PRIORITY OBJECTIVE 3

# DATA TAGGING FUNCTIONAL REQUIREMENTS

VERSION 1.0

DECEMBER 2014

UNCLASSIFIED

**A joint initiative conducted by the**
**Office of the Program Manager, Information Sharing Environment (PM-ISE)**
**and the Department of Homeland Security**

**Report Produced by the**
**Information Sharing and Access (ISA) Interagency Policy Committee (IPC)**
**Information Integration Subcommittee (IISC)**
**for the Information Sharing Environment (ISE)**

# CONTENTS

# LIST OF FIGURES



# LIST OF TABLES

# 1 INTRODUCTION

The 2012 National Strategy for Information Sharing and Safeguarding[1] (NSISS, "the strategy") identifies data tagging as a Priority Objective (PO) critical to the ability to both locate information and enable automated access control decisions. This document articulates the minimum functional requirements of data tagging standards needed to facilitate interoperable Query and Discovery, Access Control, Correlation, Audit, and Records Management capabilities across Federal networks and security domains.

Data "tags" are metadata—"data about data" applied to resources. A "tag" is an assertion describing some aspect of a resource, pairing a semantic label (or "tag name") with a corresponding tag value. For example, a document may be tagged with ***Language="English"***. The tag consists of both the name and the value, illustrated in Figure 1.

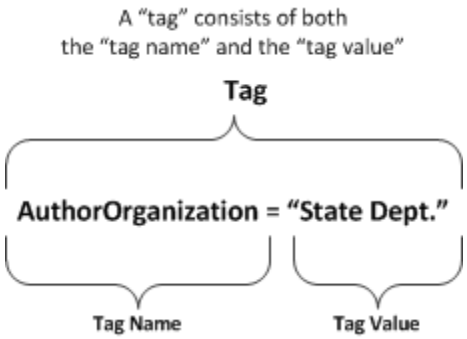



Figure 1. What Is a Tag?

The idea of metadata is not new—files have had rudimentary metadata (e.g., size, name, or date) since the early days of computer systems. Data tags extend this concept into a far richer set of metadata.

There are particularly important inter-dependencies between data tagging and other NSISS priority objectives—particularly PO 4 (FICAM on all fabrics) and PO 8 (Discovery and Access). For example, PO 3 (this PO) will define the tags assigned to resources, which may support, influence, or enable access control policies executed and enforced by PO 4 when performing the discovery capabilities defined in PO 8. Therefore agencies should examine this document in conjunction with the issuances of the POs 4 and 8 working groups and bodies.

A maturity model, provided in Appendix C, allows agencies to assess progress with respect to metadata using a common construct and scale.

[1] See http://www.whitehouse.gov/sites/default/files/docs/2012sharingstrategy_1.pdf

## 1.1 REFERENCES AND AUTHORITIES

See Appendix A.

## 1.2 RELATION TO OTHER DOCUMENTS

### 1.2.1 RELATION TO THE NSISS AND SIP

This document, the *Priority Objective 3 Data Tagging Functional Requirements Document*, is called for by the *2014 Strategic Implementation Plan* (SIP) for the NSISS. This document lays out a common set of requirements that implementations of data tagging specifications must fulfill in order to achieve the objective outlined in the NSISS.

### 1.2.2 RELATION TO OTHER PRIORITY OBJECTIVES

There is extensive interaction between Priority Objectives 3, 4, 8, and 10. For example: Priority Objective 3 data tags will enable the access control capabilities outlined in Priority Objective 4, which includes an entire series of activities around implementing data tags (that is, PO 4 depends on PO 3). Additionally, data tags will facilitate the discovery process capabilities described in Priority Objective 8. Data tags are a crucial aspect of data aggregation, described by the Data Aggregation Reference Architecture developed under PO 10. The reader is encouraged to review this document in concert with the implementation plans for these other Priority Objectives.

### 1.2.3 RELATION TO AGENCY SPECIFICATIONS

A significant number of data tagging specifications currently exist in the Federal Government, such as the Intelligence Community's Information Resource Metadata (IRM) and Information Security Marking (ISM) standards. While these specifications (listed in Appendix A) were consulted and reviewed while generating this document, these various specifications are individual instantiations of the requirements and structures set forth here; *this document does not replace those various standards, but provides a way to enable interoperability between them*. This interoperability is achieved by mapping, as described later in this document.

### 1.2.4 RELATION TO A GOVERNMENT-WIDE SPECIFICATION

A forthcoming effort will develop a *PO 3 Government-wide Data Tagging Specification* that Departments and Agencies **may** adopt if they choose, rather than developing their own. This forthcoming specification will be in alignment with and meet the requirements set forth in this Government-wide requirements document.

# 2 SCOPE

This document provides a framework for interoperable metadata tagging standards, oriented around abstract metadata concepts (vs. concrete specifications). Departments' and Agencies' internal specifications may implement these concepts in many different ways. Those specifications are not in scope for this document.

This document relates to **metadata**, *not data*. It does not attempt to address or define the data elements within a dataset or message payload, which may be defined in a data standard or specification such as the National Information Exchange Model (NIEM). This document only addresses the metadata tags that describe the data.

The metadata concepts within this document are intended to be applicable at any appropriate level of granularity: at the dataset, document, or even data element level if supported by the data specification used in a structured payload. For example, the NIEM specification defines a method for indicating the information security markings in the Safeguarding and Sharing tag area of this document.

