GEIA Standard

Data Management

GEIA- 859

Working Draft

This Standard is Being Developed under EIA Project PN 4888

DISTRIBUTION NOTICE:

Distribution is limited to participants, reviewers and commenters during the development of this standard.

Comments can be submitted to:

jforbes@lmi.org, Logistics Management Institute
(703) 917-7572

HauerCC@aol.com, Millennium Data Management, Inc
(256) 536-1096

List of Figures

List of Tables

Foreword

The identification, definition, preparation, control, archiving, and disposition of data all require a sizable investment in labor, supporting systems, and time. The purpose behind enacting consistent, high-quality data management (DM) is to make certain that the enterprise reaps a return on this investment. DM applies effective processes and tools to acquire and provide stewardship for data. A well-designed DM process ensures that customers receive the data they need when they need it, in the form they need, and of requisite quality.

When DM principles are applied using effective practices, return on investment in data is maximized and product life cycle costs are reduced. This standard is intended to be used when establishing, performing, or evaluating DM processes in any industry, business enterprise, or governmental enterprise.

This standard describes DM principles and methods using a neutral DM terminology. Sections 1 through 9 are normative. Annexes are informative.

The methods of DM have undergone significant changes as paper documents transitioned to digital data, and continue to evolve. As a result, existing formalized policy, manuals, and instructions for DM, which mostly address DM for defense products, are obsolete. They describe procedures that were adapted to efficient paper-based management of paper deliverables. This standard is intended to articulate contemporary DM principles and methods that are broadly applicable to management of electronic and non-electronic data in both the commercial and government sectors. Development of this standard began in August 2000 when the Electronic Industries Alliance’s (EIA) G-33 Committee on Data and Configuration Management initiated task PN 4888 to develop a consensus standard for data management. This is the first release of the standard. Contributors to this standard are identified in Annex A.

Introduction

SCOPE

Data is information (e.g., concepts, thoughts, opinions) that has been translated into a form that is convenient to move or process. Data can be tables of values of various types (numbers, characters, and so on). Data also more complex forms such as engineering drawings and other documents, pictures, maps, sound, and animation.

For the purposes of this standard, there are three broad classes of data with which commercial and government enterprises concern themselves. The three types are as shown in Table Intro-1.

Table Intro-1 Types of Data

Type ***Usage***	**Examples**
Product *Collaboration*	Cost, schedule, and performance data. Engineering drawings for aircraft, ships, vehicles, spacecraft; parts catalogues; software applications, and their components; operational and maintenance instructions, training materials
Business *Collaboration*	Plans and programs, financial information, inventory status, and human resource information
Operational *Transactional Records Exchange*	Orders, issues, receipts, bills of lading, and invoices

Data management, from the perspective of this standard, consists of the disciplined processes and systems that plan for, acquire, and provide stewardship for product and product-related business data, consistent with requirements, throughout the product and data life cycles. Thus this standard primarily addresses product data and the business data intrinsic to collaboration during product acquisition and sustainment. It is recognized, however, that the principles articulated in this standard also have broader application to business data and operational data generally.

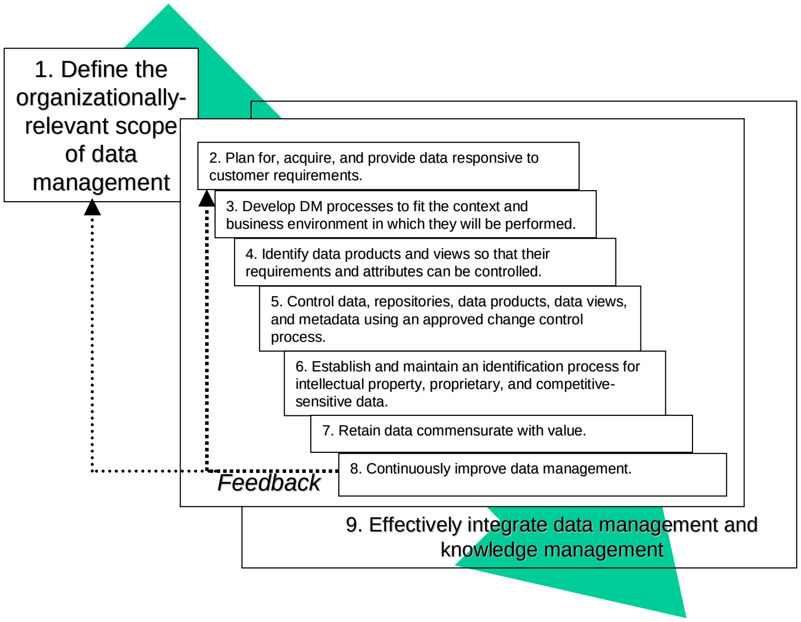
Data has many purposes including stating requirements, providing proof of achievement, establishing a basis for long-term product support, and many others. Deliverable data (customer accessible information) represents only a small fraction of the project data. In general a vast amount of design, development, fabrication, and manufacturing data remains the intellectual property of the developer/producer. Further, the value of data is not limited to its use in support of a particular product: data may have a life cycle longer than that of the product it describes. For instance, data from previous projects forms part of the foundation for new product and process design. Data also supports the enterprise in process redesign and quality. Thus data is essential to competitive position. An enterprise's data—if not properly safeguarded—can also be misused by a competitor to the competitor's advantage. For these reasons, data is an integral part of an enterprise's intellectual assets and overall enterprise knowledge.

OVERVIEW

This standard comprises nine fundamental data management principles (Figure Intro-1).

Principles are high-level descriptive statements about high quality DM; they establish what high quality DM looks like. For each principle there is a set of enablers; the enablers provide the mechanisms of DM.

Figure Intro-1 Data Management Principles


1. Define the organizationally-relevant scope of data management
2. Plan for, acquire, and provide data responsive to customer requirements.
3. Develop DM processes to fit the context and business environment in which they will be performed.
4. Identify data products and views so that their requirements and attributes can be controlled.
5. Control data, repositories, data products, data views, and metadata using an approved change control process.
6. Establish and maintain an identification process for intellectual property, proprietary, and competitive-sensitive data.
7. Retain data commensurate with value.
8. Continuously improve data management.
Feedback
9. Effectively integrate data management and knowledge management

Two different viewpoints are important to DM, corresponding to product and data life cycles. Product data (and related business data) is normally acquired or created as part of the development of a new product or similar initiative. Principle 2, which addresses the planning for and acquisition of data, and Principle 4, which deals with the identification of products, views, and related data elements, are written primarily from the perspective of the individual project. The remaining principles apply at both the project and enterprise levels. Principle 9 relates DM to knowledge management (KM).

The degree to which the DM principles in this standard apply to a product varies over the product's life cycle. Similarly they vary in applicability over the data life cycle. Some principles may not apply during every phase of either life cycle.

This standard addresses the functions of DM but not how to organize for DM. Each enterprise, for valid reasons, locates the functions of DM within enterprise elements that make sense within its own enterprise environment.

This standard is not intended for use as a compliance document or an evaluation mechanism for DM projects. It is intended for use as a source and reference document for either purpose. Appropriate application of the functions and principles in this standard enables the user to plan and implement a DM program for a product, project, or enterprise.

TERMINOLOGY

During creation of this standard, significant effort went into using wherever possible neutral terms. Neutral terms used in this standard are provided in the glossary (Annex B). There is no intent to express preference for any particular terminology set. When planning and documenting a DM program, other aliases may be substituted for the neutral terminology. Three particular sets of terms deserve special mention. The first of these is the pair of terms *program* and *project*. In practice, the term “program” is often used to represent an undertaking that is larger in scope than a “project” but such is not universally the case. This standard consistently uses the term “project.”

Second, this standard introduces some new neutral terminology used here in the context of DM for the first time. Where the terms are introduced for the first time they are explained in context. The most important of these are probably the term *data view*, *data view description,* and *bill of data* (also called a bill of information).

Data normally has and will be a by-product or the result of engineering, management, and other work efforts. Historically, prior to the widespread availability of electronic databases, a *data product* generally resulted from locating, assembling and presenting existing information in the format that a customer had specified. Since the technical work had to be done anyway, and its results recorded anyway, the cost of data of a data product was in locating, assembling, and presenting it. Given the effort and cost involved, it made sense to describe the result as a *product*. Over time there was an increasing recognition that the imposition of customer-specified formats often increases cost without creating equivalent value, prompting a move to utilize supplier formats whenever possible. More recently, manual effort is being replaced by electronic extraction from an existing database; however, and the cost of retrieving, packaging, and even “personalizing” the data is much smaller. Further, what gets captured is a snapshot of data as seen from a particular perspective, at a particular point in time. Products become *views* of the data in the repository. A *data view,* a generalization of the concept of a data product, includes the visual presentation of data by technologies such as digital images, geographical information systems, graphical user interfaces, multidimensional tables and graphs, virtual reality, three-dimensional presentations, and animation.[*]

The *data view description* provides the agreed-to content, preparation assumptions, intended use information, and (where applicable) format for a data view. The data product format can be specified in a data item description (DID), in an eXtensible markup language (XML) style sheet, or by other means.

The list of data views to be provided in accordance with a contract is a *bill of data*, an example of which is a contract data requirements list (CDRL). A bill of data is a two-way concept: a supplier may need data from the buyer in order to perform under a contract. Further, in an integrated trading partner environment both trading partners may obtain views of data, as provided for in a bill of data, from a single data repository.

Finally, references to terms such as the *enterprise*, *organization*, or *performing activity*, *developing activity,* or *producing activity* refer to that enterprise or agency that has the responsibility for performing DM. This enterprise could be commercial or a government agency. References to the customer should be interpreted as the activity that specifies requirements. A customer may be external to the developing and producing enterprise; may be an internal customer such as marketing, management, or the using department; or may even be a supplier in a conventional sense.

REFERENCES

ANSI/EIA Standard 649, Configuration Management.

EIA Standard 836, Configuration Management Data Interchange and Interoperability

Society of Aerospace Engineers Standard AS9034, Process Standard for the Storage, Retrieval, and Use of Three Dimensional Type Design Data

[*] The term *data product* is ubiquitous and the concept of a *data view* is entering the data management vocabulary. At the risk of some ambiguity, this standard retains the *term data product* until the concept of a *data view* becomes more widely accepted. However, data product is used in a generalized sense.

1.0 Principle: Define the Enterprise Relevant Scope of Data Management

Different enterprises come to different conclusions regarding the scope of DM. Traditionally, DM has been thought of as including five functions (Table 1-1).

Table 1-1 Common Functions of Traditional Data Management[†]

Identification and Definition: - Develop and maintain standard data requirement descriptions - Review life-cycle of project or program to determine needs - Identify data requirements - Ensure completeness and eliminate duplication - Provide support to program and enterprise management
Acquisition and Preparation: - Prepare internal data, as identified and prescribed above - Assure that supplier data requirements are negotiated and ordered - Ensure that externally developed data meets requirements for internal use
Control: - Implement controls for import and export of data - Implement controls for safeguarding intellectual property - Implement controls for configuration management - Implement prescribed forms/formats/screens for authorization and use requests - Prepare and maintain master inventory lists. - Assure the appropriate marking of data (e.g., for retention level, proprietary data rights, classification) - Maintain the current and historical metadata about data status and disposition (approval/disapproval, etc) - Establish and document control processes - Implement controls for import and export of data - Implement controls for intellectual property - Implement controls for configuration management - Implement prescribed forms/formats for authorization and use requests - Prepare and maintain master inventory lists. - Mark data (e.g., for retention level, proprietary data rights, classification)
Disposition: - Establish and document records, custodians and project unique records management requirements - Establish a plan for use of a document repository, whether physical or online - Publish and make documents available
Archiving - Create physical or digital files with appropriate archived information - Create project files for decision-tracking histories - Submit archival packages to higher and more general archiving facilities (internal or external) that specialize in data retention

[†] The functions in this table are distilled from GEIA data management panel experience and are intended to be representative rather than comprehensive.

Although these functions remain valid DM tasks, they are no longer a sufficient response to contemporary DM needs. However, there are DM “pockets of excellence” in particular firms and agencies that have addressed the broader foundations of DM. Further, there has been some success in documenting contemporary methods in specific organizations, in the software engineering capability maturity model (CMM), and in the requirements for International Standards Organization (ISO) 9000 certification. None of these, however, has yet brought a unifying, strategic, focus to DM. The intent of this standard it to highlight the importance of the strategic DM and supporting infrastructure. Accordingly, EIA Standard 859 defines a set of four higher-level DM tasks. They are

- DM strategy and architecture development
- DM process and infrastructure design
- DM execution
- DM process and infrastructure maintenance.

The tasks in Table 1-1 are almost all DM execution tasks. This standard adds tasks related to development of a DM strategy and architecture, related process and infrastructure design, and related process and infrastructure maintenance. The new tasks require skills that have not previously been demanded of data managers (Tables 1-2 through 1-5). It should be abundantly clear that most of the skills are not resident in existing DM enterprises, and probably will not be because of the diversity of enterprises. In a rapidly changing technological society, it is crucial to train and continuously improve the resources (i.e., people, processes, tools, budget) that support one of an enterprise’s most valuable assets, data. DM is functionally responsible for ensuring that the integrity and accessibility of data is consistent with the users’ requirements. The data manager, to be effective, should be provided the opportunity to achieve expertise through training, experience, and mentoring. Areas of concentration include planning, acquisition, preparation, control, disposition, archiving, and communication.

A particular enterprise will not necessarily agree that all five functions are part of DM as defined for that enterprise. Obviously, integration from a DM standpoint is improved by consolidating these functions and skills under one enterprise entity. However, the cost might be the sub-optimization of another process. Further, any particular enterprise may decide to simply forego one or more of these tasks—or at least to leave them unmanaged.

Principles 2 through 9 in this standard were generally written from the perspective that DM includes all five of the major tasks and each of the enumerated sub-tasks. Tables 1-2 through 1-5 provide an overview of DM tasks and subtasks as defined in the present standard, as well as related skills. Annex C provides a more complete cross mapping between tasks and skills.

Table 1-2 Strategy and Architecture[‡]

Table 1-2a
Subtasks

Subtasks
Development of DM strategies
Development of DM plans
Development of DM policies
Development of IP strategies
Integration of DM and knowledge management
Resourcing of DM requirements

Table 1-2b
Needed skills

Needed skills	
Clerical	
Budgeting	x
Cost/benefit analysis	x
Strategic planning and management	x
Program management	x
Legal	x
Technical Library management	x
Configuration Management	x
Warehouse management	
Database management	x
Electronic data administration	
Process design and engineering	x
Software engineering	
Knowledge management	x

The right hand columns in tables 1-2 through 1-5 contain skills required by the major DM tasks. The same set of skills is repeated in each table; each skill is important to one or more of the four top-level DM tasks. An “X” opposite a skill indicates that it is normally essential for one or more of the subtasks in the related table. Absence of an “X” indicates a skill is not normally essential to any of the associated subtasks, although it might be appropriate in some circumstances.

[‡] The lists of subtasks and skills in this and subsequent, similar tables in the discussion of principle 1 represent an informed estimate for purposes of illustrating the range and depth of contemporary data management functions and skills.

Table 1-3 Process and Infrastructure Design

Table 1-3a
Subtasks

Subtasks
Design of data access provisions
Development of paper data formats
Development of electronic data formats
Design of DM processes
Design and development of data envirionments
Development of provisions for interoperability and interchange
Development of training syllabi and courses
Development and management of meta data
Design of data products and views

Table 1-3b
Needed skills

Needed skills	
Clerical	
Budgeting	x
Cost/benefit analysis	x
Strategic planning and management	x
Program management	x
Legal	x
Technical Library management	x
Configuration Management	x
Warehouse management	x
Database management	x
Electronic data administration	x
Process design and engineering	x
Software engineering	x
Knowledge management	x

Table 1-4 Data Management Execution

Table 1-4a

Subtasks (previously existing tasks in italics)
Requirements identification and definition
DM risk assessments
Implementation of prescribed formats
Prioritization of data requirements
Control of data requirements
Control of deliverables received
Oversight of data preparation
Data marking
Import/export control
Preparation and maintenance of inventory master lists
Conversion from paper to electronic
Management of data collaboratively developed via IPTs or similar methods
Administrative management of intellectual property
Implementation of access provisions
Data archiving
Data disposal

Table 1-4b

Needed skills	
Clerical	**x**
Budgeting	
Cost/benefit analysis	**x**
Strategic planning and management	
Program management	
Legal	
Technical Library management	**x**
Configuration Management	**x**
Warehouse management	
Database management	**x**
Electronic data administration	
Process design and engineering	
Software engineering	
Knowledge management	

Table 1-5 Process and Infrastructure Maintenance

Table 1-5a

Subtasks
Recurring DM training
Management of electronic repositories
Management of paper repositories

Table 1-5b

Needed skills	
Clerical	
Budgeting	x
Cost/benefit analysis	x
Strategic planning and management	x
Program management	x
Legal	
Technical Library management	x
Configuration Management	x
Warehouse management	x
Database management	x
Electronic data administration	x
Process design and engineering	x
Software engineering	x
Knowledge management	

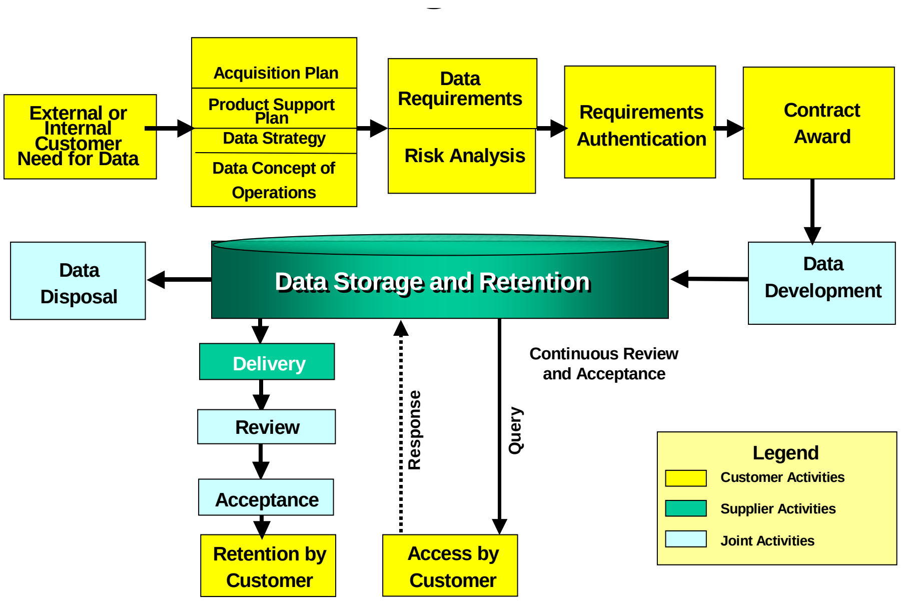
As should be apparent from Tables 1-2 through 1-5, contemporary DM requires a broad spectrum of skills. Again not all of the tasks and sub-tasks in Table 1-2 through 1-5 will be relevant to every organization performing DM. Therefore, each organization will also have its own specific requirements for DM skills.

2.0 Principle: Plan for, Acquire, and Provide Data Responsive to Customer Requirements

Introduction

In order to provide data that is responsive to customer requirements, plan for, acquire, and deliver or arrange for access to data consistent with the contemporary DM model (Figure 2-1). This principle addresses the steps in the model beginning with product need and ending with contract award.[§] Later principles discuss data development; storage, retention and disposal; delivery and access; as well as important related topics.

Figure 2-1. Contemporary Data Management Model



There are five aspects of this model important to the planning for and acquisition of data.

