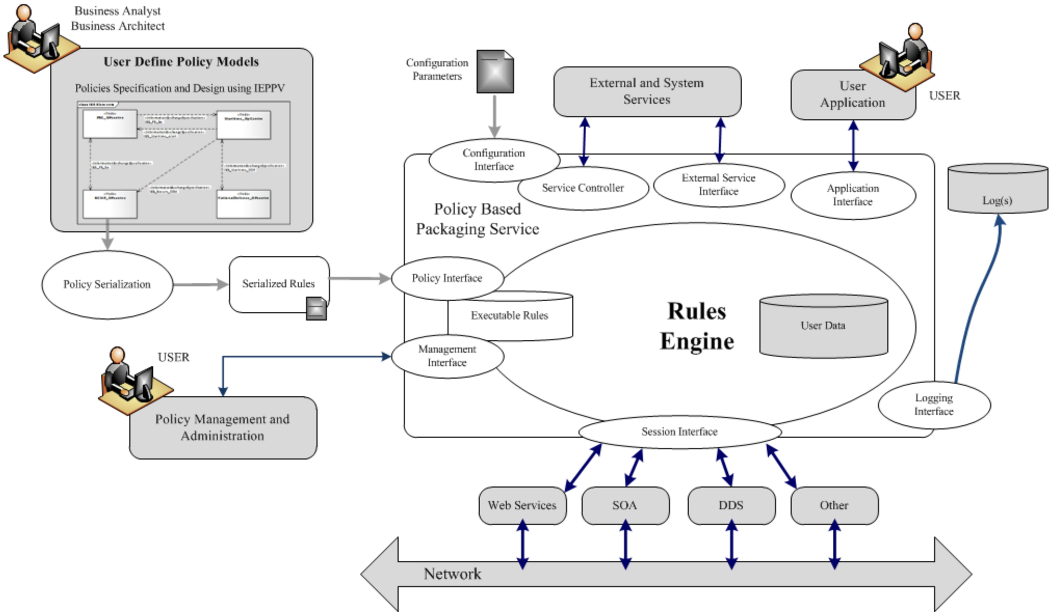


Figure 8- Policy Enforcement Point for Structured Data Packaging and Processing

For More Information

Mike Abramson, ASMG Ltd.
265 Carling Ave, Suite 630, Ottawa, Ontario, K1S2E1
Fax: 613-231-2556
Phone: 613-567-7097 x222
Cell: 613-797-8167
Email: abramson@asmg-ltd.com

Supporting data

Definitions

Adaptive Information Sharing

The ability to selectively share information content based on operational or business context (e.g., roles, relationship, risks, threats, severity, scale, and trust). This includes the ability of users (manually) or systems (automatically) to adjust active ISS policies to accommodate changes in business and operational context.

Asymmetric Information Sharing

The ability to share content with different communities, agencies or individuals conforming to legislative, regulatory, policy, contractual or service level requirements – while leveraging standard or shared protocols, interfaces and infrastructure.

Caveat

A warning or proviso of specific stipulations, conditions, or limitations to the sharing of data and information elements.

Community of Interest

A group of people interested in sharing information and knowledge in a particular topic or domain of discourse.

Data Centric

Enforce policies/rules against individual data assets; often referring to metadata or tags included within an information asset.

Decision Advantage

Enable commanders and/or decision makers, based upon information advantage and situational understanding, to make effective and informed decisions more rapidly than their adversary, thereby allowing one to dramatically increase the pace, coherence, and effectiveness of operations.

Information Advantage

Enable the provision of information needed to develop a degree of control in the information domain that permits the conduct of operations without effective opposition.

ISS Policy:

Principles, rules, and guidelines formulated or adopted by an organization to share and safeguard information holdings. They are designed to influence and determine all ISS decisions and actions, and all ISS actions take place within their boundaries.

Policy automation

The use of software services to automate the selection of a course of action (decision) and the execution of that selected course of action (execution).

Metadata

A form of structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource. Metadata is often called data about data or information about information.