This document does not attempt to dictate any sort of internal data tagging framework, terminology, lexicon, or ontology to be used purely within an agency network or system, but requires that those internal constructs be able to be mapped and translated to the constructs set forth in this document when used in interagency exchanges (in the "white space" between agencies).

This document applies to all Executive Branch Departments and Agencies who operate information technology systems on any security classification domain/fabric. This document may be useful to State, Local, Tribal, and Private Sector organizations as well.

# 3 OPERATIONAL CONCEPT

This data tagging framework in Figure 2 is organized around the concepts of Capabilities, a tiered Tagging construct, and interoperability of tagging specifications.

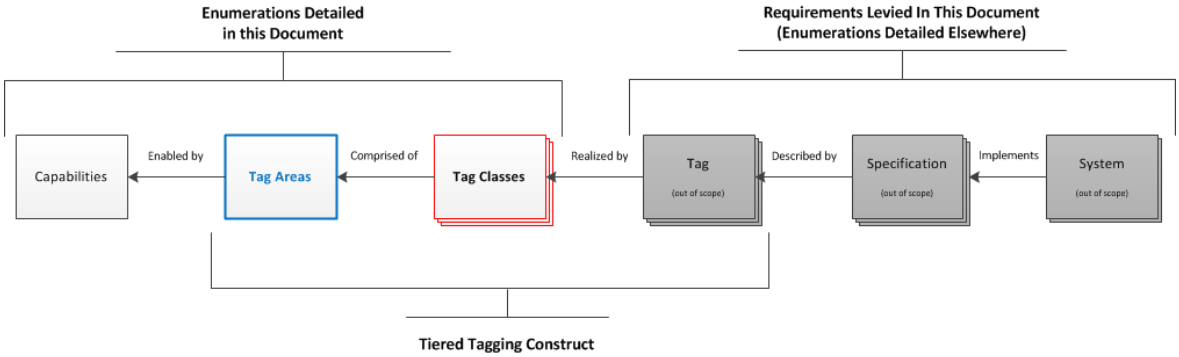



Figure 2. PO 3 Data Tagging Framework

## 3.1 CAPABILITIES

While data tags can be used for any number of purposes and can support any number of capabilities, this framework is oriented around five capabilities common across nearly all Departments and Agencies that are essential to information sharing and safeguarding:

- **Query and Discovery:** the ability to locate and obtain knowledge of the existence of, but not necessarily the contents of, a resource.[2]
- **Access Control:** granting or denying specific requests for resources based on a defined set of criteria.[3]
- **Correlation:** identifying relationships between entities within and across disparate data sets.[4]
- **Audit:** recording the sequence of actions surrounding or leading up to a specific activity or event.[5]
- **Records Management:** managerial activities involved with the creation, retention, and disposition of records.[6]

[2] Definition based on Intelligence Community Directive (ICD) 501.
[3] Definition based on Federal Information Processing Standard (FIPS) 201.
[4] Definition based on the Data Aggregation Reference Architecture (DARA), NSISS Priority Objective 10.
[5] Definition based on Committee for National Security Systems Issuance (CNSSI) 4009.
[6] Definition based on 44 USC 2901.

## 3.2 TIERED TAGGING CONSTRUCT

This data tagging framework takes a three-tier hierarchical approach to data tags:

1) **Tag Area:** an abstract, purely administrative grouping of tags that support a common Capability.
2) **Tag Class:** a logical, well-defined concept, the meaning of which is consistent across organizations (that is, it is "portable") but is still abstract.
3) **Tag:** the concrete syntactic and semantic means and encodings defined by an organization to realize the concept described by a Tag Class. Tags, and the specifications that formally describe them, are outside of the scope of this document, but will be addressed by the planned *PO 3 Government-wide Data Tagging Specification*.

| Tag* | Tag Class | Tag Areas | |
|---|---|---|---|
| C | A | A | **Abstract / Concrete** |
| | ✓ | | **Portable:** has the same connotation and denotation across organizations. |
| ✓ | | | **Translatable:** can be mapped to others by some sort of rule |

Figure 3. Characteristics of Tag Tiers

For the purposes of this framework, only the Tag Area and Tag Class tiers are in scope. The individual Tags will be covered by the *PO 3 Government-wide Data Tagging Specification*, existing cross-agency data tagging specifications, and the various department and agency specifications.

## 3.3 TAG AREAS

This framework has developed four Tag Areas:

- **Resource Description:** Tag Classes that contribute to a requestor being able to locate a resource, akin to a card in a library card catalog.
- **Reference:** Tag Classes that contribute to linking a resource with other related resources; akin to a bibliography.
- **Lifecycle:** Tag Classes that contribute to a resource moving through an organization's process, such as its maturity, review and approvals, and retention information.
- **Safeguarding and Sharing:** Tag Classes that contribute to understanding who may access (either for discovery or retrieval purposes) a resource, how it may be used, and how to properly protect the resource.