- First, the data customer can be either external or internal to the enterprise.
- Second, there are two different methods by which data is provided to the customer. The mode of delivery can be

[§] Here the term contract is intended to include formal contracts between two companies, formal contracts between a government agency and a company, interdepartmental work authorizations within a company, memoranda of agreement, and any other form of agreement that describes the duties of a supplier to perform DM for a customer. Data may also be provided through a stand-alone contract or, and more generally, as part of a larger contract for goods or services.

   - ➤ In hard copy or, increasingly, electronic form. In this case, there is normally some process by which the customer reviews and accepts the data in the form of a data product, such as a report. The customer then retains and is responsible for the representation of the data that is provided. Ownership (control authority) may or may not transfer, depending on the terms of the contract or agreement.

   - ➤ By providing the customer access to the data in a database or repository maintained by the data developer or a third party. Data products in a conventional sense may not exist and data objects may be formed at the time of need in response to customer-created queries against a database. Because the concept of a data product is so deeply embedded in the practice of data management, this standard will continue to use the term. However "data product" has a more generalized meaning than it has historically.

- ◆ Third, data development, review, acceptance, and disposal may be joint activities, conducted by the data developer and customer.

- ◆ Fourth, planning for data is deliberately linked to the overall product acquisition strategy and long-term sustainment planning through development of a data strategy and data concept of operations.

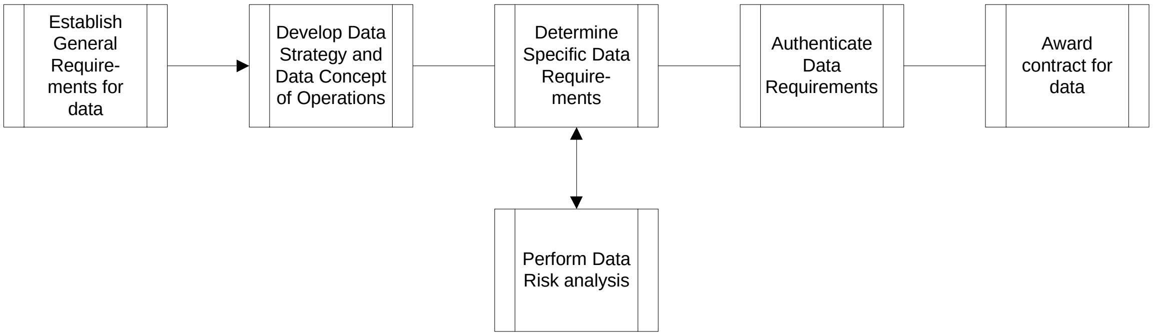
- ◆ Fifth, the data requirements authentication process, which almost always exists in some form, is preceded by a data risk analysis that examines

   - ➤ The risks of not providing for delivery or access to data

   - ➤ The risks of over-procuring data (e.g., where the data may become rapidly obsolete).

The enablers for Principle 2, which follow the logic of Figure 2-1, are diagrammed in Figure 2-2 and detailed below.

Figure 2-2. Principle 2 Enablers


Establish General Require-ments for data
Develop Data Strategy and Data Concept of Operations
Determine Specific Data Require-ments
Authenticate Data Requirements
Award contract for data
Perform Data Risk analysis

2.1 Establish general requirements for data

At the start of a project, the first step is to review the project strategy and planning to determine the anticipated general needs for data delivery or access throughout the product life cycle. Because specific data requirements may not yet be known, a practical way to proceed is to examine the data requirements of recent, similar projects. The output of this enabler is a general description of potential requirements. At this point, before establishing a data strategy, it is not necessary or even desirable to attempt to define specific data requirements. This review should consider data that may be needed to support design oversight, business oversight, manufacturing, testing, operation and maintenance; and documentation that will be needed for legal, tax, historical, internal audit, or other valid purposes. The intent is to recognize the spectrum of data views that may be needed to support the project tasks and products throughout the product life cycle. (As noted on page 11, a data view is a generalization of the concept of a data product. It includes the visual presentation of data by technologies such as digital images, geographical information systems, graphical user interfaces, multidimensional tables and graphs, virtual reality, three-dimensional presentations, and animation.)

Included are not only project specific requirements, but also related data that may be needed to meet broader enterprise or external requirements. It should be feasible at this point to anticipate the form in which the data will be required and if access or delivery is appropriate.

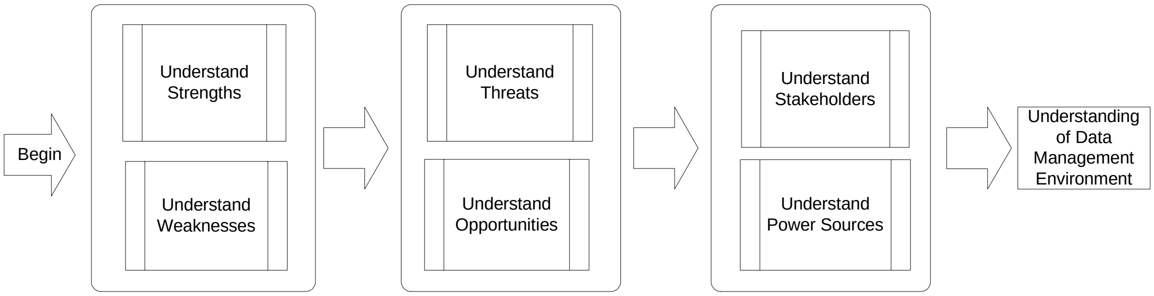
2.2 Develop data strategy and data concept of operations

This enabler is the project-level equivalent of Principle 3. Principle 3 provides a more comprehensive treatment at the enterprise level.

A data strategy is important because all decisions, including those related to data, have consequences. A data strategy creates the potential for data-related decisions to contribute to long as well as short-term goals for the project and enterprise by aligning DM with the context in which it lives. The general data requirements (Enabler 1) define the data "mission"—what needs to be done for whom, how, and why—the first step in any strategic analysis. The next step in the formulation of a data strategy is a data environmental assessment (DEA) (Figure 2-3). The DEA, in the context of the general requirements, examines:

- The internal strengths and weaknesses of the project and enterprise
- External opportunities and challenges
- Stakeholders, and power sources

Figure 2-3. Data Environmental Assessment



A strength, for instance, might be the existence of defined processes and accompanying technical infrastructure for vaulting electronic data. A weakness might be lack of training on those processes. Strengths promote and weaknesses limit the ability of the project to satisfy the project data requirements. Understanding opportunities and challenges external to the project reveals how the external environment can affect DM. An example threat could be a corporate intent to retire a database the project had intended to use. A classic opportunity is an informed customer who creates a need to develop new capabilities that have market potential beyond that customer. Understanding stakeholders and power sources is important because stakeholders have to be satisfied and power sources influence resource allocation. Taken together, strengths, weaknesses, opportunities, challenges, stakeholders, and power sources identified through the DEA define what can be done without change and what will need to change to satisfy the project data requirements. The DEA outcome is the basis for defining a course of action to solve gaps and capitalize on opportunities.

All the above determine if the needed capabilities will be in place and starts the process of creating them if they are not. It is also essential to describe how those capabilities will be employed. This is the role of the data concept of operations. The data concept of operations should cover:

- Who the customers are
- The range and depth of data to be provided, and over what period of time
- The nature of the business relationship including how it is expected to change over time, if applicable
- How data views will be generated and provided to the customer (e.g., hard copy, electronic copy, access from a data server)
- Quality assurance provisions and
- Where resources will come from

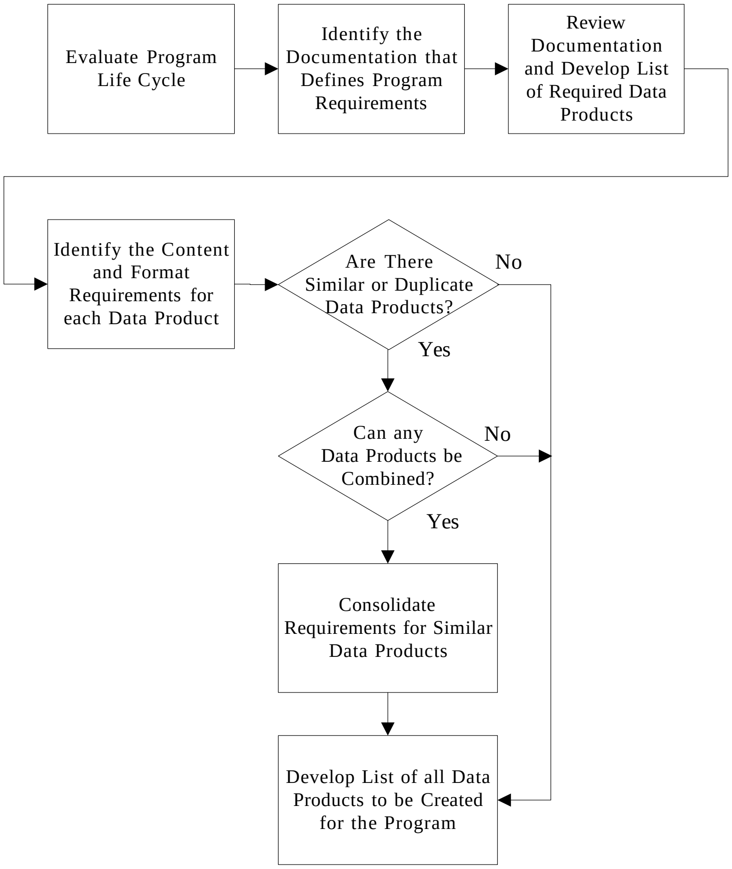
2.3 Determine specific data requirements

Next, determine specific data requirements. The steps involved include determining who needs data, what data are needed, what views of the data are required (hard copy, responses to database queries, data interchange conventions, or other specifications), when the data will be required (in terms of project milestones or calendar dates), and the delivery mechanisms

2.3.1 Determine the needs for data

The first step is to identify the data products that will be needed to support the project throughout its entire life cycle. Figure 2-4 shows the data product identification approach.

Figure 2-4 Review Project Life Cycle to Identify Data Requirements and Determine the Needs for Data


Evaluate Program Life Cycle
Identify the Documentation that Defines Program Requirements
Review Documentation and Develop List of Required Data Products
Identify the Content and Format Requirements for each Data Product
Are There Similar or Duplicate Data Products?
No
Yes
Can any Data Products be Combined?
No
Yes
Consolidate Requirements for Similar Data Products
Develop List of all Data Products to be Created for the Program

The first step in this process is to understand the project life cycle, and, in that context, to review the project requirements documentation to determine the types of data products that will need to be created. Include not only project specific requirements, but also any data that may be needed to meet enterprise or government requirements. Not all data requirements may be documented at the start of a project. There may be a need to anticipate some data requirements based on the potential for future need. As discussed under enabler 2.1, it may be useful to examine the requirements of similar projects.

It should be recognized that the need for data products may change throughout the life-cycle of the project. Management should assess the impact that any change to the project requirements has on the need for data products, or the requirements for those data products.

Once the types of data products have been determined, their specific content and format requirements need to be defined. This task includes the definition of the processes used to create each different type of data product.

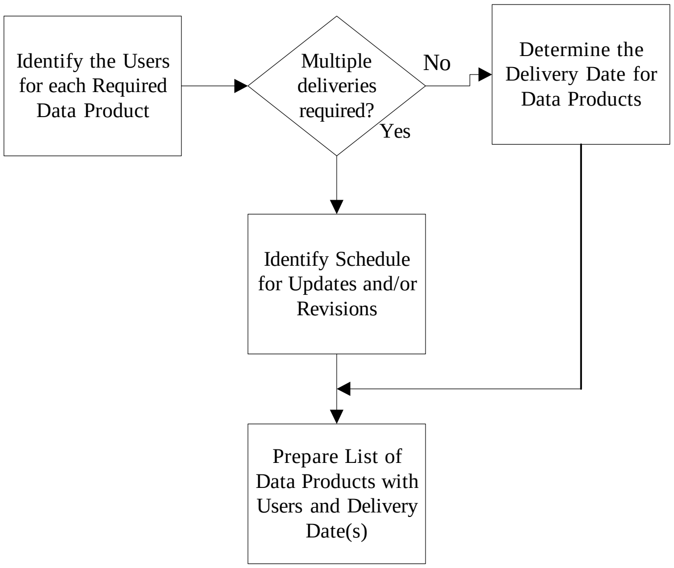
As a part of the process of defining requirements, consideration should be given to the tradeoffs between consolidating the requirements for similar data products versus providing the ability, possibly on the fly, to personalize products. Historically, when the term data product meant the production of a hard copy, it was important to consolidate requirements for similar products to minimize the number of items to be created, managed, and maintained. With electronic access to data using integrated digital environment technologies, the benefits from personalizing the presentation of data to fit individual needs can outweigh the decreasing cost of doing so.

The end result should be a consolidated list of data products needed to support the tasks and products of the project throughout their entire life cycle, the format and content requirements for those data products, and process information necessary to ensure that the data products that are created will meet the noted requirements.

2.3.2 Identify the users of the data and establish the frequency of data delivery

Identify the specific users of the data products and establish when the users need each of these items. A user may be either external to the enterprise—a customer in the usual sense—or an internal customer. Need may be defined in terms of specific calendar date, or in relation to an event in the project life cycle. Figure 2-5 shows the steps included in this process.

Figure 2-5 Identify Users of the Data Products and Establish When Data Will Be Needed


Identify the Users for each Required Data Product
Multiple deliveries required?
No
Yes
Determine the Delivery Date for Data Products
Identify Schedule for Updates and/or Revisions
Prepare List of Data Products with Users and Delivery Date(s)

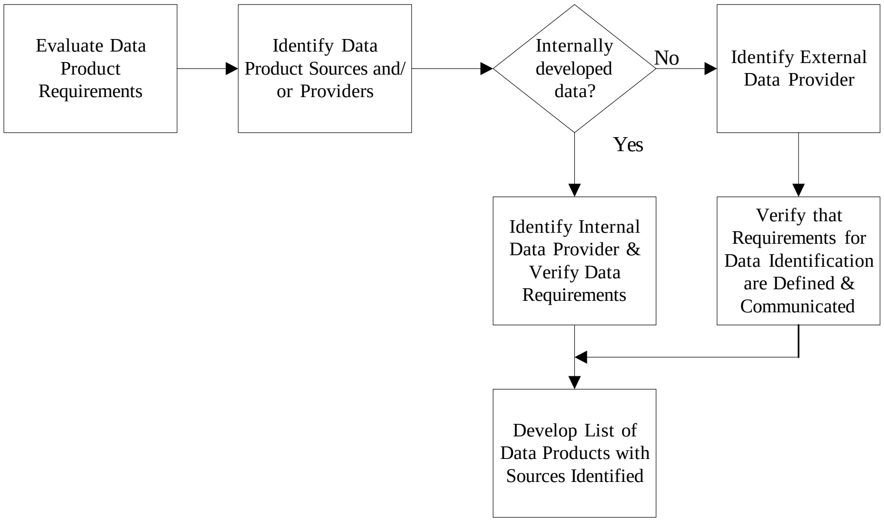
Based on life cycle requirements of the project, determine the users who need to have access to each of the completed data products, and any interim data products. If multiple areas use a data product, record all areas that require the information along with the dates needed.

Work with the users of each data product to verify the data products that they need, and the required need dates. If the data will be updated, or is periodic in nature, determine the necessary frequency for any updates.

2.3.3 Relate data requirements to the functional areas responsible for data generation and distribution

In cooperation with the functional areas who will produce the data, determine who will provide the data and how. More than one functional area may be involved in the creation of some data products. Figure 2-6 shows the steps included in this process.

Figure 2-6 Relate Data Requirements to the Functional Areas Responsible for Generating the Data


Evaluate Data Product Requirements
Identify Data Product Sources and/ or Providers
Internally developed data?
No
Identify External Data Provider
Yes
Identify Internal Data Provider & Verify Data Requirements
Verify that Requirements for Data Identification are Defined & Communicated
Develop List of Data Products with Sources Identified

The first step of the process requires as an input a complete listing of the data products required by the project.

Based on the type of data, determine the functional area (or areas) that are responsible for the generation of each of the data products. If a data product requires input from multiple areas, note the areas that provide source information and the area with final responsibility for the finished data product. Ensure that the data requirements, including marking requirements, are clearly defined and documented (see principle 6). If an internal source is responsible for the preparation of a data product, enterprise procedures should ensure that data requirements are communicated to the responsible functional area(s). If a subcontractor or other outside source is responsible for the preparation of a data product, ensure that the requirements are properly communicated. Provide the functional area with the schedule and the supporting information, make sure they understand what is required, and secure commitment. If requirements were collaboratively developed, implementation of this step is simplified.

2.4 Perform risk analysis

The management of data involves recognition of multiple sources of risk. They include the following sources:

1. Under-provisioning of the data, failure to provide data that is needed when it is needed

2. Over-provisioning data, providing data that is not useful or providing data prematurely, to the detriment of its accuracy

3. Inability to retrieve data as a result of non-existent or inadequate cataloguing and metadata

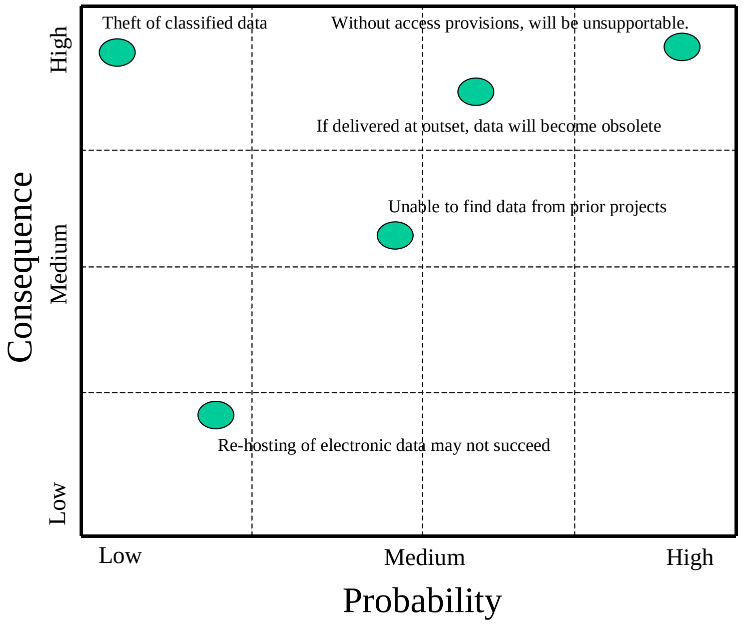
4. Data loss, whether due to misplacement, theft, or a natural disaster
5. Data obsolescence—retaining data that is of no value.
6. Compromise of intellectual property

Although the specifics of projects differ from one another, a demonstrated risk analysis method follows:

- Recognize and enumerate the sources of risk
- For each risk determine
  - The likelihood of occurrence and
  - Severity of consequence if the risk materializes.

Simple scales (e.g., high, medium, low) for evaluation often are good enough to characterize both probability and consequence (Figure 2-7). Priority for risk mitigation then belongs to those risks that, in relation to others, have combinations of higher probability and consequence.

Figure 2-7 Example Risk Portrayal



Characterization of the first two sources of risk (under and over provisioning) is an important input to data authentication. It provides a basis for understanding and defending data requirements that should be satisfied, as well as those that should not.

Understanding the third through sixth source of risk is important to process adequacy and potential process redesign. There may, be other sources of risk as well; this list is not intended to be comprehensive.

Finally, risk management is an iterative process rather than an event. Risk analysis is an inherent part of the data requirements definition and consolidation steps discussed earlier.