Policy

A definite course or method of action selected from among alternatives and in light of given conditions to guide and determine present and future decisions (Webster Merriam Dictionary). ISS policy guides the determination of which data elements are releasable under a given set of conditions.

Policy automation

The use of software services to automate the selection of a course of action (decision) and the execution of that selected course of action (execution). For the purpose of this paper, this refers to the actions to be taken by IEF services (including decision and enforcement points) to share and safeguard information assets.

Policy-Driven

A process through which user defined policy instruments are translated into machine readable rules (/instructions) and enforced by software services and systems. This process results in full traceability from policy instrument to implementation (policy decisions and enforcement).

Policy Instrument

Formal document describing a plan of action by an individual agency or community to handle information sharing and safeguarding (e.g., legislation, regulation, memorandum of understanding and service level agreements).

Quality Information

Provision of high-quality information tailored to the needs of the decision makers':

a. **Accurate.** Information that exactly, precisely, and correctly presents availability, usability and deploy-ability of C4ISR capability, systems and services;
b. **Authoritative**. information that is recognized or accepted as being true or reliable;
c. **Relevant.** Information content tailored to specific needs of the decision maker;
d. **Timely.** Information provided when and where it is needed to support the decision making process;
e. **Usable.** Information is presented in a common functional format, easily understood by the decision makers and their supporting applications;
f. **Complete.** Information that provides all necessary and relevant data (where available) to facilitate a decision;
g. **Concise**: Information is provided in a form that is brief and succinct, yet including all important information;
h. **Trusted.** Information that is accepted as authoritative by stakeholders, decision makers and users; and
i. **Secure.** Information is protected from inadvertent or Malicious Release to unauthorized persons, systems or organizations.

Releasable Data

A dataset that conforms to the policy rights of the data receiver according to the relevant policy instruments.

Responsible Information Sharing

Compliant with legislation, regulation and policy; consistent with agency strategy, policy and direction; and accountable through governance and oversight:

- Maximize the volume, variety and quality of information that is discoverable and accessible by authorized users;
- Protect sensitive (classified, private, confidential and legally significant) information from unauthorized access/release and tampering;
- Protect information sources and processing methods;
- Protect civil rights/liberties; and
- Ensure that information is assured in its content, safe in transmission and use, and safeguarded from the threat of malicious acts, unauthorized use, clandestine exfiltration or compromise by remote intrusion.

Structured Data
A data set defined by fixed fields within a record or file. Relational databases and spreadsheets are examples of structured data.

Semantically Complete
Preserving the explicit meaning and intent of the information during packaging, processing and exchange.

Semi-Structured Data
A form of structured data.
A data set that is not fixed in location like traditional database records, but are structured, because the data are tagged and can be accurately identified (e.g., XML Document).

Sensitive Information
Information elements identified as classified, private, confidential or legally significant.

Acronyms

DODAF	Department of Defense Architecture Framework
IEF[xvi]	Information Exchange Framework
IEF RA[xvii]	Information Exchange Framework Reference Architecture
IEPPV[xviii]	Information Exchange Packaging Policy Vocabulary
IEPPS[xix]	Information Exchange Policy-based Packaging Service(s)
IEPMS[xx]	Information Exchange Policy Management Service(s)
ISE	Information Sharing Environment
ISS	Information Sharing and Safeguarding
JC3IEDM[xxi]	Joint Consultation, Command, Control and Intelligence Information Exchange Data Model
MODAF	Ministry of Defense Architecture Framework
NAF	NATO Architecture Framework
NATO	North Atlantic Treaty Organization
NIEM[xxii]	National Information Exchange Model
PM-ISE[xxiii]	Project Manager Information Sharing Environment
SOPES	Shared Operational Picture Exchange Services
SOPES IEDM[xxiv]	SOPES Information Exchange Data Model

UPDM[xxv]	Unified Profile for DODAF and MODAF

Endnotes

[i] Charles E. Phillips, Jr. et al, SACMAT '02 Proceedings of the seventh ACM symposium on Access control models and technologies, Pages 87-96, http://dl.acm.org/citation.cfm?doid=507711.507726