Each Tag Area supports one or more Capabilities, as shown in Figure 4 below.

| | | Tag Areas | | | |
|---|---|---|---|---|---|
| | | Resource Description | Lifecycle | Reference | Safeguarding & Sharing |
| Capabilities | **Query & Discovery:** obtaining knowledge of the existence of, but not necessarily the contents of, a resource. (ICD 501) | ✓ | | ✓ | |
| | **Access Control:** granting or denying specific requests for resources based on a defined set of criteria (FIPS 201) | | | | ✓ |
| | **Correlation:** identifying relationships between entities within and across disparate data sets (DARA) | ✓ | ✓ | | |
| | **Audit:** recording the sequence of actions surrounding or leading up to a specific activity, or event (CNSSI 4009) | | ✓ | | |
| | **Records Mgmt.:** managerial activities involved with the creation, retention, and disposition of records (44 USC 2901) | | ✓ | | |

Figure 4. Capabilities

The flowchart diagram in Figure 5 shows how the various Capabilities and Tag Areas fit together in a user- and data-oriented approach, rather than an architectural approach. The following narrative describes the flow:

- A query is processed using the tags in the Resource Description Tag Area to return a set of relevant resource items – 25 items in the example in Figure 5.
- Tags in the Safeguarding and Sharing Tag Area are used to identify which of the 25 relevant items are discoverable – 15 in the example.
- Of the 15 discoverable items, tags in the Safeguarding and Sharing Tag Area determine which are accessible, 10 items in the example, and the items are retrieved.
- The remaining items are discoverable but not retrievable, and the requester can be provided with point of contact information at which to further inquire about the items.
- Retrieved items may cite other items using tags in the Reference Tag Area, which can result in a new query.
- The Audit capability supports the entire process and immutable audit logs are created throughout the sequence of events.
- Throughout the entire process, the Records Management capability uses the tags in the Lifecycle Tag Area to manage the scheduling, retention, and disposition of Federal records.

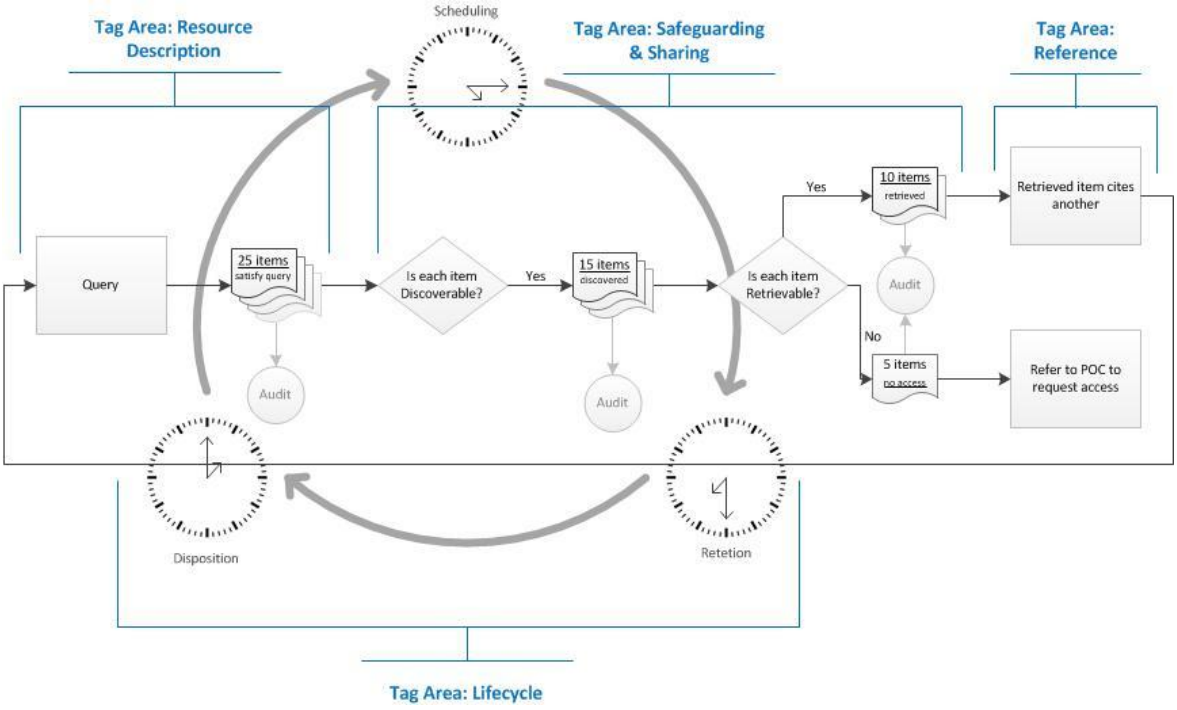



Figure 5. Tag Enablement Example

## 3.4 TAG CLASSES

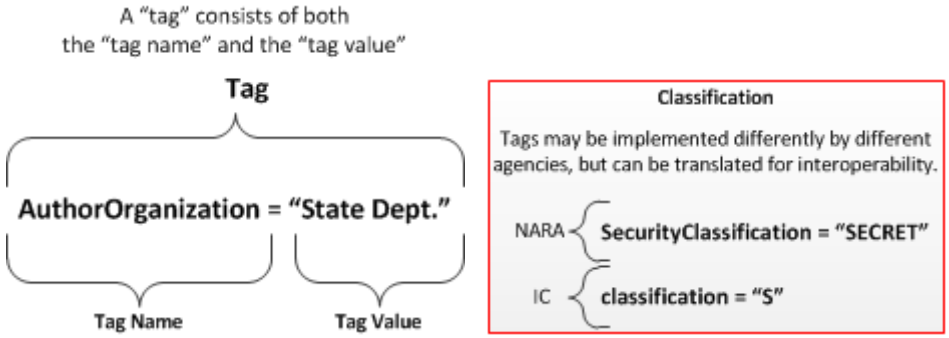
Tag Classes are the logical, abstract concepts defined in and required by this framework. Where this framework is adopted, these concepts are consistent across Departments and Agencies. The collection of tag classes included in this framework was drawn from existing agency specifications such as the IC's ISM & IRM and DoD's DDMS by "rolling up" the specific implementations to their higher level, abstract concepts. Section 5 further explains and enumerates the Tag Classes.