2.5 Authenticate data requirements

Authentication is the capstone task prior to contracting for or authorizing internal development of data. The purpose is to make sure that requirements as defined in a bill of data are valid, complete, and make sense from a business standpoint. In authenticating data requirements, address the following questions as a minimum.

- Are a DM strategy and DM plan in place to guide overall data acquisition for the affected project? Are the strategy and plan adequate? Does the proposed procurement of data follow the strategy and plan?
- Does the proposed bill of data respond to user requirements? Specifically does the content as well as the types, formats, and delivery or access timelines respond to user and enterprise needs?
- Has a risk assessment been performed? Is the risk assessment reasonable—i.e., are the risks understood and the approach to risk mitigation reasonable?
- Have requirements been adequately integrated to resolve duplicate requirements? Did the integration effort look for and resolve implied or missing requirements? Were non- or low-value added requirements identified and resolved?
- Are adequate quality assurance measures specified so that the data received, generated, and used will be appropriate and suitable?
- Have data rights issues been addressed adequately?
- If future access or contingent requirements are involved have data maintenance, configuration management, and (if appropriate) deferred data delivery, deferred data ordering, or third-party escrow been addressed?

2.6 Contract for data

Contract award is the final enabler for this principle. As noted earlier in the discussion of this principle, the term contract is intended to include formal contracts between two companies, formal contracts between a government agency and a company, interdepartmental work authorizations within a company, memoranda of agreement, and any other form of agreement that describes the duties of a supplier to perform DM for a customer. Data may also be provided through a stand-alone contract or, and more generally, as part of a larger contract for goods or services. Here what is contracted for can take many forms including the following.

- Data access under agreed-to provisions (i.e., over what period of time, who may access, purposes of access, limitations, etc.) This approach is becoming increasingly important and does away with delivery, per-se. "Delivery" when it is needed at all is effected by notification that the data is available to be accessed.

- Conventional delivery at a specified time or in conjunction with a specified event.

- Deferred data delivery. Used when it is in the buyer's interest to defer the delivery of data. As an example, when design is still evolving and what is desired is technical data that correspond to the final design. Establishes an obligation on the part of the supplier to deliver data up until some specified time period (e.g., two years) after contract termination or the date of acceptance of the last item other than technical data or computer software.

- Deferred data ordering. Used when a firm requirement for a particular data item(s) has not been established prior to contract award but there is a potential need for the data. Under this provision, the buyer may order any data that has been generated in the performance of the contract, or any subcontract, until some specified time (e.g., three years) after contract termination or acceptance of all items other than technical data or computer software.

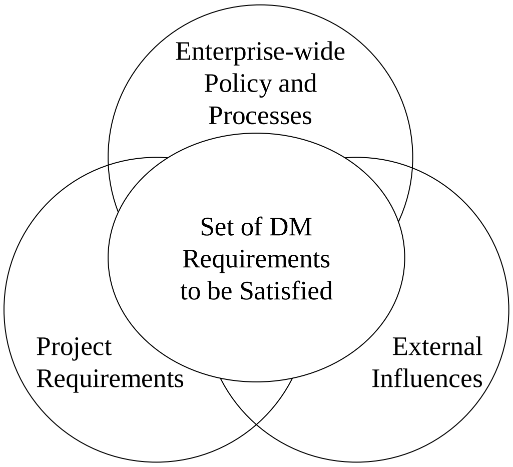
- Third party data escrow. Used when it is not in the buyer's interest to take immediate delivery of data but the buyer needs assurance that data will be available to the buyer in the event that the supplier goes out of business, decides to stop supporting the related business line, or for any of a number of similar reasons might be unable or unwilling to provide needed data. Simultaneously, by placing the data in the hands of a disinterested third party, protects supplier technology and intellectual assets until release under specified conditions. Can include provisions for periodic updates (e.g., when versions change) and verification.

3.0 Principle: Develop DM Processes to Fit the Context and Business Environment in Which They Will be Performed

Introduction

To be effective, DM solutions, processes, and practices should supported by a realistic analysis and understanding of the business context and environment in which they will be performed. The business context and environment are characterized by both internal and external factors; DM solutions are necessarily conditioned by those factors. Requirements to be satisfied come not only from projects themselves but also from future expectations related to projects, from enterprise policies and processes, and from the environment external to the enterprise. Taken together these sources define the context and business environment in which DM will operate (Figure 3-1).

Figure 3-1 DM Requirements


Enterprise-wide Policy and Processes
Set of DM Requirements to be Satisfied
Project Requirements
External Influences

In a particular circumstance the project-specific, enterprise-wide, and externally imposed requirements can be complementary. In this instance, planning for their solution amounts to identifying all of the requirements and deciding, within available resources, which can be satisfied, when they can be satisfied, and how. But the requirements can also be in conflict. An example would be a requirement for enhanced data sharing and a simultaneous requirement to improve controls over intellectual property. Regardless of whether they are synergistic, additive, in conflict, or—more likely—a mixture of the three, it is DM's task to identify and then address the full set of requirements, including examining trade-offs where appropriate.

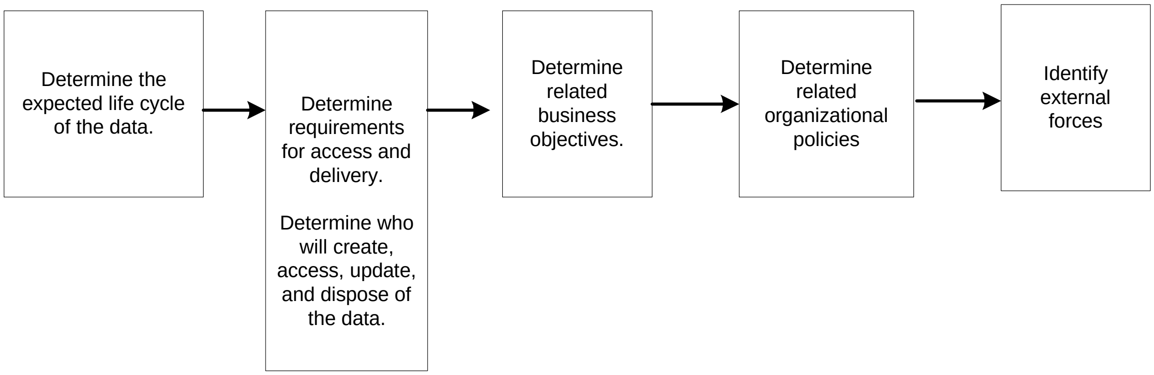
The identification of project-specific requirements is addressed by Principle 2. Enterprise requirements and the requirements arising from the environment external to the enterprise are integrated with project-specific requirements by the current principle. This principle addresses the four (4) major components of a successful DM solution

- Deriving the complete set of DM requirements
- Determining the shape of the preferred DM solution
- Comparing the proposed, best solution to existing and planned enterprise process infrastructure
- Developing needed adjustments that fulfill the total set of DM solution requirements by resolving gaps and conflicts.

3.1 Determine the complete set of requirements that the DM solution must address.

Prior to developing new DM strategies and solutions, identify the general set of requirements to be addressed. This includes not only the requirements for data, but also the broader requirements that relate to data capabilities and data processes. To identify these broader requirements it is important to understand, as a minimum, the intended use of the data, related business objectives, technology issues, and external constraints. The steps listed in Figure 3-2 outline a process for developing a complete set of requirements.

Figure 3-2. Process for Understanding Requirements


Determine the expected life cycle of the data.
Determine requirements for access and delivery.
Determine who will create, access, update, and dispose of the data.
Determine related business objectives.
Determine related organizational policies
Identify external forces

As illustrated in Figure 3-2, among the essential considerations are the following, although this list is not exhaustive.

1) Determine the expected life cycle of the data and expected use of the data. Is data being developed or acquired against a one-time requirement or is there likely to be a recurring requirement for the data?

2) Determine who will create or acquire the data. At least three situations can apply: customer developed or acquired, developed or acquired for the customer, or collaboratively developed (Table 3-1)

Table 3-1 Creation and Acquisition of Data

Who creates or acquires	Passes through "hands" of data manager	Considerations
Customer developed and provided	Yes	• Important to understand what the customer expects in the way of data inventory management and protection for data provided
Developed or acquired for the customer	Yes	• Realm of "traditional" DM and, assuming reasonably unambiguous requirements, the easiest for which to plan.
Collaboratively developed	Probably no.	• Growing in prominence as a result of increased use of integrated product/process teams (IPTs) and other trust-based relationships. • DM task is to put in place the means and processes for IPT-level self-management and then to oversee that self-management.

3) Determine the expected requirements for access, for delivery, for maintenance, for storage, for protection, and for disposal over the life cycle. As an example, it is reasonable to expect that data created during the early design phases of a project will be important during later phases. As another example, it is important to determine if the customer is likely to want delivery or access.

4) Determine who, over the life cycle, will have access to, be responsible for updating, and be responsible for disposal of data (Table 3-2). Assess which of these cases is the most likely, and plan accordingly.

Table 3-2 Responsibility for Updating and Disposing of Data

Case	Comments
Customer has requested delivery with no provision for updates	Customer may take responsibility for maintaining the data current, or it is possible that the customer has not yet considered the need for update and disposal.
Customer has, or is likely to, request either delivery or access sometime in the future	Customer will probably want the data developer to maintain the data current

5) Determine if there are related business objectives and considerations. For instance,

   a. Will the data potentially be reused or repurposed?

   b. Is there any provision for warranting the correctness of the data? If so, then clearly this needs to be taken into account and planned for from both a process and financial standpoint.

   c. Is any of the data important from an intellectual property standpoint?

   d. If there is, or is likely to be, a requirement to provide for continued long-term access, what are the provisions for assuring access if the enterprise ceases to exist in its current form (e.g., as a result of reorganization, a buyout, or a decision to abandon the line of business)?

   e. Does the enterprise want to be in the business of data warehousing (for either electronic or non-electronic data) or will this responsibility be outsourced to a third party?

   f. Is the data requirement indicative of an emerging market for the enterprise, a market the enterprise is maintaining, or a market the enterprise intends to withdraw from? The appropriate investments (time, infrastructure, acquisition of staff, training) are different for each case.

6) Determine what enterprise policies pertain? For instance,

   a. Does the enterprise have in place policies that promote either centralized or decentralized management of data?

   b. Does the enterprise have in place policies related to retention and disposal?

   c. Is there a formal or informal policy related to disaster planning and recovery, such as maintaining data in more than one physical location?

7) Identify what external forces apply.

   a. Has the customer requested compliance with national or international standards? If so, it will be important to understand the extent to which changes in the standards, some of which cannot necessarily be foreseen, create new requirements?

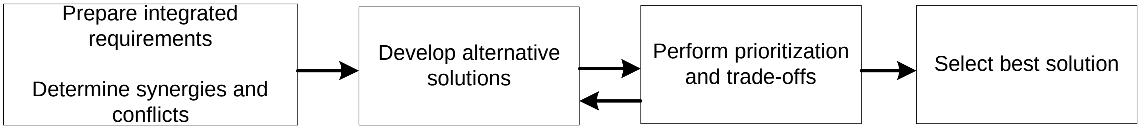
   b. Similarly, has the enterprise adopted national or international standards that apply?

   c. Are there either national or international legal requirements (e.g., with respect to intellectual property such as patents or copyrights) that have to be respected?

3.2 Determine the shape of the DM solution.

Next, consider the broad characteristics of the DM solution and what it must address, but not "how" the solution will be implemented. A complete solution includes an analysis of all factors: internal, project-specific, and external. Figure 3-3 illustrates the essential steps, a matter of applying standard systems engineering precepts to the development of a DM solution. Although the process in Figure 3-3 is portrayed as linear, it generally is iterative, especially requiring some cycling back and forth between alternative development and prioritization.

Figure 3-3. Process for Determining the Shape of the DM Solution


Prepare integrated requirements
Determine synergies and conflicts
Develop alternative solutions
Perform prioritization and trade-offs
Select best solution

Assemble the requirements identified by Enabler 3-1 in a form that can be worked with for purposes of DM solution design. One way to do this is to list the requirements according to their source and relative priority. As an example, group requirements in terms of:

- ◆ Current external customer contract requirements for specific deliverables or access requirements.

- ◆ Current internal customer contract requirements for specific deliverables or access requirements.

- ◆ Supplier requirements—i.e., requirements for data to be provided to suppliers rather than external or internal customers. An example might be a set of envelope drawings that a supplier needs in order to proceed with detailed design.

- Derivative, intermediate datasets required to satisfy external or internal customer requirements. As an example, even though a formal logistics support analysis may not be called for in the contract from an external customer, the logistics support analysis may still be needed in order to determine sustainment requirements and design a supply chain.

- Anticipated near-term new contract or internal requirements for deliverables or access requirements.

- Anticipated longer-term contract or internal requirements.

- Requirements imposed by enterprise policy and practice.

- Requirements, such as those found in the uniform commercial code or law, which come from the environment outside the enterprise.

Normally the requirements from the process just described are not fully independent of each other (Table 3-3)

Table 3-3 Interdependent Requirements

Case	Comments
Duplicative requirements	◆ Important to detect and look for ways to combine requirements so they can be satisfied with a single rather than multiple efforts.
Conflicting requirements	◆ Can be simple—such as needs for the same or similar data, but at different points in time. ◆ Can be more complex — such as competing needs for data sharing and protection of intellectual property rights.
Synergistic requirements	◆ Example is case where the data created in response to one requirement when integrated with or augmented by data created in response to a second requirement provides a solution to a third requirement. ◆ Almost always introduces time dependencies; important to capture interdependencies in a process chart, network chart, or by similar suitable means.

The method for portraying integrated requirements varies from circumstance to circumstance. In some cases, a simple database is sufficient. In others, particularly where process and capability considerations are important, a narrative report may be required. A strong DM solution depends on a clear understanding of the relationships.

Develop a set of alternative solutions for the comprehensive set of requirements. Each solution is a scenario: a particular combination of processes, enterprise elements, and infrastructure elements. The processes, enterprise elements, and infrastructure elements do not need to already exist. In fact, it would be a mistake to consider only elements that exist since doing so almost certainly constrains improvement. In particular, do not be limited by pre-existing policies and practices. These policies and practices are essentially the “residue” of previous decisions. Although valuable because they represent enterprise learning they can also be counterproductive if that learning is not relevant to the problem being studied. Whether or not the cost (monetary, labor, time, and or good will) to change policy or practice is worth it can be considered as part of the evaluation process.

Generating alternative solutions is typically the most difficult of the steps because it involves idea generation; it is harder to envision what could be than what already is. Consider creating alternative solutions through some form of brainstorming exercise since it has proven to be effective. As an aid to creating alternatives, this is the point in the process to identify and consider incorporating best practices. When a set of candidate alternatives has been generated, perform a top-level, first-order analysis to separate feasible from infeasible alternatives. For instance, there is no point considering an alternative that violates rules of physics or would require resource investments beyond that which has any likelihood of being available.

Prioritize the feasible solutions in terms of their ability to satisfy the requirements and cost of implementation. This is almost always a matter of comparing multiple solutions in terms of their ability to satisfy multiple objectives. Perform this step by doing the following:

- Prioritize the requirements in terms of high, medium, low; on a nine-point scale; or through some similar notation.

- Evaluate the ability of each solution to satisfy each requirement, again using an appropriate and complementary scale.

- Evaluate the implementation cost of each solution on an appropriate scale (e.g., low, medium, high). Highly precise cost estimates are not normally worth the effort to prepare at this stage of analysis. Since more than financial costs are normally involved, be certain to consider non-monetary as well as monetary costs.

- Determine which alternative does the best job of satisfying the most important requirements at the most attractive cost and risk. Although this process can be painstaking, the tradeoffs associated with alternative solutions are frequently those most important to the decision making process. Well-crafted solutions yield positive results for the enterprise, better processes and practices for the future, and even the development of competitive edge methods.

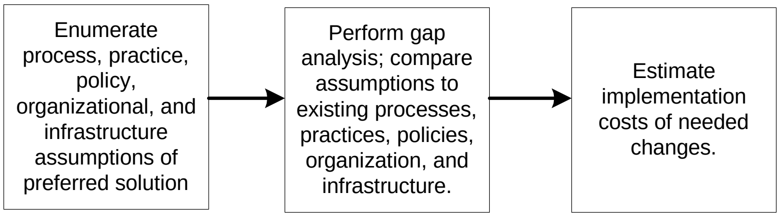
The structured approach described above may show that no proposed alternative solution satisfies enough of the priority requirements to be satisfactory. If this happens, then back up a step and consider additional alternatives. Even after considerable effort, it is possible to not have a feasible alternative. In this case it may be necessary to go back one more step to make sure the requirements were correctly understood and potentially challenge them. Those who established requirements did not necessarily have available to them an understanding of the feasibility or affordability of solutions that would satisfy the stated need.

The best solution is then derived using all relevant considerations The previous step will have identified the best overall solution, framed in the context of benefits, risks, gains, and losses that other solutions may represent. Review the ranking for reasonableness and, in addition, consider other factors that are relevant but may be difficult to quantify.

3.3 Compare the proposed, best solution to existing and planned enterprise capability (infrastructure and processes).

Given that most organizations have some DM capability already in place, it is important to compare the needs of the proposed, best solution to existing and planning enterprise capability (Figure 3-4).

Figure 3-4. Process for Comparing Proposed Solution to Existing and Planned Enterprise Capability


Enumerate process, practice, policy, organizational, and infrastructure assumptions of preferred solution
Perform gap analysis; compare assumptions to existing processes, practices, policies, organization, and infrastructure.
Estimate implementation costs of needed changes.

Conduct a detailed examination of the preferred solution in order to identify process, practice, policy, enterprise, and infrastructure characteristics that are required to implement it.

Perform a gap analysis by comparing the needed characteristics to those of existing processes, practices, policies, organizational alignments, and infrastructure. For instance, if the preferred solution involves electronic storage of quantities of digital data over extended periods of time, determine if the enterprise's infrastructure plans are supportive. Similarly, if the preferred solution involves processes different from those in place then process reengineering is required.

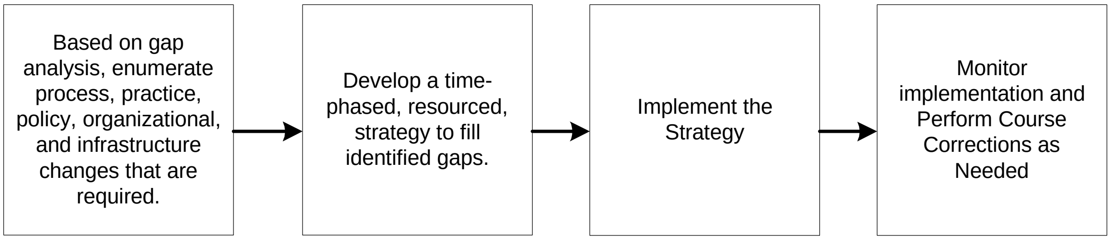
Identify any conflicts or roadblocks (e.g., enterprise plans to remove infrastructure that will be needed) that will have to be overcome or considered. Current processes, practices, policies, and infrastructure capabilities may not be in easily retrievable form. Part of the gap analysis effort may entail knowledge capture and a subsequent documentation effort sufficient to perform the gap analysis.

Finally, determine the monetary and non-monetary implementation costs of the needed changes. Since this estimate provides a basis for advocating for and allocating resources, estimate the resources in financial terms (e.g., in dollars) within an acceptable range of uncertainty (e.g., 10 percent)

3.4 Make needed adjustments in processes, practices, policy, organizational alignment, and infrastructure.

Using the comparison results derived in the previous enabler, make needed adjustments to resolve the gaps (Figure 3-5).