[ii] 9/11 Commission Report, http://www.google.ca/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CBwQFjAA&url=http%3A%2F%2Fwww.9-11commission.gov%2Freport%2F911Report.pdf&ei=nc_HU4WmGcmlyATxz4DIDg&usg=AFQjCNHGlh1FX-h5OMtRGzOky5zJfmQnbQ&bvm=bv.71198958,d.aWw&cad=rja

[iii] Critical Infrastructure Protection: Improving Information Sharing with Infrastructure Sectors GAO-04-780: Published: Jul 9, 2004. Publicly Released: Jul 27, 2004. http://www.gao.gov/assets/250/243318.pdf

[iv] http://www.healthit.gov/providers-professionals/health-information-exchange/what-hie

[v] NATIONAL STRATEGY FOR INFORMATION SHARING AND SAFEGUARDING, December 2012, http://www.google.ca/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CBwQFjAA&url=http%3A%2F%2Fwww.whitehouse.gov%2Fsites%2Fdefault%2Ffiles%2Fdocs%2F2012sharingstrategy_1.pdf&ei=CdLHU4XDPNSvyAT5oYLwCQ&usg=AFQjCNEkmIIqWdE9oYHP3dxA2K-8kkjw1A&bvm=bv.71198958,d.aWw&cad=rja

[vi] DHS, Information Sharing and Safeguarding Strategy, January 2013, https://www.google.ca/url?sa=t&rct=j&q=&esrc=s&source=web&cd=3&ved=0CCoQFjAC&url=https%3A%2F%2Fwww.dhs.gov%2Fsites%2Fdefault%2Ffiles%2Fpublications%2F12-4466-dhs-information-sharing-and-safeguarding-strategy-01-30-13--fina%2520%2520%2520.pdf&ei=CdLHU4XDPNSvyAT5oYLwCQ&usg=AFQjCNGjKn_VT-dQxiuu7JcjwTY-eTIqjQ&bvm=bv.71198958,d.aWw&cad=rja

[vii] Ministerial Directive in the Canadian Security Intelligence Service: Information Sharing with Foreign Entities; http://www.google.ca/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&uact=8&ved=0CCkQFjAB&url=http%3A%2F%2Fwww.cba.org%2FABC%2Fresolutions%2Fpdf%2F13-08-A-pdf.pdf&ei=C9THU6TuBIufyATn54HYAg&usg=AFQjCNEOZe0E1byizQbsb311dUopJBbMeg&bvm=bv.71198958,d.aWw

[viii] http://fas.org/sgp/news/2005/12/wh121605-memo.html

[ix] UPDM Version 3 Request for Proposal,

[x] SOPES Profile in the predecessor to IEPPV. The SOPES profile is currently integrating into UPDM version 2.1. Additional information on the SOPES Profile can be found in Annex A to the SOPES IEDM Version 1.0 specification, http://www.omg.org/spec/SOPES/

[xi] National Information Exchange Model, www.niem.gov

[xii] Emergency Data Exchange Language,

[xiii] Common Alerting Protocol,

[xiv] Unified Modeling Language,

[xv] Multilateral Interoperability Programme,

[xvi] http://www.asmg-ltd.com/sect_5a.html#Information_Exchange_Framework_%28IEF%29

[xvii] IEF RA RFP, http://www.omg.org/cgi-bin/doc.cgi?c4i/2013-9-11

[xviii] IEPPV, http://www.omg.org/spec/IEF-IEPPV/

[xix] IEPPS, see IEF RA, http://www.omg.org/cgi-bin/doc.cgi?mars/2014-3-17

[xx] IEPMS, see IEF RA, http://www.omg.org/cgi-bin/doc.cgi?mars/2014-3-17

[xxi] JC3IEDM, https://mipsite.lsec.dnd.ca/

[xxii] http://www.niem.gov

[xxiii] http://www.ise.gov

[xxiv] SOPES IEDM, http://www.omg.org/spec/SOPES/

[xxv] UPDM, http://www.omg.org/spec/UPDM/