## 3.5 TAGS

Tags are the first and only concrete layer in the tiered tagging construct, and consist of a name+value pair that together convey some information about the resource with which the tag is associated. One or more tags may combine to provide the information required by a Tag Class. For example, if the Tag Class is "Author," one agency may use a single tag to convey both the author's organization as well as the specific author, such as "Author=FBI/Agent Smith". Another agency may use two tags: "AuthorAgency=FBI" and "Author=Agent Smith".

Regardless of which tags an agency implements, a specification defines and formalizes those selections. The specification provides an explicit name for each tag, the allowed values that each tag can be assigned, and the meanings of those values.

Both the single-tag and multiple-tag authorship models described above are acceptable under this decentralized-yet-compatible framework. As agencies define their tagging specifications (or adopt an existing tagging specification), the tags map back to the Tag Class that they support, enabling construction of machine-readable rules to perform automated translation. In the example given above, the two models can be easily translated by either splitting the single tag into two or combining the two tags into one, depending on the direction required.




Portable Tag Classes enable interoperability

Figure 6. Tag Concept, Tag Portability

# 4 REQUIREMENTS AND ACTIVITIES

As data tagging is "the process or act(s) of associating a data object with characterizing metadata for some purpose",[7] it has both an organizational (people, governance, and process) aspect and a functional (specification, implementation) aspect. This document levies requirements on both of these aspects, levying organizational requirements on the Agency and functional requirements on the Tag, Specification, and System.

Table 1. Functional Requirements

| | | |
|---|---|---|
| **AGENCY-ASSIGNED ACTIVITIES** | | |
| **AA1** | | Within six (6) months of the approval of the Priority Objective 3 Government-wide Tagging Specification ("the PO 3 specification"), Agencies shall select and publish a data tagging specification compatible with the PO 3 specification either by 1) adopting the PO 3 specification as-is, 2) adopting the PO 3 specification with modifications or extensions, 3) adopting another Agency's specification that is itself compatible with PO 3, or 4) developing their own specification. |
| | **AA1.1** | If adopting a modified or extended PO 3 specification, or if developing their own specification, Agencies shall define and publish translation rules required for automated interoperability between their specification and the PO 3 specification. |
| | **AA1.2** | Within two (2) years of the approval of the PO 3 specification, Agencies shall ensure that resources newly created are tagged in accordance with their selected data tagging specification. |
| | **AA1.3** | Within two (2) years of the approval of the PO 3 specification, Agencies shall ensure that existing resources are tagged in accordance with their selected data tagging specification when those resources are migrated to a new system, updated, or otherwise altered. |
| | **AA1.4** | Within two (2) years of the approval of the PO 3 specification, Agencies shall ensure that any resources being shared with an external organization are tagged in accordance with the PO 3 specification when they leave the Agency, regardless of the selected data tagging specification used within the Agency. |
| **TAG REQUIREMENTS** | | |
| **T1** | | Tags shall have unique names within the organization. |
| **T2** | | Tags shall be traceable back to zero[8] or one Tag Class and documented accordingly. |
| **T3** | | Tags that are traceable back to a Tag Class shall have semantic meanings consistent with the Tag Class that they realize. |
| **T4** | | Tags that are traceable back to a Tag Class shall have tag values whose meanings are consistent with the meanings in the PO 3 specification. |
| **T5** | | Tags shall have a defined syntax for its possible values (e.g., CVE, regex, etc.). |
| **SPECIFICATION REQUIREMENTS** | | |
| **Sp1** | | An agency's selected data tagging specification shall cover all tags used within the organization. |
| **Sp2** | | An agency's selected data tagging specification shall be change controlled. |
| **Sp3** | | An agency's selected data tagging specification shall be discoverable within the organization. |
| **Sp4** | | An agency's selected data tagging specification shall include machine readable translation rules between the indigenous specification and the PO 3 specification |

[7] Priority Objective 3 Implementation Plan, citing the definition agreed to by the IISC.

[8] Tags that trace to zero Tag Classes are considered non-interoperable extensions and should be minimized to the extent possible.

| SYSTEM REQUIREMENTS | |
|---|---|
| **Sy1** | Systems implementing data tagging shall bind or otherwise reliably associate tags and the resources that they describe. |
| **Sy2** | Systems implementing data tagging shall ensure enforcement of the specification that they implement. |
| **Sy3** | Systems implementing data tagging shall assign tags at an appropriate level (cell, record, collection of records, etc.) |
| **Sy4** | Systems implementing data tagging shall allow query by, and refinement of query by, data tags. |
| **Sy5** | Systems implementing data tagging shall apply tags to structured, semi-structured, and un-structured resources. |
| **Sy6** | Systems implementing data tagging shall protect tags on resources against tampering or unauthorized modification. |

# 5 TAG CLASSES

Tag Classes are administratively grouped into Tag Areas purely for the sake of convenience and comprehension and will be presented below organized in the same manner. A Tag Class's membership in one Tag Area vs. another has no impact on the underlying implementation by a Department or Agency.