Figure 3-5. Process for Making Needed Adjustments in Processes, Practices, Policies, Enterprise, and Infrastructure


Based on gap analysis, enumerate process, practice, policy, organizational, and infrastructure changes that are required.
Develop a time-phased, resourced, strategy to fill identified gaps.
Implement the Strategy
Monitor implementation and Perform Course Corrections as Needed

Define the needed changes in processes, practices, policy, organizational alignment, and infrastructure. Some of this work will have been completed in the prior enabler.

Using the comprehensive list of needed changes, develop a time-phased, resourced strategy to resolve the gaps. There are many reasons for devoting an appropriate amount of effort to the development of this time-phased strategy; some are listed below.

- Not all requirements need to be satisfied up-front. Since money has a time value, projects to put in place new capabilities should be initiated lead-time away from their need. There is also a secondary benefit of lead-time away implementation: there is less uncertainty what the requirements are.

- The enterprise may not be able to provide all of the resources up-front. A carefully considered plan enables the enterprise to implement changes with the highest return on investment first.

- Enterprises can only absorb so much change at once in any event. Particularly where changes in processes, practices, and organizational alignment are involved, it is important to develop buy-in through appropriate education and other forms of outreach and communication. Time needs to be allocated for outreach and communication.

The next task is to implement the defined strategy.

Finally, monitor implementation and make course corrections as needed. Aside from detecting implementation problems, it is unlikely that the solution was as completely correct as initially envisioned. Further, there will be fact-of-life changes in requirements as time goes on that will need to be addressed.

4.0 Principle: Identify Data Products and Views so That Their Requirements and Attributes can be Controlled

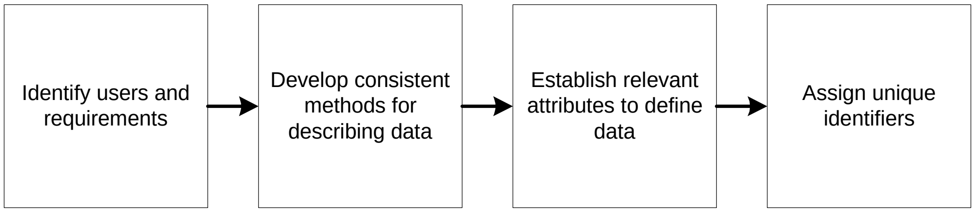
Introduction

Data is of value to the enterprise when it can be located or accessed by users. Metadata, *or data about data*, is essential for data managers and others to identify, catalog, store, search for, locate, and retrieve data. Metadata includes attributes and relationships and is further described in enabler 4-1. Careful consideration of requirements when selecting elements of metadata enhances the ability of users to locate data regardless of storage medium or the amount of data stored. Creating standard processes for selecting metadata provides for consistent, uniform, repeatable processes that can be tailored to specific business requirements. Further, using uniform processes saves time, reduces cost, and allows projects to reap economies of scale through adoption by multiple users or enterprises that exchange data.

Not all data is delivered as a data product; if anything the trend is away from delivery and towards access as needed. When access is provided for, an authorized user can retrieve data that has been grouped or organized to meet specific needs—what is referred to in this standard as a *data view*. Data views, whether implemented as queries, XML schema, or by other means are described by metadata. Particularly where the data views are complex, and when it is important to ensure that the same view is provided each time it is needed, it is important to define and control the metadata.

The purpose of this principle is to ensure that metadata is selected to enable effective identification, storage, and retrieval of data so that the creation and retention of data can be properly managed. Figure 4-1 illustrates the process at the top level.

Figure 4-1. Data Product Identification Enables the Control of Requirements and Attributes


Identify users and requirements
Develop consistent methods for describing data
Establish relevant attributes to define data
Assign unique identifiers

The process begins with identification of users and a review of the requirements to identify data that need to be developed or procured. This can be an iterative process that may reveal additional data requirements.

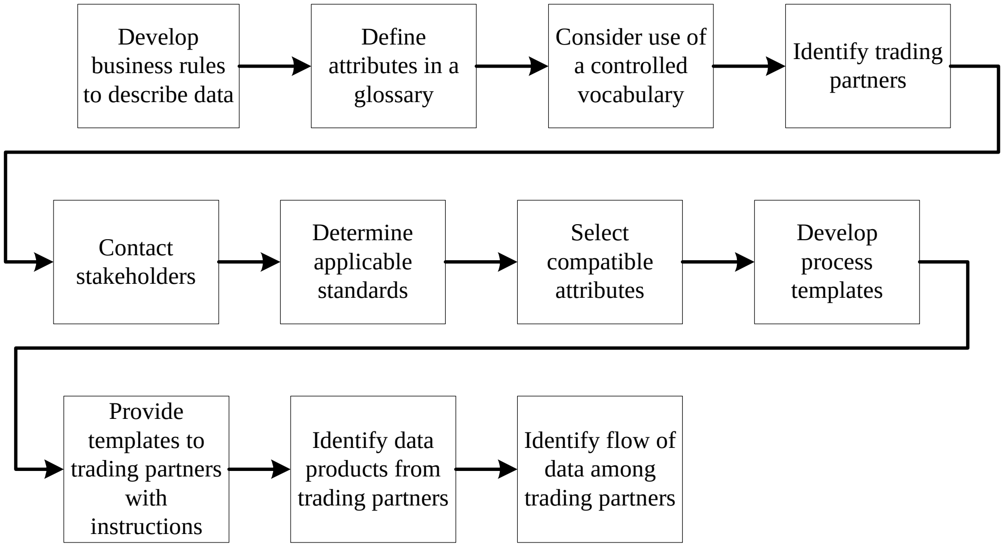
Develop consistent methods for describing data. Doing so avoids the confusion that comes from calling a data element an “author” in one context, “person_author” in another, “document_author” in a third and so on when they all are describing the same thing. Use the consistent methods from step two to establish relevant attributes for the project's data and then assign unique identifiers, what are usually called "keys" in database terminology. The unique identifiers are the attributes (e.g., document number and version number) that make it possible to unambiguously distinguish one product from another. Each of these steps is described in more detail in the lower level enablers.

4.1 Develop consistent methods for describing data

While the types of data to be managed vary among enterprises and projects, the process for establishing metadata can be standardized. Consistent development and use of metadata enables effective communications across enterprises exchanging data as well as within and between enterprises over time. The process for selecting metadata should be coordinated with users or other enterprises to ensure compatibility among those who will exchange data. Process templates can be used to provide a consistent, repeatable method for identifying the data products and flow of data among enterprises

Attributes are the properties that uniquely characterize the data, such as document number, title, date and data type. A metadata record consists of a set of attributes necessary to describe the data in question. While identification of attributes initially occurs during the early stages of planning, it should be seen as an iterative process throughout the data life cycle. New methods of data storage and new types of data may evolve, requiring different ways of storing and retrieving data. Changes to metadata should support multiple paper or electronic storage and retrieval approaches, while maintaining the integrity of existing attributes. See Figure 4-2.

Figure 4-2 Process for Consistently Describing Data


Develop business rules to describe data
Define attributes in a glossary
Consider use of a controlled vocabulary
Identify trading partners
Contact stakeholders
Determine applicable standards
Select compatible attributes
Develop process templates
Provide templates to trading partners with instructions
Identify data products from trading partners
Identify flow of data among trading partners

Develop business rules to consistently describe data throughout the life cycle. Select attributes from a "controlled vocabulary," which is a limited set of consistently used and carefully defined terms. This is critical to ensure effective retrieval. Without basic terminology control, inconsistent metadata diminish the quality of search results. Ideally the controlled vocabulary is not project specific but is created at the enterprise or higher level in the form of a standard data dictionary, standard ontology, or similar means and applied consistently to all projects (see enabler 4.1.1 below).

4.1.1 Ensure data interoperability between team members

During selection of metadata attributes, identify team members who potentially create data, update data, exchange data, enter data into a repository, or search for data. Contact team members to obtain input and coordinate requirements. While it is desirable to standardize attributes it may be expensive to do so if modification to existing data systems is required. An alternative is for each to map to a neutral standard. In any event, standards invoked by a customer should be flowed down to team members and understood by all parties. Use of standards, such as EIA-836, *Configuration Management Data Exchange and Interoperability* and the Universal Data Element Framework (UDEF) enhances the ability to exchange data.

4.1.2 Apply processes to characterize data and data products to ensure adequacy and consistency

Processes should be developed to map the flow of data throughout the life cycle. The use of a template provides a consistent, repeatable method to identify data products and the flow of data between users. Use of templates helps ensure consistency across the enterprise in defining data products Data owners and users are identified in the process, along with any requirements associated with metadata. A template, for instance, could help identify commonly needed fields for any product, the associated metadata, and valid entries for the data.

Once processes are developed and tested, users should be trained in using the templates to identify the data products. Users should be provided with the templates along with instructions for use and possible tailoring. The purpose, expected results, and any ground rules should be identified to assist users in accomplishing their goals. Consistent use of the templates helps in the exchange of data among users. Table 4-1 is intended as a representative sample of some types of attributes that may be selected by an enterprise. Specific titles and descriptions are defined by the project or enterprise to meet specific requirements. A glossary, often referred to as a data dictionary, is required to define each attribute. An attribute such as “Document Type” can mean different things on different projects and to different enterprises although the use of a controlled vocabulary acts to restrain proliferation.

Table 4-1. Metadata Examples.

Attribute	Description
Author	Originator of the document or file
Classification	Level of security classification or business sensitivity
Contract Identifier	Contract number or other identifier
Date Modified	Date of revision
Date Originated	Date of document or file. May be date of creation, date of approval, or date entered into repository.
Document Number	Unique number assigned to a document using a numbering convention developed by the enterprise or project
Document Owner	Individual authorized to make or direct changes to the document
Document Size	Physical size of document such as 8-1/2 x 11, 3 x 5, roll microfilm, etc
Document Type	Defined by project to describe general content type, such as report, plan, agenda, test procedure
Environmental Requirements	Defines any environmental considerations for storage
File Format	Software application used to create the file, such as Word, PowerPoint, ProE, Adobe Acrobat, etc. Sometimes includes version, such as Word 6.0
File Size	Size of electronic file, usually identified electronically by the system when entered into a repository
File Type	Describes physical characteristics such as hard copy, microfilm, electronic, etc
Enterprise Identifier	Identifies enterprise, department, or project
Related Document ID	Identifies other documents to which the document is related
Related Product ID	Identifies products to which the document is related
Revision Identifier	Unique identifier for data revision or version
Rights	Rights and limitations in access and use of data
Storage Medium	Electronic, file cabinet, card catalog, etc.

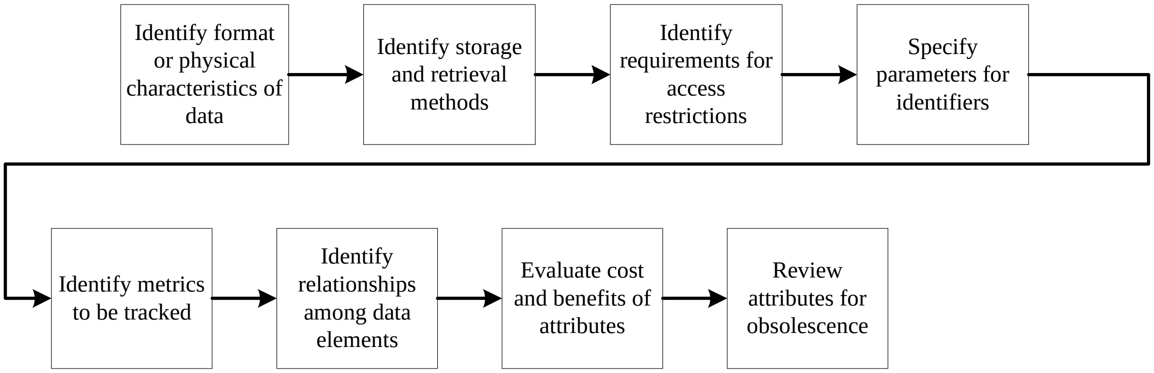
Table 4-1. Metadata Examples.(Continued)

Attribute	Description
Subject	Subject matter of the document or file
Submittal Date	Date of formal submittal to customer, trading partner, supplier, etc
Title	Document title or other descriptive information defining the content of the document or file

4.2 Establish relevant attributes to refer to and define data

Figure 4-3 shows the factors that should be considered when selecting attributes.

Figure 4-3. Develop a Process for Selecting Attributes


Identify format or physical characteristics of data
Identify storage and retrieval methods
Identify requirements for access restrictions
Specify parameters for identifiers
Identify metrics to be tracked
Identify relationships among data elements
Evaluate cost and benefits of attributes
Review attributes for obsolescence

Cataloging, storing, and retrieving data depend on understanding the format of the data to be managed. Electronic files are managed differently than hardcopy paper or microfilm, so the physical characteristics should be taken into consideration when establishing attributes. File format, or the software application used to create or view the file, is relevant for retrieval of electronic files but not for data stored only in hardcopy format. The storage medium and file formats influence readability and reproducibility of the content. Microfiche, for instance, can pose important readability limitations.

The storage medium and file formats also influence the selection of attributes. Selection of attributes to support identification of storage medium is useful in planning for storage facilities. For example, identifying the file size of data to be stored electronically helps identify the resource allocation.

Access to data is restricted based on proprietary issues, security issues, or other limits in data rights. Attributes are selected to identify data that requires special handling or limited access. This protects the enterprise from inadvertent disclosure of data to inappropriate parties. For more information, see Principle 6.

Requirements for tracking and reporting metrics should also be considered when selecting attributes. Metrics are typically used to monitor throughput and ensure that the process is operating as intended, or to ensure that resources are properly allocated. Enterprises that routinely track certain metrics should assist in creating standard attributes to enable the collection of metrics. For more information, see Principle 8.

Identify relationships and their importance in regard to other data elements in order to efficiently identify and manage related objects. It is important to weigh the cost of creating and entering metadata attributes as well as the potential benefits. If users are required to complete numerous metadata entries when placing a document in a repository, it is likely that documents will be entered with missing or erroneous entries, or that documents will not be entered into the repository at all. Potential attributes should be evaluated based on whether there is value added in tracking and locating data. The set of required attributes should be kept as small and simple as possible to allow a user to create simple descriptive records and provide for effective retrieval. Any existing metadata standards should be tailored to meet needs.

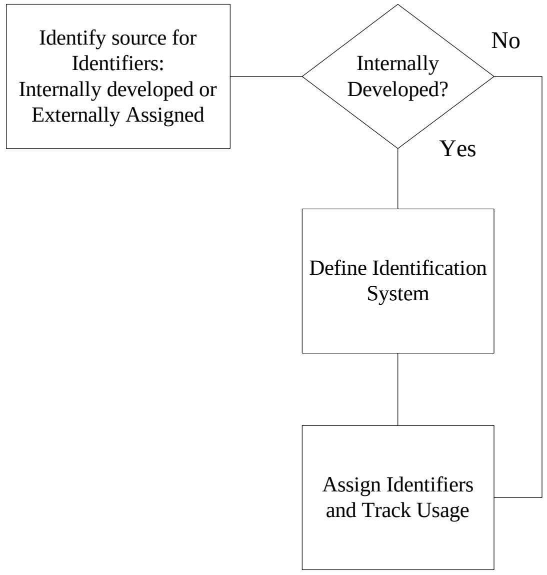
Metadata attributes change over time due to evolving requirements throughout the life cycle. These changes include changes to the data repository (e.g., facility or system upgrades) as well as obsolescence. Part of the overall DM process includes periodic reviews of metadata attributes.

When modifying attributes consider the impact on legacy data. In a large repository, it may not be feasible to update metadata attributes of existing data and it may be necessary to develop translation tables or similar mechanisms.

4.3 Assign identifying information to distinguish similar or related data products from each other

Identifying information is assigned to uniquely identify or name specific data. The identifying information for data commonly consists of a title, unique identifier (e.g., document number), the source of the document, date, and the revision. Figure 4-4 shows the steps included in the assignment of identifiers. The requirements for document identification are discussed in EIA-649, National Consensus Standard for Configuration Management, and EIA-836, Configuration Management Data Exchange and Interoperability.

Figure 4-4. Assign Identifying Information to Distinguish Among Similar Data Products


Identify source for Identifiers: Internally developed or Externally Assigned
Internally Developed?
No
Yes
Define Identification System
Assign Identifiers and Track Usage

Ensure that a unique identifier is needed. Unique identifiers are only assigned to the data that needs to be tracked and controlled to meet on-going needs for the data. The identifier provides an identification method to differentiate between similar documents, and enables consumers to identify the information they need to perform their assigned tasks. It also helps to minimize the delay in retrieving the desired information, and the problems caused by the use of incorrect information.

5.0 Principle: Control Data, Data Products, Data Views, and Metadata Using Approved Change Control Processes

Introduction

This principle provides guidance that will ensure the integrity and timeliness of data, data elements, data structures, and data views by applying the principles of configuration management.

DM and configuration management (CM) are two disciplines critical to the success of any project. They are strongly related and interwoven in their scope, application, and elements. Both are disciplines whose ultimate purpose is assuring the integrity of the products they support. One of the functions of each of these disciplines is to control change or, for some kinds of data, to protect it from change. It should be recognized that not all data requires formal change control or the same level of control—it is a matter of balancing cost and benefits. This principle addresses the body of data for which some level of control is appropriate.

A critical factor to consider is when a data product is ready to be placed under formal data management control. This process infers a transfer of control (or stewardship) from the author or originating IPT to the data management control process.

The data product needs to be in a state of maturity that makes control both meaningful and productive. When judging this level of maturity, the end item condition of the data product must be known and compared to the state of the data at the transfer point. Considerations of the following factors are critical:

1) the format and media are in concern with the end item requirements;

2) the data is accurate and at an appropriate level of completeness;

3) the timing of the transfer is appropriate to the data product’s end use (too early is just as critical as too late);

    a. too early often imposes unnecessary control when it is not yet appropriate;

    b. too late can cause time problems with the end use.

4) the data product has been reviewed by an appropriate level of authority (e.g., engineering manager, integrated product/process team lead). DM receives the control-ready product from appropriate predetermined sources.

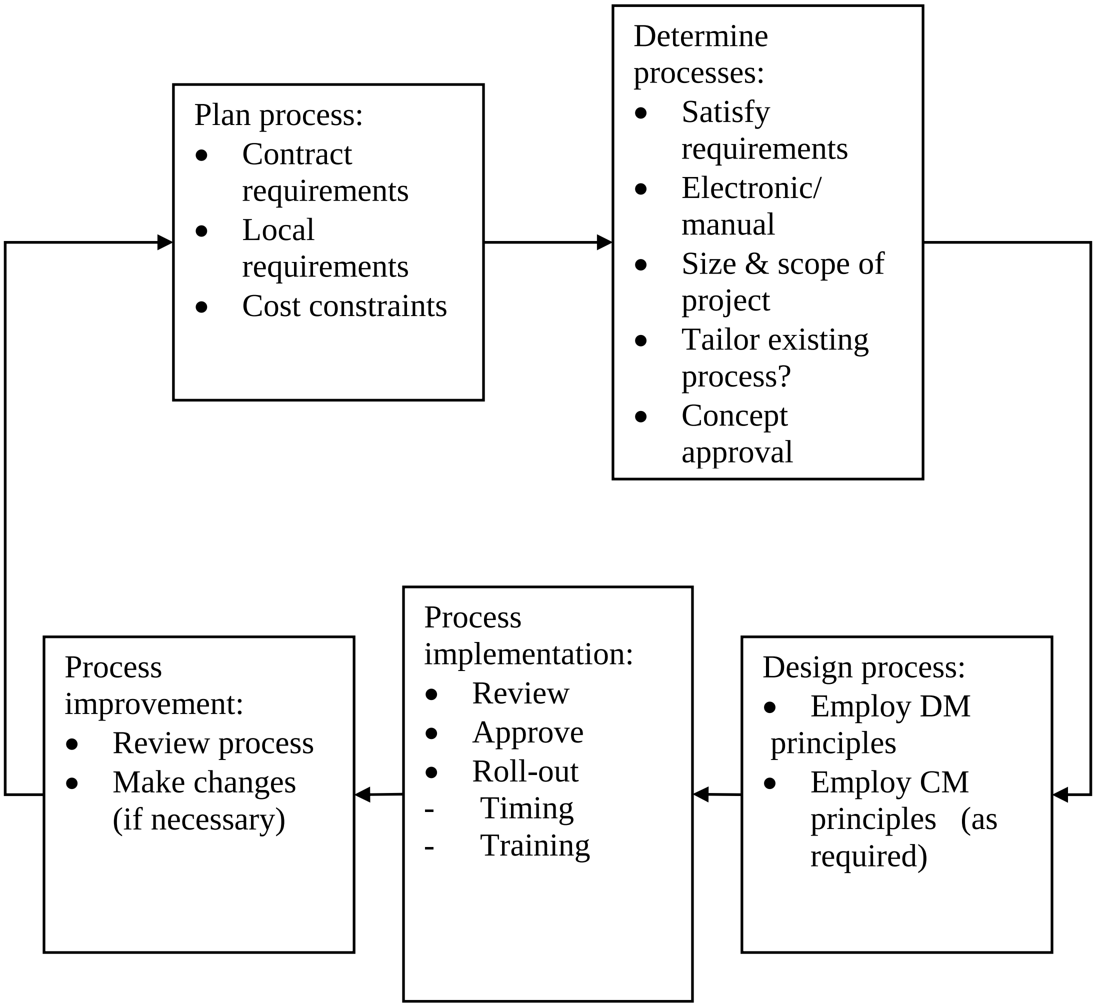
In large programs, it may be appropriate to develop a formal process to cover this transfer as part of the general change control process.

The change control functions and principles defined in EIA-649 *National Consensus Standard for Configuration Management* are appropriate for DM. For that reason, what is discussed here borrows heavily from CM and describes how the change control process applies to DM.

The levels of control, which can be formal or informal, are defined by the requirements of a project. It is important that the levels of control for data be identified and communicated at the beginning of a project. (Where DM is provided as matrixed support, such identification should be accomplished jointly by project management and DM.) The following discussion applies to data under formal change control. Much of the process definition below specifies a formal and thorough methodology that can be tailored as needed.

Figure 5-1 summarizes the steps needed to provide this control.

Figure 5-1. Establishing Control



5.1 Control the integrity of data, data elements, data structures, and data views

DM ensures that data products satisfy requirements. Doing so, in part, requires that the integrity of the data products (see Principle 3 for definition) and associated data elements is maintained using a consistent change control process and that changes are approved by an authorized approval authority.

Data retains value commensurate with its accuracy, timeliness, and relevancy to the business. The value added by the DM processes is the preservation of this worth. Business revenues are dependent, in part, on the compliance of the data to requirements and intangibly on customer satisfaction. Relevancy of the data will vary throughout the project. Data can vary in maturity and therefore importance. Data within a project undergoes continuous development—working data, mature data, released data, submitted data, approved data, archived data, and possibly delivered data. At each of these stages, data possesses different levels of value and importance. Recognition of these relative stages is important with regard to the level of control that is imposed on the data. Some data are placed under change control at change control inception, some are considered for formal control at a later date, and some may never be brought under any formal control.

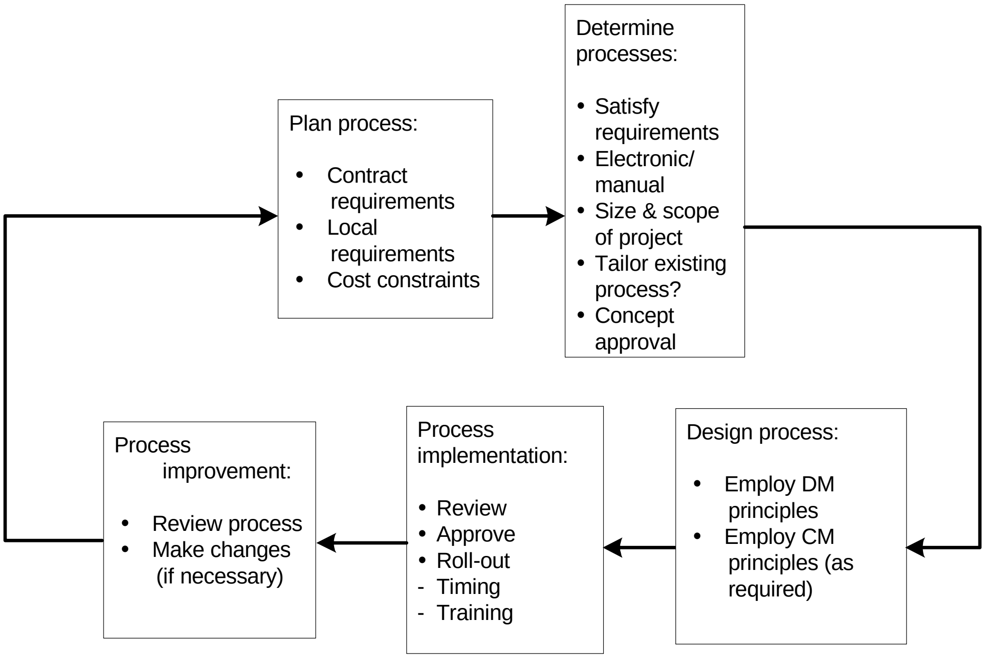
A given project may also require different levels of control dependent on customer imposed requirements, enterprise requirements, or maturity of the project. Decisions need to be made early in the process outlining which data elements require configuration control. It is also important to consider at what point in time control needs to be imposed and what level of control is necessary. The over application of controls is just as in-appropriate as too little control and the application of control too soon is as inappropriate as too late. An appropriate change control process ensures efficient and effective request for change processing without impeding design development, production, or operational readiness.

5.1.1 Establish a change control process that imposes the appropriate level of review and approval

Control of data within a project is as important as is control of the product’s design.

DM applies the doctrine of change control for data elements that require control. Figure 5-2 represents a basic change control process, which can be tailored to meet the particular requirements of the project.

Figure 5-2. Establish consistent change control process



5.1.2 Provide a systematic review of proposed changes within the change process

One of the most important reasons for an organized change control process is the thorough review that is applied to a proposed change. This process provides critical information for the status accounting of the change, the change history, and the ultimate disposition of the change. While it is not always necessary to exactly mimic the CM change process, the basic principles of good CM should be adopted.

The change authority plays a vital role in the configuration change control process. This authority evaluates requests for change based on information developed as a result of an administrative and technical review to approve (and implement) the change, disapprove the change, defer the change, or return it to the originator for rework. As needed this change authority may take the form of a single project designated individual or be implemented as a more formal Change Control Board (CCB).

The change authority should evaluate the:

- ◆ validity of the proposed change;
- ◆ interface effect on other data under control;
- ◆ impact on other project areas other than the area which is recommending the change;
- ◆ effect on established delivery schedules;

- life cycle cost effectiveness and the availability of funds; and
- other factors pertinent to the project.

The configuration change control process begins with the preparation of a request for change. The form and format of this request, usually a formalized document that exists either on paper or electronically, is dependent on the project. While this is not strictly governed by the configuration management requirements for the project, it may be advantageous to implement a format tailored after the project's CM request for change. The formality of the process depends on the requirements of the project.

After recording receipt of the request for change, an administrative review of the request for change and associated supporting documentation (making up the request for change package) may be required to determine if the request for change is acceptable for processing.

The project manager may designate a subject matter expert as the sponsor for the request for change and be responsible for conducting a thorough review. The key is selecting the correct reviewers for the requested change. Ultimately, they should represent those affected by the change.

This review provides the needed information for the change authority, which ultimately dispositions the change (approve, disapprove, defer, or return the change for rework), to make a reasonable, economic, and informed decision.

This kind of process, in the teaming environment under which many projects are being conducted, also provides team "buy-in" for the participants. It allows the team to actively play a role in both the technical and programmatic content of the project.

A structured means for disseminating the change package to the selected reviewer(s) may be important. Consideration should be given to electronic workflow, but paper distribution may be more realistic in certain circumstances such as small projects. The change authority and/or the technical manager has the responsibility to ensure that the appropriate person(s) sees and reviews each change. This responsibility normally may be delegated.

5.1.3 Determine the impact of change to include associated products, data, data elements, data structures, and data views

A significant benefit accrued in applying CM techniques to DM is in the management and control of data through providing a process for determining the impact of change.

There are several considerations relative to the execution of this process. As noted above, the selection of the reviewer(s) is critical to the successful accomplishment of the impact assessment. Reviewer(s) should be subject matter experts, competent in understanding the technical area(s) associated with the proposed change.

It is often helpful to provide a set of criteria by which the assessment should be made. These criteria, which typically address, cost, schedule, and performance, may be simple or as technically complex as needed.

What areas are impacted by the change and the extent and significance of the impact needs to be determined. Some of the areas with regard to changes to a data element that could have a significant influence are:

- Changes to requirements
- Changes to specifications
- Changes to customer furnished information
- Changes in supplier data
- Contract changes
- Changes in de-facto performance
- Changes that would impact cost or schedule.

Any of these are likely to be significant. Obtaining a precise statement of the impact and its potential consequences may prove more elusive, reinforcing the need for effective selection of reviewers.

The review process entails ensuring balance between thoroughness and the timeliness of the review. Once this review is conducted and any impacts (and their potential consequences) are determined, create a concise written statement and forward to the designated approval authority for action.

5.1.4 Gain approval or disapproval of changes to data, data elements, data structures, and data views (data products) by a designated approval authority

After completion of the processes for conducting reviews of changes (enabler 5.1.2 and 5.1.3), the process for determining change disposition follows logically. The formality of this process (e.g. a structured change control board (CCB), or a single person authority) is dependent on the size, scope, and contractual requirements placed on the DM process.

The change authority dispositions the change in one of three ways:

- The first is to approve the change and forward the change to the proper authority for implementation and ultimate close out.

- The second is to disapprove the change. In this case, the disapproval is noted in the CCB record and the change originator/sponsor is notified.
- The third choice is to defer the change. This could occur for a variety of reasons. For example, additional information may be required to make an informed decision, there may be a flaw in the supporting documentation submitted, or there are unresolved funding issues. The change is returned to the originator or sponsor for further correction, amplification, or clarification

Regardless of the disposition, a notification is prepared and issued to those impacted by or otherwise interested in the disposition. The form and formality of the notification vary from circumstance to circumstance. It is good practice to document not only the disposition but also the position (for or against with reasoning) of each party to the decision. The notification also provides essential information for updating status accounting records, the “official” records of change dispositions.

5.2 Establish and maintain a status accounting process, reporting tool and mechanism

A unique change control number is assigned to each request for change and entered into a change status accounting tracking system. The tracking system should include, as a minimum and as applicable, the date of request for change, request for change control number, priority, classification (if required), originator, request for change title, affected data item(s), date of receipt by the change authority, CCB meeting date (if there is a CCB otherwise disposition date), request for change approval, disapproval, and deferral status.

In addition to tracking change history, data sources and other project related information should also be accounted for and tracked. This data includes, but is not limited to, items delivered to the contractor by subcontractors. Examples of the type of metadata that may be recorded within the status accounting database are listed below:

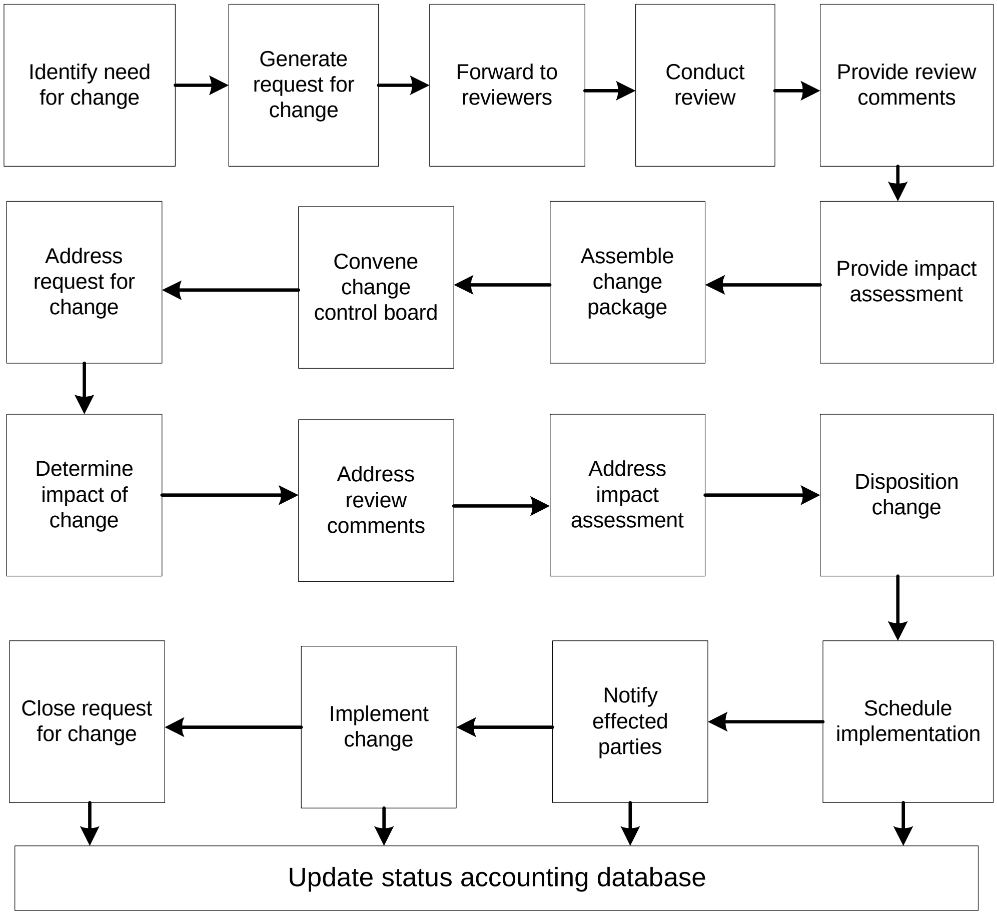
- Identification of data item
- Source of the data item
- Date delivered
- Contract required date
- Contract (or subcontract) reference
- Format of item (hardcopy, disk, CD, file)
- Destination (who received a copy or who was notified of receipt)

- Storage location of original
- Security classification (if applicable)
- Export/import information (if applicable)

The establishment of a status accounting mechanism is as important as the establishment of a status accounting process. This mechanism consists of the tools that support and complement the status accounting process.

The process for maintaining metadata, pictured below in Figure 5-3, defines how the metadata is generated, gathered and introduced into the status accounting database. The process largely defines the information that is stored in the database and sets the database requirements.

Figure 5-3. Metadata is maintained for project use in a status accounting database



Project requirements, availability of technical database expertise at affordable cost, and reuse opportunities drive the architecture of the database. These elements have a bearing on whether the database is stand-alone, network centric, Internet centric (web based), a simple flat file, relational, homegrown, or a commercial product. Reuse of existing status accounting databases should be considered. Reuse can save time and money.

Table 5-1, below, lists examples of the types of functions that should be considered.

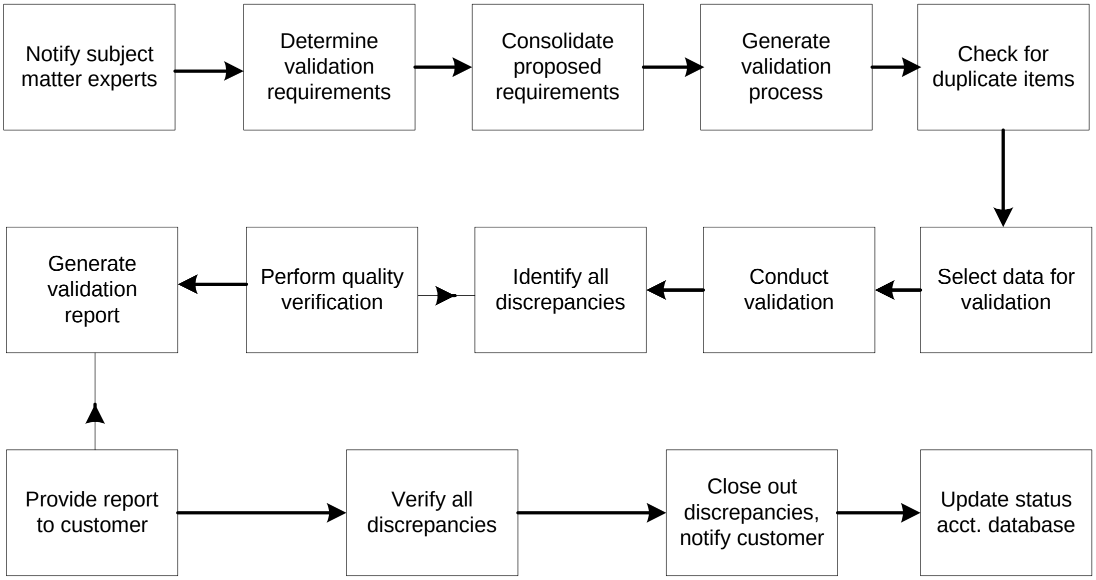
Table 5-1 Example Elements of Database Functionality

Administrative Functions	User Interfaces	Data Relationships and Functions
Database security User permissions Administrative overrides Administrative rights	Screen layouts Sign-on screens View screens Input screens Output screens Reports Ad hoc User generated	Data fields Data formats Data relationships Metadata requirements Metadata definitions Search mechanisms Search criteria

5.3 Establish and Maintain an Internal Validation Mechanism

There are several key validations required within the DM process (Figure 5-4). These relate to the status accounting process itself (as discussed in Enabler 5-2), the data stored in a repository (if used), and the data contained in the change status accounting database.

Figure 5-4. Status accounting data and stored data require validation to ensure integrity


Notify subject matter experts
Determine validation requirements
Consolidate proposed requirements
Generate validation process
Check for duplicate items
Select data for validation
Conduct validation
Identify all discrepancies
Perform quality verification
Generate validation report
Provide report to customer
Verify all discrepancies
Close out discrepancies, notify customer
Update status acct. database

Validation of processes can be done by anyone, but are frequently, best accomplished by the element of the enterprise charged with DM. There are several reasons for this:

- The self-validation process helps to build lessons learned in a more meaningful fashion when conducted by the process owners

- ◆ The enterprise element charged with DM is most familiar with the DM processes

- ◆ Correction of deficiencies is normally more expedient when responsibility is centralized

- ◆ Self-examinations can be conducted prior to, or in conjunction with, formal quality validations thus expediting and adding integrity to a formal validation.

If a repository is available for a project to store significant data, the integrity of that data needs to be maintained. A self-validation can assess the completeness and uniqueness of the data items within the repository, adequacy of metadata, and similar essential characteristics. The validation should also ensure that each data item is worthy of storage and retrieval. Validations are a convenient time to reassess the continued value of data and to dispose or archive it if appropriate. The validation can also examine the adequacy of protections for intellectual property.