In each of the diagrams below, the Tag Area is represented by a blue box, and a Tag Class by a red box. Individual Tags are outside of the scope of this document and are not shown in the diagrams.

Each of the Tag Classes identified in these sub-sections is described in detail in the table in Appendix B.

## 5.1 RESOURCE DESCRIPTION TAG CLASSES

The Resource Description Tag Area includes Tag Classes that are most similar to a card in a library's card catalog, such as authorship, subject, title, etc. These Tag Classes help a requestor find the item(s) being sought. The Tag Classes in Resource Description, shown in Figure 7, are based primarily on the Dublin Core, which is the foundation for many other data tagging efforts, including within the IC and NARA.

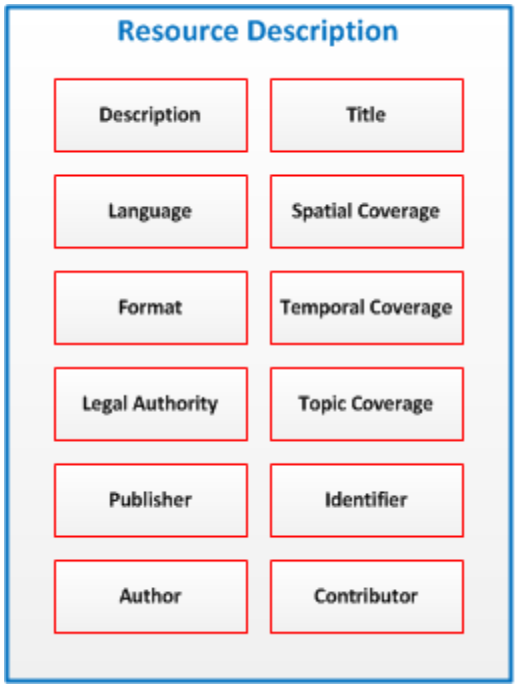



Figure 7. Resource Description Tag Classes

## 5.2 REFERENCE TAG CLASSES

The Reference Tag Area includes Tag Classes that allow a requestor to identify other resources that have a direct linkage to the current resource. For example, an intelligence product may provide a reference to the various reports on which it was based (Citation), or a multi-part video may contain a reference to the next and previous parts in the series. The Tag Classes in Reference, shown in Figure 8, are based primarily on the Dublin Core.


Figure 8. Reference Tag Classes

## 5.3 LIFECYCLE TAG CLASSES

Unlike the Resource Description and Reference Tag Classes, which are focused on enabling actions taken by a requestor, the Lifecycle Tag Classes, as shown in Figure 9, focus on enabling actions taken by an organization. The Tag Classes included in the Lifecycle Tag Area enable an organization to track a resource during its movements through the organization, such as through a data lifecycle, a review-and-approval process, and Federal Records Act activities such as retention and disposition.

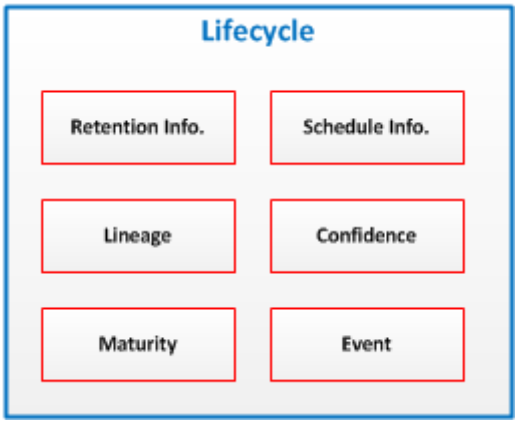



Figure 9. Lifecycle Tag Classes

## 5.4 SAFEGUARDING AND SHARING TAG CLASSES

The Safeguarding and Sharing Tag Classes, as shown in Figure 10, like the Lifecycle Tag Classes, are focused more on enabling an organization's actions than a user's. This Tag Area includes the Tag Classes needed to protect resources from unauthorized access (such as with classified information) or use (such as with licensed information).

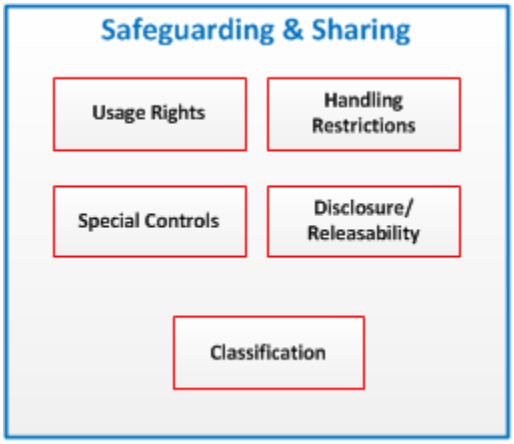



Figure 10. Safeguarding and Sharing Tag Classes

# 6 USE CASES/FUNCTIONAL SCENARIOS

The following functional scenarios are intended to describe which data tags are needed in cross-agency information sharing to support each capability. The functional scenarios are notional, and may be used to develop functional tests for capabilities using the various tags, but are not intended to detail comprehensive eventual functional testing that would be developed in a test plan.