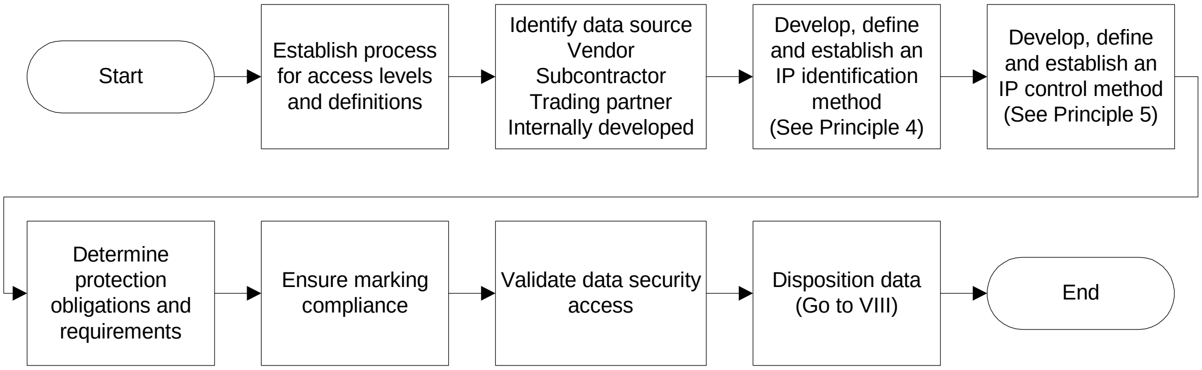
6.0 Principle: Establish and Maintain an Identification Process for Intellectual Property, Proprietary, and Competition-sensitive Data

Introduction

Intellectual property (IP) is a term used to describe real but intangible assets, embodied in such items as patents, copyrights, trademarks, and trade secrets. IP is at the center of an enterprise's competitive position and ultimately contributes to financial success. For this reason, protection of IP is necessary to maintain an enterprise's competitiveness. In many cases, it is also necessary to comply with legal obligations to suppliers and customers.

IP assets come from a variety of sources. In addition to internally developed data, IP is received from suppliers, subcontractors, and trading partners. All of this data is identified and tracked for protection based on data rights. The process flow diagram at Figure 6-1 illustrates, at a relatively high level, the management process for IP and its relationship with other DM principles. Lower level processes are delineated under the enablers.

Figure 6-1. Principle 6 flow diagram


Start
Establish process for access levels and definitions
Identify data source
Vendor
Subcontractor
Trading partner
Internally developed
Develop, define and establish an IP identification method (See Principle 4)
Develop, define and establish an IP control method (See Principle 5)
Determine protection obligations and requirements
Ensure marking compliance
Validate data security access
Disposition data (Go to VIII)
End

The rights obtained from the provider through documented agreements, such as statements of work, license agreements and contract negotiations, determine how IP is managed. The ability to deliver the information obtained from a supplier to a third party, as well as the obligations and requirements to limit access and use are also defined in these documents.

There are several varieties of proprietary data. Some examples include general business information (available to the general public), information to be used only within the enterprise (internal use only), information developed by the enterprise that has monetary value (enterprise proprietary information), and enterprise-developed information that has been officially registered with a legal authority (registered proprietary information). Data not specifically typed but that might be construed as providing an enterprise advantage within industry is considered to be competition sensitive data. Examples include best practice information, proposal information and tools implementations.

Enterprise policies for IP management provide a standardized way to type, mark, and identify; control and track ownership; manage rights to use and sell; control access; distribute; and disposition IP within the enterprise. Management of IP requires the following:

- Identify items that need to be protected and tracked.
- Store items in a protected environment or repository with limited access.
- Control access to and distribution of data dependent on data type and source.
- Provide security as required by agreements and legal obligations.

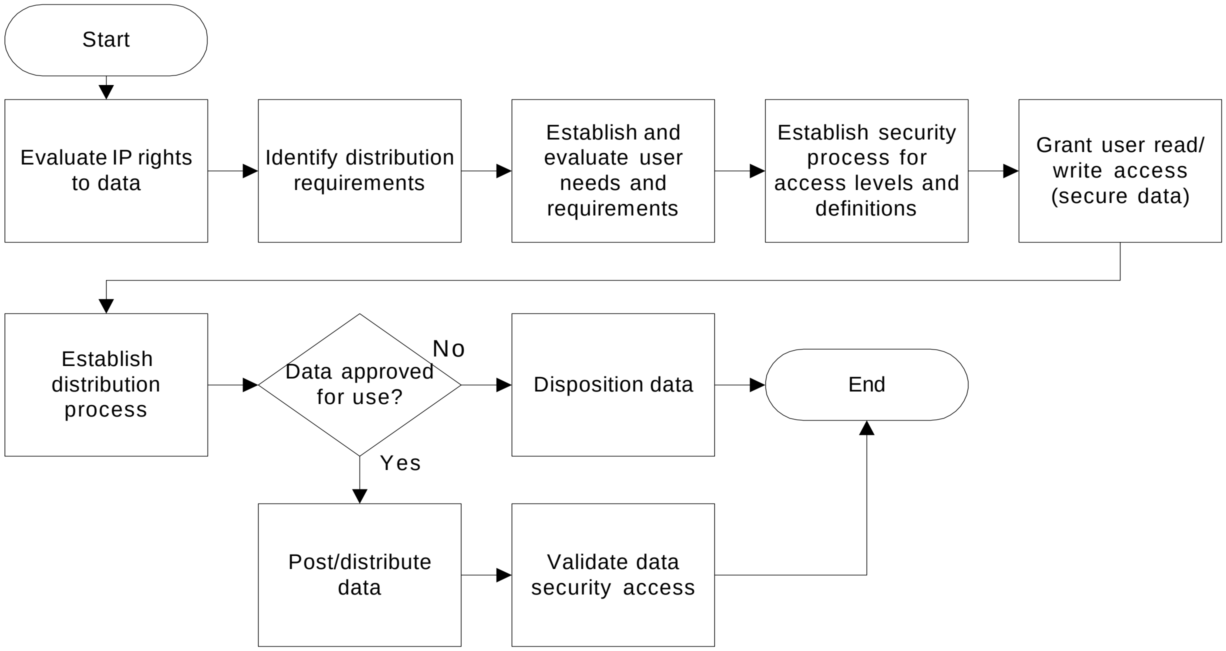
In some instances there is a need to sell, purchase, or license IP for purposes such as establishing standards, developing business relationships, creating new and larger markets, and realizing strategic goals. These transfers should take place under stipulated conditions and careful control to protect the data rights of the data originators and owners. Regardless of the type or source of IP, manage it as an asset of the enterprise. Failure to successfully manage IP can have personal, enterprise, national, and international implications.

The three enablers discussed below provide a basis for flexible and tailorable IP management.

6.1 Establish and maintain a process for data access and distribution

To effectively manage IP, a method for managing data access and distribution needs to be in place. Access and distribution of data is critical to the protection of data rights. The process to support Enabler 6-1 is delineated in Figure 6-2 and defined in the sub-enablers.

Figure 6-2 Process for managing data access to intellectual property, proprietary and competition sensitive data


Start
Evaluate IP rights to data
Identify distribution requirements
Establish and evaluate user needs and requirements
Establish security process for access levels and definitions
Grant user read/ write access (secure data)
Establish distribution process
Data approved for use?
No
Yes
Disposition data
End
Post/distribute data
Validate data security access

Review types and varieties of IP that are to be addressed, and create a method of controlling access and distribution. In a manual environment, IP may be managed through limited access facilities such as locked files or areas. In an electronic environment, electronic methods such as organizational and role based access control are generally required to limit the electronic access of data.

Where enterprise policies and procedures do not exist, document the access constraints for the various types and varieties of data. Once this process is defined, it can be applied at all levels of the enterprise to address all sources of data.

6.1.1 Define access requirements.

Review documented agreements to verify that access rights granted support the intended use by the enterprise. If rights to data are not authorized, evaluate data to determine the currency of the business need within the enterprise. For those items no longer current, or where a need has expired, schedule disposition in accordance with the enterprise or department retention schedules and authorization for the intended use. Contract negotiations, subcontract negotiations, licensing agreements, royalty payments, and similar legal documentation define the rights to data. Data is not distributed or used until the legal right to do so has been verified.

Review contractual requirements and legal rights and responsibilities prior to providing access or distribution of data to trading partners, subcontractors, suppliers and customers. If access is authorized through a documented agreement, verify the type of data needed by the user, as well as the distribution method and access level required to support the user's needs.

When interchange data environments are required or used, define the levels of and definitions for access rights and establish the mechanism for authorizing that access.

6.1.2 Ensure entitlement to access and use of data is validated and documented by the proper authority.

Ensure that the owner of the data (enterprise or individual representative) has authorized or validated the user's need for access. Maintain records of access rights granted, distribution methods, and account authorizations for verification and validation purposes. Review these records regularly to ensure that data remains secure and access rights are current.

Before data is distributed, validate the information is approved or authorized for use. If not authorized, evaluate data to determine the reasons. Data is distributed or used only after authorization by a review authority. If authorized, distribute the data in accordance with the defined process and the user rights. This distribution may be performed manually, through email, by means of an electronic interchange data environment or any other method that meets the requirements of the process.

Validate the security of the data on a periodic basis as part of an audit activity. Failure to secure the data, or allow unauthorized use can result in monetary fines or other penalties for both enterprises and individuals. Information to be considered for audit includes:

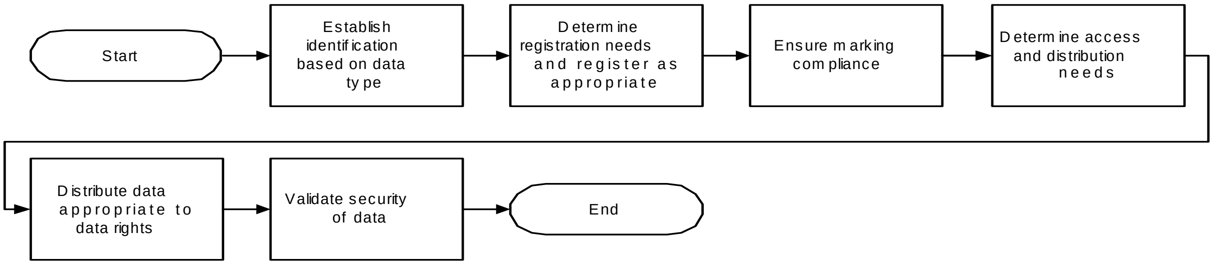
- IP is properly identified by type and source
- IP is properly marked and tracked
- Patents exist where appropriate
- Copyrights are registered where appropriate
- IP rights granted are current and followed
- Import and export evidence exists where appropriate
- IP user access rights are reviewed
- IP distribution is reviewed
- IP disposition schedules and methods are followed

As data or information is disposed of or considered of no ongoing value, it is handled in accordance with Principle 7.

6.2 Establish and maintain an identification process for IP, proprietary and competition sensitive data

The process to support Enabler 6.2 is delineated in Figure 6-3 and defined in the text

Figure 6-3 Process for identification, tracking and protecting intellectual property, proprietary and competition sensitive data


Start
Establish identification based on data type
Determine registration needs and register as appropriate
Ensure marking compliance
Determine access and distribution needs
Distribute data appropriate to data rights
Validate security of data
End

6.2.1 Distinguish contractually deliverable data

At the enterprise level, policies are documented to define the process for distinguishing and managing IP from other data. Prior to design of a product, a process should be used to determine the data requirements for development of the product. When the product contains information that is deliverable to a customer, an evaluation occurs regarding IP and the legal responsibility to protect it, regardless of source. Negotiations occur with potential suppliers to establish an agreement to use or resell the data. The outcome of those negotiations is documented and forms the basis for what can legally be contracted to another party.

When a potential or contracted customer requests delivery of a product where legal rights to deliver to a third party do not exist, resolution must be reached with either the customer or the supplier, usually through negotiations. Even if data is not contractually deliverable, identify and secure it to protect the rights of the provider. Data is used in accordance with the rights granted by the provider, through contracts, subcontracts, license agreements, or other legal documentation. Always evaluate the obligations and legal responsibility for data protection.

6.2.2 Establish and maintain identification methods

Enterprise processes should exist for identification methods that address data within the enterprise. Data and data requirements are defined and identified with unique identifiers, as delineated in Principle 4. At the project level, the identification methods should be documented if they deviate from an enterprise policy or an enterprise policy does not exist. This includes another layer of identification for IP to ensure the data is handled in accordance with IP policies and legal obligations. Internally generated data can then be easily identified and typed for protection.

Evaluate internally developed and funded data to determine if a patent, trademark or copyright is feasible in the business environment. Register U.S. patents and trademarks with the United States Patent and Trademark Office (http://www.uspto.gov/). In the United States, copyrights are automatic. In some instances (e.g., protection of data rights in a global market) it is advantageous to register a copyright. Register copyrights with the United States Copyright Office (http://www.copyright.gov/).

Review data obtained from an external source to determine if it is registered IP. Verify documented rights to data prior to use to ensure that appropriate protection of the data occurs.

An enterprise policy or process for import and export control should address the legal obligations for importing and exporting data outside the country of origin. Review data prior to export to ensure compliance with enterprise processes and legal obligations. Obtain additional information and assistance for United States policies through the Bureau of Export Administration, U.S. Department of Commerce (http://www.bxa.doc.gov/bxahelp.htm).

6.2.3 Establish and maintain tracking mechanisms for identification of data

Identification and control of data are addressed in Principles 4 and 5 respectively. There are however some additional elements of metadata that need to be tracked for IP. Tracking mechanisms and evidence are fundamental for the following items:

- Distribution is appropriate to rights granted
- Appropriate maintenance of data is possible
- Configuration status of IP is maintained
- Import and export forms are maintained
- Licensed quantities and locations are tracked
- Appropriate rights are negotiated or granted for updated items
- Distribution (list of names, addresses, restrictions, etc.) is appropriate to rights granted

6.2.4 Ensure compliance with marking conventions and requirements

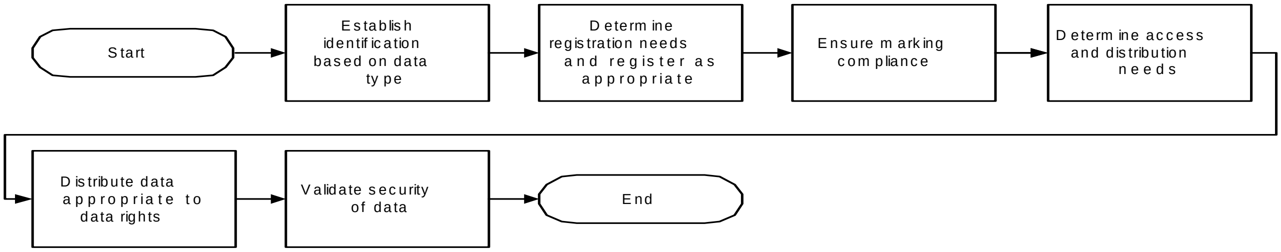
Once IP has been identified, it should be marked appropriate to its type or variety. When proprietary information or IP is provided to the U.S. government, mark it using government notices or legends. See the glossary for definitions and marking references. Disclosure of proprietary information in any other context requires a nondisclosure agreement or other legally binding type of documented agreement. These legal agreements restrict the use and disclosure of the information being shared.

If the information is provided to a non-U.S. citizen, export control requirements need to be satisfied prior to disclosure. This includes printed, electronic or verbal disclosure of information.

6.3 Establish and maintain an effective data control process

The process to support Enabler 6.3 is delineated in Figure 6-4 and defined in the text of the sub-enablers.

Figure 6-4 Process for controlling, tracking, and protecting intellectual property, proprietary and competition sensitive data


Start
Establish identification based on data type
Determine registration needs and register as appropriate
Ensure marking compliance
Determine access and distribution needs
Distribute data appropriate to data rights
Validate security of data
End

6.3.1 Establish and maintain control methods

Processes should exist for control methods that address data within the enterprise. Data is controlled so that changes to data are reviewed and authorized by the appropriate personnel and results are provided on a need to know basis. Details of the change process are defined in Principle 5. For IP, control systems are different based on owners and use of data and include appropriate approval mechanisms and updated documented agreements for data rights. This provides another layer of control for IP to ensure the data is handled in accordance with IP policies and legal obligations.

Evaluate internally developed and funded data to assess the impact of the change. Determine if a patent, trademark or copyright requires updating or re-registration as a result of changes. If re-registration is appropriate, register U.S. patents and trademarks with the United States Patent and Trademark Office (http://www.uspto.gov/). Register new copyrights with the United States Copyright Office (http://www.copyright.gov/).

Changes to the data may or may not impact documented agreements for data rights. Review documented agreements to assess the impact of the change. Areas of particular concern exist where the right to use the updated item is not part of the original agreement. In those instances, new agreements must be negotiated. Establish review and disposition methods for IP changes based on the business needs.

6.3.2 Establish mechanisms for tracking and determining status of data

The mechanism for tracking IP continues when tracking changes to IP. When changes occur, the ability to trace users of IP data assists in determining the distribution for approved updates. As with other IP issues, changes need to be tracked and the data rights reviewed before distribution.

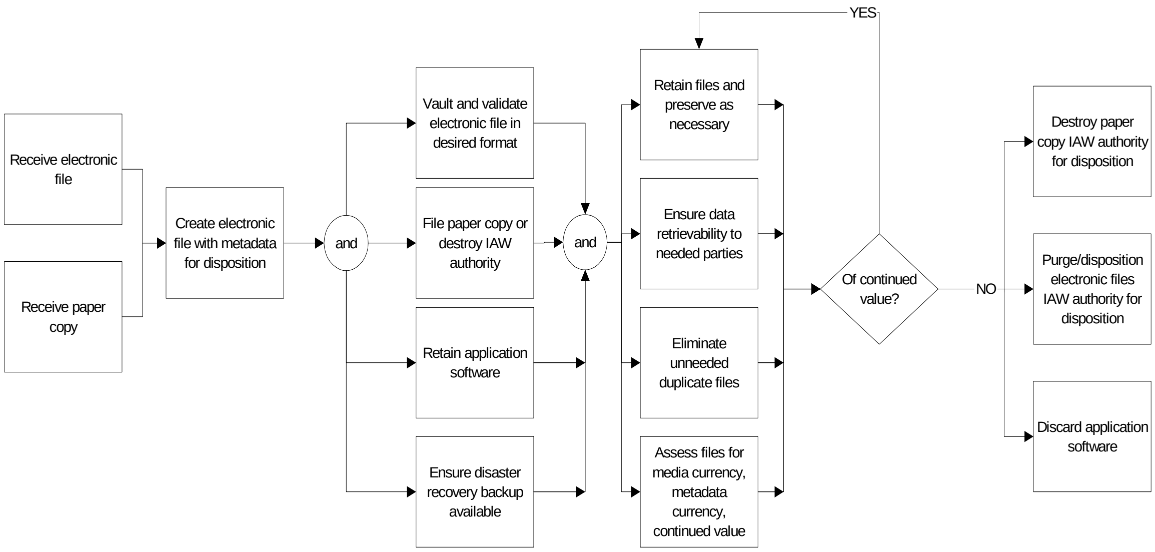
At some point, rights to data expire or are no longer of value to the enterprise. If there is an enterprise retention policy, or a legal obligation to maintain the data, retain the IP information, including the documented agreements that define the data rights. Principle 7 provides guidelines for data retention and storage.

7.0 Principle: Retain Data Commensurate With Value

Introduction

The purpose of this principle is to delineate methods for ensuring adequate retention and preservation of data assets that are of value to the enterprise and effectively disposing of data assets that are no longer of value. Figure 7-1 illustrates the overall process.

Figure 7-1 Planning decision tree for data of sustained value


Receive electronic file
Receive paper copy
Create electronic file with metadata for disposition
and
Vault and validate electronic file in desired format
File paper copy or destroy IAW authority
Retain application software
Ensure disaster recovery backup available
and
YES
Retain files and preserve as necessary
Ensure data retrievability to needed parties
Eliminate unneeded duplicate files
Assess files for media currency, metadata currency, continued value
Of continued value?
NO
Destroy paper copy IAW authority for disposition
Purge/disposition electronic files IAW authority for disposition
Discard application software

Any data assets that are of potential business, project, or operational value should be retained until their value is depleted. Data of sustained value to the enterprise should be retained and evaluated on an ongoing basis as notionally shown in Figure 7-1. The enablers described herein provide a basis for enterprise behavior that ensures data is retained commensurate with its potential entrepreneurial, legal, contractual, and other worth to the enterprise and customer. Quality, accurate, and up-to-date data aids in critical business decisions. Timeliness of the decision-making process with value added data increases competitive advantage. The right data at the right time is cost effective and reduces lead-time to decision making and business processes. Non-value added data should be removed from the enterprise’s inventory.

7.1 Plan to ensure data is available when later needed

Data assets should, upon their creation and initial storage, have planned retention requirements identified and documented. Such business processes would cover archive formats, frequency of reviews, purge planning, disposition funding, and related activities. Clearly defined methods for data retention help assure the data will be available when and if needed. One such method is to develop an enterprise policy on data retention. Considerations for such a policy are detailed in Enabler 7-2.

Effective control of data is best accomplished through defined process ownership and accountability. To ensure proper planning for eventual disposition of data assets, identify an appropriate data steward for planning disposition date(s) early in the data life cycle. These stewards should be trained in the organization’s retention and disposal processes. They manage physical custody of their assets to ensure, as a minimum, that electronic data with wide applicability is stored in a retrievable storage media; that inactive data is archived, that hard copies are protected, and that data are identified and catalogued for retrievability. They ensure that movement of data assets and their backups are known to them. Further, they ensure planning is in place to control data assets near the end of their useful lives such that the enterprise does not store items that no longer retain value.

Data stewards ensure that data is stored at authorized locations. They maintain backup copies at locations separate from the masters for best disaster recovery practice. They assure backup copies are not maintained or stored at unapproved locations, such as personal residences. They make certain that the physical whereabouts of data assets are known and easily retrievable by those in need.

Maintaining control of the repository and the associated processes or data holdings is a DM function. Pay special attention to changes of stewardship. These changes can result from a number of factors including::

- Enterprise charter changes, corporate mergers, or corporate divestitures.
- Personnel changes resulting from changes in position responsibilities, retirement, attrition, or similar actions.
- Changes in management during the data life cycle - for example, from an on-site location in the early life cycle to an off-site location when data is archived

To ensure proper planning for storage in protected environments, investigate cost and facility availability to meet upcoming needs for data protection environments. Paper files may be somewhat protected from fire, for example, with inert gas systems. Protection of compact disks (CDs) and other heat and light sensitive items in warehouse environments may be ensured with effective cooling and humidity control. Protection of electronic files from viruses may be enhanced through an effective virus protection program.

An essential element of preservation planning is to ensure planning for adequate protection of data against potential disaster commensurate with the value of the data. Review risk areas where singular versions of data could be lost in the event of a disaster and develop plans to mitigate such risks. Mitigation may occur by developing duplicate copies or scanning paper documents to provide backup as well as by maintaining electronic versions (tapes, CD, etc.) of files and associated software. Storage of the duplicated data at a separate locale or even region of the country is prudent to overcome risk of local disasters such as hurricane, tornado, or earthquake. Additionally, effective planning for disaster recovery in the event of a crisis, such as facility fire, server crash, flood, etc. mitigates the potential of data asset loss and implicit inability to retrieve such assets. For some enterprises, a disaster has been cause to cease operations as no disaster planning was incorporated in the enterprise's basic business planning.

7.2 Maintain data assets and an index of enterprise data assets

Data stewards are to ensure that accurate and complete records are identified, available for view by those with a valid business need, controlled, retained, protected, and subsequently disposed of in accordance with the requirements set forth for the data assets.

Retention dates are the latest of dates set by law, by local policy, by ascertained potential need to the enterprise, or potential need to the customer. (Society of Aerospace Engineers Standard AS9034, “Process Standard for the Storage, Retrieval, and Use of Three Dimensional Type Design Data” provides related process descriptions and related information for aerospace-specific application.)

To assure data assets are well identified, use appropriate identifying data pertinent to the items – metadata - to enable its retrievability. Such metadata may include date, contract number, author, title, general topic key words, owning enterprise document number, document serialization, default retention date and data steward. Effectively identified data assets enable timely retrieval and destruction as appropriate.

To protect data assets from unauthorized viewing, place security requirements on data assets as appropriate (Principle 6). Enterprise data assets may be of a proprietary, government classified, or other sensitive nature that warrant protection against unauthorized viewing. In a paper environment, protection may be assured through physical security. In an electronic environment, protection may be assured through segregation of data by server, by firewall, by password, and similar means.

When evaluating how to store data, the following questions may assist in decision making:

- Will the data ever be used as direct source information in the creation of new material?

- How long will this electronic data likely have value to the enterprise? If a long duration, electronic data migration might be considered more heavily than if a short duration.

- How long will this paper data likely have value to the enterprise and how frequently might it be accessed? If a long duration, electronic conversion may be warranted

- How difficult will it be to retain compatible (and potentially obsolescent) equipment and software in order to provide for retrievability and readability of the files over their anticipated lives?

To ensure data assets are readable in future years, storage of data in neutral formats is preferable if the data is not anticipated to be manipulated later or used in the creation of new material. Neutral formats work well for data that is only for reference and not a source for future work.

To ensure data assets are readable in native formats for later manipulation, retain necessary computer resources to recall and install, view, revise, print images or refresh to newer technology. Alternatively, plan to periodically migrate data assets to current software applications and hardware formats for continued currency and availability for retrieval. The decision process to retain obsolete computer resources or to refresh to newer technology is a business case, driven by economics pertinent to the predicted likelihood of data reuse. In the case of retaining obsolete resources, this process may involve extending date expiration-sensitive licenses or arranging software support into out-years. By retaining computer resources, the enterprise ensures pertinent records are viewable and editable upon later need. In the case of migrating data assets to current formats, periodic migration to current software with its correlated validation for accuracy occurs. Failure to continue migration to current formats can be costly to the enterprise. It is time consuming and often expensive to locate a supplier or enterprise with the capability to migrate to current technology media.

Hard copy data, while not as susceptible to some of the issues electronic media face, have their own vulnerabilities. Many inks fade over time, degrading legibility of data on Mylar. Copied material is somewhat prone to ink lifting, particularly if exposed to heat, and legibility degrades with each succeeding generation of copies.

To mitigate electronic data loss due to shelf-life limitations of storage media, perform periodic refreshes and data validation. Consider migration to state-of-the-art software formats and storage media. Some CDs, whose shelf life has been viewed as very long term, are now rendered useless because of the acid content in the inks used to print labels. Inventories of data assets maintained in current software formats allow for fast retrievability and easy readability. Maintaining a data refresh schedule facilitates proper attention to electronic data before media shelf lives are exhausted.

Table 7-1 Representative Refresh and Migration Intervals

MEDIA	REFRESH/MIGRATION	ANTICIPATED LIFE
File server	Backup daily	5 years
Disks, compact (CDs and their variations)	5 (10) years	25- 100 years
Disk (Diskette), 8-inch	Migrate to current formats	1-3 years
Disk (Diskette), 5 ¼-inch	Migrate to current formats	1-3 years
Disk (Diskette), 3 ½-inch	Migrate to current formats	1-3 years
Tape, cassette (magnetic tape, cartridge)	Migrate to current formats	10-30 years
Tape, magnetic (magnetic tape, reel)	Every 5 years migrate to current formats	10-30 years
Magnetic tape, compact	Every 5 years migrate to current formats	10-30 years
Removable disk, ZIP/JAZ	Every 5 year, migrate to current formats	5-15 years
Microfilm/microfiche	Consider for migration	40 years
PC hard drives	Back up to other locations	5 years

Table 7-1 provides representative refresh and migration intervals. The data in this table is gathered from experiential as well as supplier sources. It is for general information only.

To minimize retention of duplicate media, review duplicate data assets and determine the need to store multiple media beyond that needed for disaster recovery. When records exist in more than one physical medium without specific need, unnecessary duplication wastes physical space (in a paper environment) and digital storage (in an electronic environment). Additionally, and perhaps most importantly, retaining multiple copies exacerbates configuration control problems. On the other hand, an enterprise need to be cognizant of business needs or specific legal and contractual requirements that may mandate that multiple media be retained.

Providing special protection and backup for vital data records provides greater likelihood of retrieval for items of key legal or strategic significance (Principle 6). Such protection may include greater physical protection such as a rodent-proof container, fireproof vault, temperature and humidity controlled area, etc. It can also mean added security (access restriction) protection. Data with the greatest potential loss to the enterprise may be a candidate for special attention. Added measures to prevent data getting to the possession of unauthorized personnel need to be carefully weighed against the inherent advantages to ensuring broad availability of important data. In an acquisition and merger environment, proper attention to the rights of the current and prior data owners is paramount to doing business. Special requirements may exist for retention and/or disposition relative to previous or new corporate ownership. For example, data pertaining to a certain project, with certain tax application, created during a certain period, etc. may have to be reviewed by another enterprise as part of the disposal process.

7.3 Assess the current and potential future value of the enterprises' data holdings

Periodically assess data to make sure the enterprise does not retain non-value added data. As a minimum, data should be re-evaluated at the time originally designated for disposition, its original default retention date. Re-evaluation considers the latest of dates set by law, by local policy, or by ascertained potential need to the enterprise or potential need to the customer.

Enterprises need to periodically reassess the value of current holdings to the enterprise's future. Thee are some key points in time that are natural candidates for reassessment. Examples include planning for a physical move; upon transition from one contract to another; upon reassignment of data stewards, or upon corporate reorganization or sale.

The frequency for review of holdings can be easily managed in an electronic environment. Many enterprises have computer reporting of potentially obsolete data by virtue of a metadata sort, for example, date-driven reports that flag original metadata entries as suggesting they are obsolete. At such time, the enterprise reevaluates the continued value of such data. Upon reevaluation, readjust default retention dates or, if the data is determine to be non-value added, arrange for disposal. When evaluating the current and future value of data, some questions that may be asked as follows.

- What data is currently stored and for what purpose?
- Is the data accurate and up-to-date?
- What are the criteria for the storage life of data?
- What are the costs associated with data retention?
- Will retaining the data enhance network security?
- Is there potential for this data to satisfy legal requirements?

Value assessments and disposition processes pertinent to particular data should cease immediately upon knowledge of a potential or initiated lawsuit and not be re-initiated until after such action is completed.

Active data is the data that needs to be readily available to the enterprise for regular reference. Inactive data is data that still retains value but is not considered to have regular, continuing need. Such data needs to be retrievable, but can be stored less centrally. The data retention program should include a periodic review to determine when data is no longer being actively used, can be classified as inactive, and be moved to an archival location. Inactive data is usually relocated to an archival database, if electronic, or an off-site location, if electronic or paper. Relocation of such inactive data frees up physical and/or digital storage space. Supplier off-site storage may be an option for storing historical hard copy that is space consumptive. Suppliers who specialize in data storage may be able to offer rates that allow cost-effective retention. If backup or additional copies of these archived files are available, storage of the additional copies at a separate locale or even region of the country is prudent to overcome risk of local disasters such as hurricane, tornado, or earthquake. Conversely, if these archived files have no backup version or second copy, they are likely of greater value to the enterprise with local storage, since it normally facilitates faster retrieval.

Review both active and inactive data periodically to determine if they add value. If they do not, arrange for disposal.

7.4 Disposition data

Upon determination that data retained is not of continued value, arrange for disposal.

An enterprise should discontinue use of personnel energy, physical space, and other resources to non-value added data. Such non-value data should be purged from the enterprise records and disposed of in a manner that makes best business sense. When data is removed, its associated metadata should likewise be removed. A record should be retained to identify what documents were destroyed, when they were destroyed, and who authorized the destruction. An enterprise needs to have an effective process that ensures qualified individuals or teams of potentially interested parties authorize destruction. In some cases, as in merger environments, other enterprises may be required to jointly authorize destruction.

The enterprise should assure data assets that may be of interest to other parties are prevented from wide or public availability upon destruction. If there is potential value to entities outside the enterprise, such measures as electronic deletion, shredding or burning may be appropriate methods for destruction. Effective destruction processes ensure that the enterprise’s non-value-added data - its trash - is not compromised, becoming another enterprise's treasure.

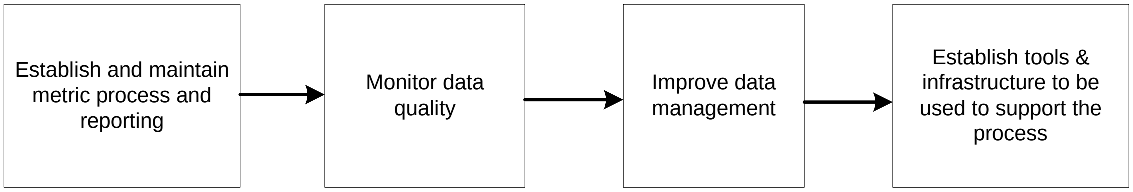
Failure to dispose of data in a timely, appropriate manner may exact a cost in poor use of space (square footage expenditure) or compromise of data.

8.0 Principle: Continuously Improve Data Management

Introduction

In a rapidly changing technological society, it is crucial to continuously improve the quality of the resources that house one of an enterprise's most valuable assets, data. DM is the function responsible for ensuring that the quality of data is consistent with the users' requirements. The purpose of this principle is to provide a basis for implementing a process for data quality improvement. Figure 8-1illustrates a methodology.

Figure 8-1. Improving Data Management


Establish and maintain metric process and reporting
Monitor data quality
Improve data management
Establish tools & infrastructure to be used to support the process

8.1 Recognize the need to continuously improve the quality of data

As indicated on Figure 8-1, metrics are essential for determining where to improve and to monitor improvement. Metrics should be designed to positively motivate, rather than keep score, and should focus on future strategy rather than providing a compilation of past history. Identify the users most directly involved with metrics and performance measurements and make them active members of the DM team. This may include personnel familiar with administrative, financial, technical and, where applicable, contracting issues. In order to effectively facilitate continuous improvement, the following questions need to be considered:

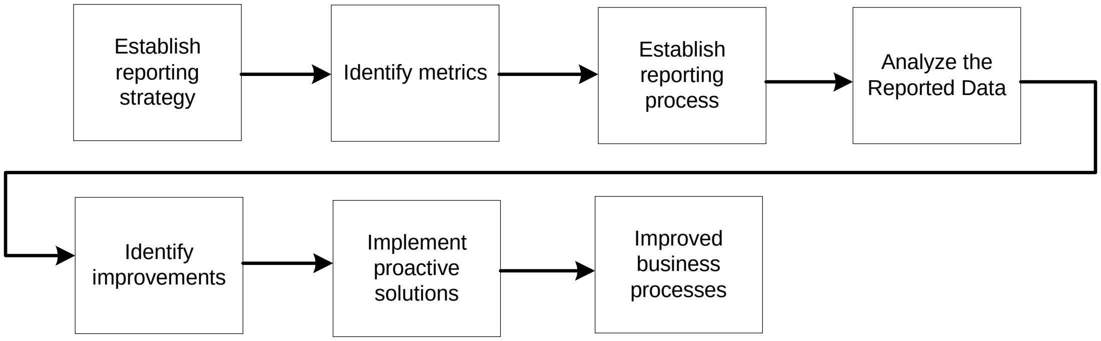
- What type of data is required, by whom and when?
- Who will use the data?
- How will the data be used?
- What is the user's infrastructure?
- How will the data be delivered?
- Where is the data maintained?

The answers to these questions establish a meaningful process and should be considered not only for the specific project but also reviewed for the overall impact to the enterprise.

8.2 Establish and maintain a metric process and reporting strategy

The steps in Figure 8-2 assist in establishing a reporting process that identifies and utilizes appropriate, advantageous metrics to improve the quality of data and the process that reports the quality. A consistent and repeatable process based on this enabler permits proactive actions by the enterprise and specific projects.

Figure 8-2. Process and Reporting Strategy


Establish reporting strategy
Identify metrics
Establish reporting process
Analyze the Reported Data
Identify improvements
Implement proactive solutions
Improved business processes

Metrics vary from enterprise to enterprise and from project to project. These process measurements should be simple and accurate indicators of performance, yet provide sufficient data to allow analysis. Examples of metrics that may be useful to assess the volume and performance of data activity are shown in Table 8-1.

Table 8-1 Examples of Data Management Metrics

Metric Name	Definition	Suggested Reporting Frequency
Data Schedule Status	Summarizes delivery status, contains number delivered early, on-time and late by month	Monthly
Electronic Delivery Status	Summarizes progress towards electronic delivery; contains the percentage of deliverables made electronically for each month	Monthly
Project Data Report	Summarizes the project data traffic; identifies the number of correspondence items transmitted between prime trading partners each month.	Monthly
Data Acceptance Rate	Percentage of submittals approved and disapproved by customer on first submission	Monthly

Table 8-1. Examples of Data Management Metrics (Continued)

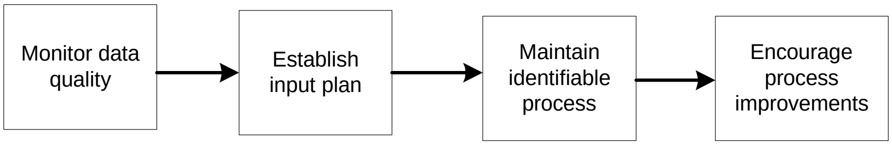
Metric Name	Definition	Suggested Reporting Frequency
Review Cycle Time	Measures the cycle time of the review and approval process both internally and externally	Monthly Issues addressed more frequently
Problem Reports	Measures the number of problems reported and time to problem closure.	Monthly
Customer Satisfaction	Measures customer satisfaction trends using survey methodology	Annual or on Project Closeout

For example, analyzing the review and approval process may identify areas for potential cycle time reduction resulting in improved data deliveries and schedule performance. If a particular area is revealed to have a longer turnaround for approval, the causes can be examined and corrected or the process can be changed to accommodate needs and achieve objectives. Publication of results and findings provides a method to standardize metrics across the enterprise.

8.3 Monitor the quality of data to improve data and processes

Monitoring the data quality (Figure 8-3) through the use of metrics ensures that changes in initial input quality are identified. Degradation of the data below the metric goal identifies a need to re-evaluate the goal and possibly update the process. As quality improves, the process can be changed to accommodate more stringent goals. The value of this activity is the on-going assurance that the quality of the data meets or exceeds requirements through an up-to-date identifiable process that also contributes to achieving enterprise goals.

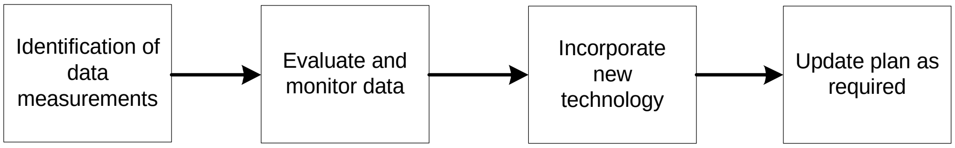
Figure 8-3. Monitoring Data Quality



Develop and implement a process improvement plan, with resources planned and allocated, to improve process performance and capability. Update the plan at defined intervals or as required. The plan should include identification, evaluation and incorporation of new technology innovations into the defined process. The plan should also include measures of effectiveness from which metrics can be derived.