## 6.1 FUNCTIONAL SCENARIO 1 – ACCESS

**Capability Demonstrated:** Access

**Tag Area/Class:** Safeguarding and Sharing/Handling Restrictions

**Narrative:** A DHS law enforcement officer investigating a case involving a threat to protected critical infrastructure requires access to a certain document ("resource") that is tagged "PCII." A tag implementing the Handling Restriction Tag Class applied to the resource allows the access control system to determine which access control policy should be applied, and determine the conditions for access to that particular resource.

**Outcome:** The system evaluates the PCII access control policy and grants or denies access to the resource.

**Function of Tag:** The tag allowed the access control system to identify the resource as being PCII, which would not have been easily discernable from the contents of the resource itself.

## 6.2 FUNCTIONAL SCENARIO 2 – CORRELATION

**Capability Demonstrated:** Correlation

**Tag Area/Class:** Lifecycle/Lineage

**Narrative:** An FBI Special Agent is investigating a foreign national who has applied for a U.S. Visa, reviewing potentially derogatory information. The foreign national's name is present in multiple data holdings, some of which are copies of each other. A tag implementing the Lineage Tag Class identifies these copies as coming from the same underlying data holding. This prevents correlating data with a copy of itself, regardless of in which system it is copied and stored.

**Outcome:** The system links information without self-reinforcing feedback loops (a form of circular reporting).

**Function of Tag:** The tag allowed the system to trace the origin of the information, which is not easily discernable from the contents of the resource itself.

## 6.3 FUNCTIONAL SCENARIO 3 – DISCOVERY

**Capability Demonstrated:** Discovery

**Tag Area/Class:** Resource Description/Coverage

**Narrative:** In September 2012, a FBI Special Agent uploads a document to the LEEP/LEO system. During the upload process he tags the document with a temporal coverage of "March 2012". Later, an intelligence analyst narrows her search by specifying that only documents covering Q1 2012. Even though the document was uploaded in September, the tag implementing the Temporal Coverage Tag Class identifies the document as covering March 2012 and the document is included in the results.

**Outcome:** The system returns documents that cover only the specified time, regardless of when they were uploaded.

**Function of Tag:** The tag allowed the system to differentiate between the date that the information was uploaded and the date that the information was about. Additionally, a document may include several dates ("Subject born in April 1980 committed a robbery in August 2004"); the tag allows the contributor to identify which date(s) is relevant.

## 6.4 FUNCTIONAL SCENARIO 4 – RECORDS MANAGEMENT

**Capability Demonstrated:** Records Management

**Tag Area:** Lifecycle/Retention Information

**Narrative:** The Department of Homeland Security (DHS) maintains a computer system subject to the Federal Records Act. Based on a NARA-approved records control schedule, each document is tagged with its permanent/temporary/non-record status and its approved disposition date.

**Outcome:** The system automatically destroys temporary and non-record documents in accordance with the approved records control schedule, and automatically transfers permanent records to NARA. When these DHS documents are shared with another agency, the tag on the document notifies the receiving agency of the approved disposition dates.

**Function of Tag:** The tag allowed the system to recognize which documents are subject to which approved retention policies, and to automatically enforce those policies.

## 6.5 FUNCTIONAL SCENARIO 5 – AUDIT

**Capability Demonstrated:** Audit

**Tag Area:** Lifecycle/Lineage, Lifecycle/Audit

**Narrative:** An Intelligence Community (IC) agency receives information via some mechanism. This information flows through a number of filtering, analysis, and exploitation steps before resulting in a finished product, which itself flows through a number of review steps. At each stage, the information is tagged to reflect that it was handled/touched by a certain process, system, or individual.

Later it has been determined that this information is false and must be retracted from all places where it was disseminated. Using the tags associated with the resource, all recipients are notified.

**Outcome:** A transparent record of action is generated and associated with the document.

**Function of Tag:** The tag allows the system to record the flow of a resource throughout the enterprise, without having to alter the resource itself.

# APPENDICES

This page intentionally blank.

# A. REFERENCES AND AUTHORITIES

A. The White House. (December 2012). National Strategy for Information Sharing and Safeguarding.

B. Comprehensive National Cybersecurity Initiative 5 (CNCI-5) Enhance Shared Situational Awareness (ESSA). (17 July 2013). CNCI-5 Information Sharing Architecture (ISA) Shared Situational Awareness (SSA) Requirements Document v2.0.

C. Comprehensive National Cybersecurity Initiative 5 (CNCI-5) Enhance Shared Situational Awareness (ESSA). (10 February 2014). Information Sharing Architecture (ISA) Access Control Specification v1.1.

D. Data.gov Agency POC Working Group. (January 2010). Recommendations for Metadata Within Data.Gov

E. Department of Defense. (21 December 2012). Department of Defense Discovery Metadata Specification (DDMS) v5.0.

F. International Organization for Standardization (ISO). (14 June 2012). Dublin Core Metadata Element Set (DC MES), Version 1.1. (ISO 15836).

G. Office of the Director of National Intelligence. (14 January 2013). Intelligence Community Technical Specification XML Data Encoding Specification for Access Rights and Handling Version 2.

H. Office of the Director of National Intelligence. (14 January 2013). Intelligence Community Technical Specification XML Data Encoding Specification for Enterprise Data Header Version 2.

I. Office of the Director of National Intelligence. (14 January 2013). Intelligence Community Technical Specification XML Data Encoding Specification for Information Resource Metadata Version 9.

J. Office of the Director of National Intelligence. (14 January 2013). Intelligence Community Technical Specification XML Data Encoding Specification for Information Security Markings Version 1.

K. Office of the Director of National Intelligence. (14 January 2013). Intelligence Community Technical Specification XML Data Encoding Specification for Need-To-Know Metadata Version 8.

L. Office of the Director of National Intelligence. (27 February 2012). Intelligence Community Technical Specification XML Data Encoding Specification for Intelligence Publications Version 9.

This page intentionally blank.

# B. TAG CLASS DEFINITIONS

| TAG CLASS | DESCRIPTION | ABSTRACT CONCEPT IMPLEMENTED | EXAMPLE TAG |
|---|---|---|---|
| **Author** | An entity primarily responsible for making the resource. | Creator | NARA: Creator<br>IC: AuthorInfo |
| **Contributor** | An entity responsible for making contributions to the resource. | Contributor | |
| **Description** | A brief account of the resource. | Description | NARA: Description<br>IC: Description |
| **Format** | The encoding or data type of resource, providing information on how to interpret, open, or view the contents. | Format | |
| **Identifier** | An unambiguous (unique) reference to the resource | Identifier | NARA: RecordID |
| **Language** | The specific language in which the resource is written | Language | IC: Language |
| **Legal Authority** | The particular documented legal basis for mission activities associated with the creation, retention and use of a resource. | | |
| **Publisher** | The entity responsible for making a resource available ("releasing the resource"). | Publisher | IC: Publisher |
| **Spatial Coverage** | The geographic region(s) about which the resource provides information. | Coverage | NARA: SpatialCoverage<br>IC: Region |
| **Temporal Coverage** | The time period(s) about which the resource provides information. This is separate from the date that the resource was created or published. | Coverage | NARA: TemporalCoverage<br>IC: Temporal |
| **Title** | A name given to the resource | Title | NARA: Title<br>IC: Title |
| **Topic Coverage** | The subject(s) (in the thematic / issue sense of the word, not the person sense) about which the resource provides information | Subject | |
| **Citation** | A bibliographic reference | | IC: BibliographyEntry |
| **Related Resource** | A link to another resource that contains complementary, contradictory, clarifying, other otherwise related information. | Relation | IC: Relation |
| **Confidence** | A description of the level of belief in the accuracy of the information within the resource | | |
| **Event** | Information pertaining to an event within the resource's lifecycle (e.g. authored, published, approved, rescinded, viewed, forwarded, etc.) | | |
| **Lineage** | Information pertaining to where a resource originated and where it has travelled or been routed. | | |
| **Maturity** | Information pertaining to the resource's point within a lifecycle | | |
| **Retention Info.** | Information pertaining to the resource's authorized retention and disposition under the Federal Records Act | | |
| **Schedule Info.** | Information pertaining to the resource's assignment to and categorization under an authorized Records Schedule | | |

| TAG CLASS | DESCRIPTION | ABSTRACT CONCEPT IMPLEMENTED | EXAMPLE TAG |
|---|---|---|---|
| **Classification** | A single indicator identifying the highest level of classification contained within a resource | | NARA: SecurityClassification |
| **Disclosure/ Releasability** | Information pertaining to countries, organizations, or communities approved to receive the resource. | | |
| **Handling Restrictions** | Limitations not related to classification or releasability, such as Controlled Unclassified Information designations. | | |
| **Special Controls** | Indicator(s) identifying the sensitive compartmented information, special access program/special access required, or related that are contained within a resource. | | |
| **Usage Rights** | Restrictions on commercial, intellectual, or proprietary information, such as copyrights. | | |

# C. DATA TAGGING MATURITY MODEL AND CONCEPT DIAGRAM

| | LEVEL 1 – AD HOC | LEVEL 2 – REPEATABLE | LEVEL 3 – ENHANCED | LEVEL 4 – MANAGED | LEVEL 5 – OPTIMIZED |
|---|---|---|---|---|---|
| **People** | • Limited understanding of metadata | • Awareness of the importance and role of metadata to interoperability and information sharing<br>• Assigned roles supporting metadata lifecycle | • Motivated to apply and update metadata on resources | • Trained on metadata best practices and the organization's metadata policies, procedures, and standards | • Manage metadata as part of normal business process<br>• Provided opportunities to give optimizing input and feedback on metadata use and management |
| **Governance** | • No formal metadata governance process<br>• No organizational metadata policies or procedures<br>• Rudimentary, often informal, agreements between individual local users of metadata | • Organizational metadata policies and procedures are developed<br>• Individual point-to-point agreements are formalized and standardized | • Governance bodies formed to manage metadata and interoperability across the organization<br>• Point-to-point agreements are migrated to enterprise-wide model | • Metadata lifecycle operations and use are evaluated against approved policies and standards<br>• A culture of metadata interoperability is promulgated throughout | • Optimization decisions are made on a regular basis<br>• Policies and procedures developed jointly with other external organizations |
| **Process** | • Little to no process documentation<br>• Processes are unpredictable, poorly controlled, and reactive | • Best practices are identified and made available<br>• Change control process for metadata specifications established but not consistently followed<br>• Processes are predominately reactive | • Metrics for evaluating the performance of metadata use are established<br>• Change control process for metadata specifications enforced and adhered to<br>• Processes are standardized and proactive | • Performance is determined based on established metrics<br>• Processes are consistent with established policies and procedures<br>• Processes are controlled and measured | • Processes are in place to evaluate new approaches to optimizing metadata for advancing interoperability and information sharing across the organizations and to other, external organizations |
| **Specification** | • No "complete picture" of all specifications in use<br>• Systems adhere to multiple specifications, many of which are undocumented<br>• Each system has local metadata definitions, syntax, semantics, and encodings | • Inventory of existing metadata specifications completed<br>• Existing metadata specifications are documented<br>• Commonalities between specifications are identified | • A baseline set of metadata definitions is identified, approved for use, maintained, and stored in a repository<br>• Consistent metadata definitions, syntax, and semantics are established for widely-used metadata concepts (Tag Classes) | • Organization standardizes on a single data tagging specification describing syntax, semantics, and encodings<br>• Specification is used consistently across the organization<br>• Specification is easily accessible via a repository | • Specification is interoperable across organizations, in accordance with the *PO 3 Government-wide Tagging Specification* |
| **System** (Implementation) | • Few systems implement metadata<br>• No enforcement of specification compliance due to free-hand tagging | • Few systems implement metadata, but those that do enforce compliance with their specification | • All relevant systems implement and enforce approved specification, but tagging remains manual | • Resources are semi-automatically tagged at the lowest appropriate level<br>• Automated tools for metadata management | • Resources tagged automatically at the lowest appropriate level |
| **Use** | • Almost no automated sharing or decision-making based on metadata | • Semi-automated sharing and decision-making based on metadata, with human review/verification prior to execution | • Some automated sharing and decision-making within the organization based on metadata. | • Extensive automated sharing and decision-making within the organization based on metadata | • Automated sharing and decision-making across organizations based on metadata |

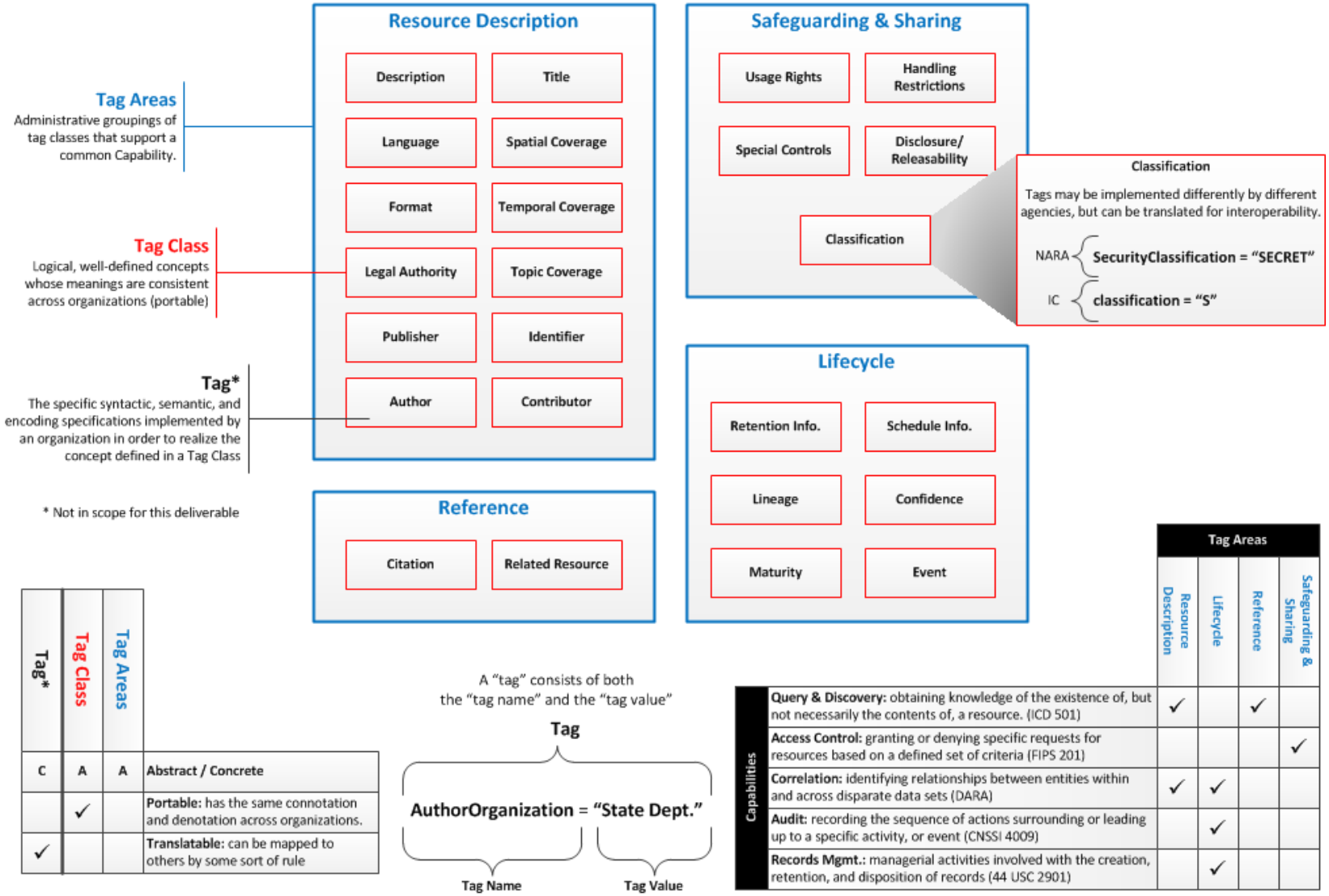



Figure C-1. Priority Object 3: Data Tagging – Functional Concept