The purpose of implementing a strategy (Figure 8-4) for on-going improvement is to ensure the plan is current and continues to provide direction and meet requirements. The value gained from this plan is the ability to readily identify improvements toward an objective of continuous improvement and preventative maintenance.

Figure 8-4. Improvement Strategy

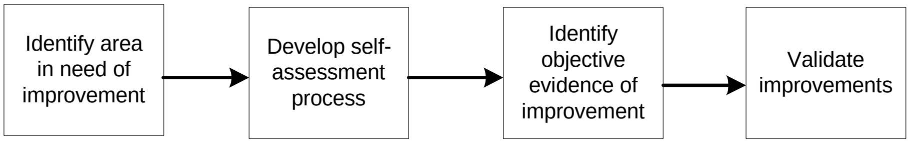

Identification of data measurements
Evaluate and monitor data
Incorporate new technology
Update plan as required

After establishing baseline requirements, a continuous improvement plan should be implemented. A systematic assessment of the process should be applied through planning, measurement, causal analysis and defect prevention, execution and process refinement. The results provide the basis for modification of systems and personnel retraining, as required. Adherence to this plan ensures that goals are reviewed on a consistent basis and adjusted as improvements in quality are made.

8.4 Improve data management through a systematic and self-diagnostic process

The four steps listed in Figure 8-5 are necessary in the development of a systematic approach to self-analyzing the identified improvement process. This self-assessment improves the ability to identify, analyze, address and correct data issues. The end result will be ready access to pertinent and accurate data.

Figure 8-5. Self-Diagnostic Process


Identify area in need of improvement
Develop self-assessment process
Identify objective evidence of improvement
Validate improvements

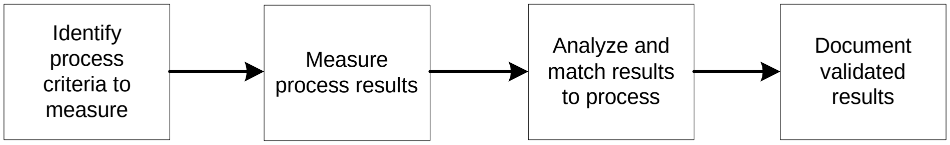
There are numerous process improvement projects used by enterprises today. Any systematic process used should address/identify the following:

- Prevent errors from recurring by identifying and eliminating the cause
- Periodic reviews, audits and publication of metric results
- Lessons learned, documented and published for reuse
- Identification of best practices, documented and published for reuse

For maximum gain and return on the costs of development, an effective improvement process includes periodic reviews and audits.

Develop objective evidence of improvements to provide concrete indication of the process results and/or actions and how they correlate to the process improvement (Figure 8-6). Validated documentation ensures that documented proof of improvements achieved as a result of the process improvement is readily available.

Figure 8-6. Develop Objective Evidence of Improvement

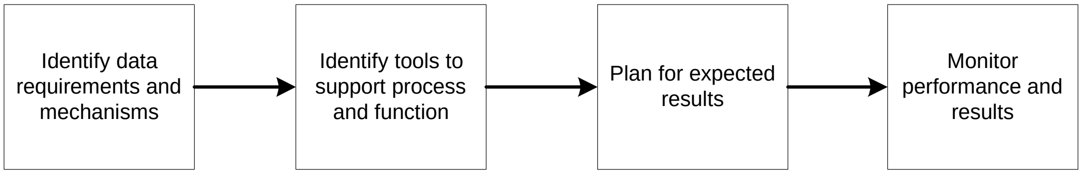

Identify process criteria to measure
Measure process results
Analyze and match results to process
Document validated results

Creation of a process assessment matrix, or other tracking method, listing the stated criteria, would be advantageous in indicating the process/steps taken to accomplish the results of objective evidence. Efforts expended and lessons learned to improve the process should be documented to leverage the efficiency of future measurements.

8.5 Establish the necessary tools and infrastructure to support the process and assess the results

Establishing the necessary tools and infrastructure (Figure 8-7) is essential. Identifying the data requirements at project inception and the mechanisms on which the process will be based reduces confusion and increase productivity. Knowing the expectations in the beginning benefits the project and the enterprise by allowing time for planning and achievement of the expected results thus saving time and money.

Figure 8-7. Tools and Infrastructure Support the Process and Assess Results


Identify data requirements and mechanisms
Identify tools to support process and function
Plan for expected results
Monitor performance and results

Tools may vary depending on the requirements and available infrastructure. At a minimum, the tools should have the capability to support the enterprise process and perform the following functions:

- Data scheduling
- Action tracking
- Data delivery

The process infrastructure at a minimum should include:

- Resources
- Information systems dependencies
- Training approach
- Other essential elements needed to support process improvement

Metrics enable:

- Performance monitoring
- Progress demonstration
- Project completion

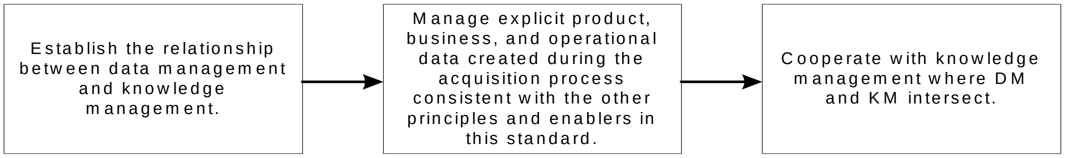
The evidence that metrics provide creates a level of confidence in the process being measured. A properly established set of metrics supports goals and process improvement, in addition to providing a basis for assessing the improvements and assessing trends.

9.0 Principle: Effectively Integrate Data Management and Knowledge Management

Introduction

This principle describes the interdependent relationship between DM and knowledge management (KM). Since knowledge management and data management are naturally interdependent, the objective of this principle is to distinguish the roles of each so that, in practice, KM and DM efforts are complementary. Ultimately, if DM and KM are managed such that they are complementary, the enterprise and its trading partners achieve increased return on their intellectual assets investment. Figure 9-1 is a top-level process view.

Figure 9-1. Understanding the Interdependence of DM and KM


Establish the relationship between data management and knowledge management.
Manage explicit product, business, and operational data created during the acquisition process consistent with the other principles and enablers in this standard.
Cooperate with knowledge management where DM and KM intersect.

9.1 Establish the relationship between data management and knowledge management

Table 9-1 illustrates the relationship between data and knowledge. The scope of knowledge spans both explicit information (data that is recorded on media) and tacit information (information that is held in the minds of individuals). Explicit data includes both structured data, such as is found in traditional databases, and unstructured data, such as is found in technical manuals, drawings, and reports. The term “unstructured” is used in the information technology field to contrast with the structured data found in databases. Unstructured data often does have some structure, such as a table of contents.

Table 9-1 Relationship Between Data and Knowledge

Type	Explicit		Tacit
	Structured	Unstructured	
	Transactions	Collaboration	
Examples	Purchase order Purchase order acknowledgement Invoice Remittance advice Request for quote Shipping schedule	Technical report, Analysis report, Specification, Manual Parts list Drawing …	Mental models Informal recipes Rules of thumb Lessons learned Communities of practice …
Normally responsibility of	Database administration	Data management	Knowledge management

The key insights regarding knowledge management, are:

- Knowledge comprises both explicit and tacit information

- Data can be structured or unstructured

- Structured data, which supports transactions, has historically been the domain of database administration

- Unstructured data, which supports collaboration, generated as part of the *acquisition process*, has historically been the domain of DM. There is also other unstructured data well outside the scope of historical DM. Examples are the books and articles that are managed by librarians; human resource files managed by that function, etc.

- Knowledge management, as it is emerging, primarily addresses the management of human aspects of knowledge, such as the fostering and support of communities of practice.

Knowledge management as a named field is relatively new, dating to approximately the mid-1990s. Thus the KM field is still emergent. Data administration and data management date back at least 40 years and have historically been separate fields. In past years, DM dealt mostly with paper products that described the product of an acquisition process. Data administration dealt mostly with automated data. With the impact of information technology on data, these two fields are migrating closer together. As a minimum, DM practitioners need to possess effective electronic skills to host and to search for data electronically.

As was discussed in the introduction to the standard, data can be defined in terms of broad classes (product, business, and operational). The scope of DM, within the context of this DM standard, extends primarily to product, business, and operational data as created and maintained during the life-cycle process. Since DM provides a solution for management of explicit product, business, and operational data created during the acquisition process; it inherently provides that part of the solution for knowledge management.

9.2 Cooperate with knowledge management where DM and KM intersect as KM methods develop

The emerging field of knowledge management is developing methods for managing tacit data and the human aspects of knowledge. All of these methods are subject to rapid change as enabling technologies emerge and mature. Since tacit data and the human aspects of knowledge management are outside the domain of DM as defined in Enabler 9-1, the best strategy for DM is to cooperate with knowledge management as KM methods develop while leveraging the capability that KM is developing. KM implementations generally stress three elements:

1. Organizing content (repositories and structure)
2. Connecting people (for purposes of facilitating data exchange)
3. Managing change (promoting knowledge sharing behavior and provide tools that facilitate knowledge management within the enterprise and with its trading partners.)

For any enterprise, DM and KM intersect with regard to element 1 and potentially element 3 (with respect to tools). Because the state of KM is fluid, the methods in use are enterprise dependent. Therefore integration with regard to element 1 and element 3 requires understanding the state of KM capability in the enterprise with respect to these elements.

9.2.1 Understand state of KM in the enterprise

There are three steps required to understand the state of KM in the enterprise.

1. Determine through interviews or similar means if the enterprise has designated either a formal or informal knowledge management individual or group. If there is no such individual or group then KM activity is ad hoc at best.
2. Determine through interviews or similar means if the enterprise has accomplished a KM capability self-assessment that is an internal assessment of its KM capability.
3. Use existing self-assessment or perform self-assessment.

If the enterprise has accomplished such a self-assessment and it is reasonably current, or if one is intended within a reasonable time period, use the results of that self-assessment.

If there is no such self-assessment accomplished or planned, then it is necessary to assess the value of developing a self-assessment, considering such elements as corporate culture and vision as well as cost. Such an assessment involves broadening the role of DM to encompass KM. Such assessments may yield tremendous potential benefit if a project is in the early phases of the life cycle. While inherently of less value in the later phases of the project’s life cycle, projects with long anticipated future use may still warrant a more encompassing DM role.

If the assessment suggests merit to further pursuit, then perform a self-assessment.

9.2.2 Coordinate DM and KM efforts.

Based on the results of enabler 9.2.1, coordinate DM and KM efforts for elements 1 and 3. The specifics depend on both the results of the KM self-assessment and the state of DM in the enterprise.

10.0 Application Notes

1. EIA 859 is a principles-based guidance document. It is not intended to be a compliance document in a contract for purchase and acquisition of a product. It may be used as a source for relevant information when preparing such items as a request for proposal or other business contract agreements. Following EIA 859 principles enables users to establish, plan, document, and communicate appropriate and consistent DM programs and plans for an enterprise or project environment.

2. This standard should be used to design, plan, implement, and sustain effective DM solutions for the enterprise or the project. DM practices should be selected and applied in context and to the extent appropriate.

3. Because this is a principles-based standard, the standard can be used to communicate and convey effective DM solutions and processes. Example users might be:

   - an practitioner who is seeking best methods for DM and associated solutions;
   - an author who is drafting a DM plan;
   - an manager who is developing criteria for measurement of a process.

4. This standard provides the basic terminology for DM. It has been coordinated with EIA-836 and ANSI/EIA-649A to ensure consistent use of common terms and definitions.

5. Annex D, “Non-Commercial Practices For Data Management,” identifies Department of Defense (DoD) data management procedures and relates the existing practices to this standard where applicable. The annex is intended to identify DoD-specific data procurement, management, and administration practices but not to direct or limit the use of any particular practice or business strategy

-

Annex A

Participants

Robert Armstrong
Robert.Armstrong@wpafb.af.mil

*Elizabeth M. Arre
Naval Sea Systems Command
202-781-3383
arreem@navsea.navy.mil

Jesus Arroyo
Jesus.Arroyo@HSC.com

*Nancy Brailsford
Lockheed Martin Federal Systems
(607) 751-4126
nancy.brailsford@lmco.com

Dillard Broadway
dillard.broadway@msl.redstone.army.mil

Susan Brock
susan.brock@eglin.af.mil

Michiel Budlong
michiel.budlong@baesystems.com

Steven Buzinski
Air Force Research Laboratory
(315)330-2407
steven.buzinski@rl.af.mil

Elizabeth J. Carter
Boeing
425-717-3765
elizabeth.j.carter@boeing.com

Karen L. Chandler
Boeing
karen.l.chandler@boeing.com

Gary L. Comet
cometg@GDLS.com

Rudy Cornelius
Rudy.Cornelius@msl.redstone.army.mil

Suzi Demichelle
suzi.demichelle@fscnet.vandenberg.af.mil

Peter Eirich
Peter.Eirich@jhuapl.edu

Professor Sam Epstein
AFIT
samuel.epstein@wpafb.af.mil

Lisa Fenwich
lfenwick@cmstat.com

+ Task Leader
* Primary Author

Patricia J. Ferrell
Det 1, AFRL/WSC
(937) 255-4627
Patricia.Ferrell@wpafb.af.mil

Robert Firment
Lockheed Martin Marietta (Ret.)
(321)268-9874
rfirment4@cfl.rr.com

Patricia Fisher
CM Data Repository
(321) 494-5086
Patricia.Fisher@rc.patrick.af.mil

+James A. Forbes
Logistics Management Institute (LMI)
703-917-7572
jforbes@lmi.org

Sandra Franchi
L-3 Communications
856-338-2742
Sandra.B.Franchi@L-3Com.com

Nick J. Frato
nick_j_frato@raytheon.com

Kathryn Fuszner
Air Force (Civil Service)
(850) 882-9307 x5115
kathryn.fuszner@eglin.af.mil

*Herman G. Gaines
NAVICP
herman.gaines@navy.mil

Lina Gaitlan
Lina.Gaitan1@losangeles.af.mil

Josefina H. Gerende
Naval Air Systems Command
(805) 989-3522
gerendejh@navair.navy.mil

*Vicki T. Girardi
U.S. Army Aviation and Missile Command
(256) 955-7621
Vicki.girardi@rdec.redstone.army.mil

+Cynthia Hauer
Millennium Data Management, Inc.
(256) 536-1096
HauerCC@aol.com

Dana Hazarna
dana.hazama@eglin.af.mil

William Hosmer
william.w.hosmer@saic.com

*Diane Howell
Northrop Grumman Shipbuilding
(228) 935-2971
dghowell@northropgrumman.com

*Bonnie Johnson
General Dynamics Decisions Systems
480-675-2805
Bonnie.Johnson@gdds.com

+ Task Leader
* Primary Author

*Donna Johnson
Naval Sea Systems Command
202-781-3383
johnsonds@navsea.navy.mil

Dena Joyce
dena.joyce@eglin.af.mil

Grace Keinath
United Defense LP
407-243-4946
grace.keinath@udlp.com

Lisa Kelley
USAF
(310) 336-4966
Lisa.Kelley@losangeles.af.mil

William Kellogg
william.kellogg@lmco.com

Jeff Klein
jeff.r.klein@boeing.com

Jim Knowles
Army Material Command
KNOWLESJ@hqamc-exchg.army.mil

Steve Larson
Microsoft Business Solutions
(701) 281-7614
stevlars@microsoft.com

Major Jeff Lathrop
USAF
Jeffrey.lathrop@wpafb.af.mil

Caroline Lee
caroline.lee@eglin.af.mil

David L. Lee
Northrop Grumman Space Technology
310.813.1432
david.k.lee@trw.com

Bob Leibrand
NAVICP Philadelphia
(215) 697-1095
Robert_C_ Leibrand@icpphil.navy.mil

*Stan Littlefield
United Defense, LP
(763) 572-4872
Stan.Littlefield@udlp.com

*Jan F. Lundy
Raytheon Systems
jflundy@raytheon.com

Leroy P. Maestas
SMC DET11/CID
(719) 556-2527
Leroy.Maestas@cisf.af.mil

Doris Martinez
doris.martinez@navy.mil

Pamela McCarthy
pamela.v.mccarthy1@jsc.nasa.gov

Stephen McGlone
U.S. Army Materiel Systems Analysis Activity
(309) 782-6521
steve.mcglone@us.army.mil

Paulette L. Meadows
Naval Air Warfare Center, Weapons Division
(760) 939-4579
Meadowspl@navair.navy.mil

Robert J. Meinhart
meinhart@pica.army.mil

Dick Menken
MenkenRE@phdnswc.navy.mil

John Montgomery
john.montgomery@redstone.army.mil

Joseph Mystkowski
Northrup Grumman
mystkjo@mail.northgrum.com

Iris Nichols
Iris.Nichols@Hill.af.mil

Kim Northup
Wright Patterson Air Force Base
(937) 904-0853
Kim.Northup@wpafb.af.mil

Elaine O'Reilly
Northrop Grumman
(516) 575-1457
Elaine.OReilly@NorthropGrumman.com

Karen Ragland
karen.ragland@wpafb.af.mil

Leslie Reed
NAVICP Philadelphia
(215) 697-3887
leslie_d_reed@icpphil.navy.mil

*Joe Roman
Lockheed Martin NE&SS-SS
865.914.6523
joe.roman@lmco.com

Graham Rutherford
graham.rutherford@lmco.com

Jeannie M. Sage
Boeing Commercial Airplanes
425-373-8563
jeannie.m.sage@.boeing.com

C.V. Schmidt
schmidtcv@navair.navy.mil

+ Task Leader
* Primary Author

Tom Schneider
US Army Material Systems Analysis Activity
schneidert@ria.army.mil

Ronald L. Schuldt
Lockheed Martin Enterprise Information Systems
303-977-1414
ron.l.schuldt@lmco.com

Kristen Shaffer
Army Material Command
(703) 617-5706
ShafferK-Contractor@hqamc-exchg.army.mil

Dick Smith
Dick.Smith@wpafb.af.mil

Paul M. Smith
Acquisition Support Office, USAF
(407) 384-3826
Paul_Smith@peostri.army.mil

Janice Strickland
Janice.Strickland@losangeles.af.mil

*Brenda Sutherland
NASA Marshall Space Flight Center
256-544-6552
brenda.sutherland@msfc.nasa.gov

Jenny Trenda
jenny_trenda@udlp.com

Jim Van Dyke
Boeing Commercial Airplanes
(425) 294-2644
james.j.vandyke@boeing.com

K. Vannoy
kvannoy@ems.jsc.nasa.gov

Tom Wajda
twajda@cse.l-3com.com

Kathy Walker
kathy.walker@lmco.com

*Karen Wheeler
Lockheed Martin Federal Systems
(607) 751-5733
karen.wheeler@lmco.com

Evette Wilkerson
wilkersonee@navsea.navy.mil

Joanne R. Wilson
The Boeing Company
joanne.r.wilson@boeing.com

Ruthie Young
Ruthie.Young@afscn.com

+ Task Leader
* Primary Author

Annex B
EIA-859 Glossary

Term	Source	Acronym	Definition
Acceptance			Formal acknowledgment that a product, data, or service conforms to requirements. Acceptance may occur before, at or after delivery.
Access			Formal arrangement between parties to be able to view/read data. (Note: Access, in itself, does not authorize use, reproduction, manipulation, altering or transfer of possession of data.)
Acquisition			The activity performed to obtain or come to have as one's own.
Approval	EIA-649A		Authorization from a designated authority that a product, process, or information is complete and suitable for use.
Approval authority			Designated entity with the authority to approve, disapprove, or otherwise disposition a change request and direct its implementation.
Approved	EIA-649A		A state signifying approval.
Archive			A place where public records, documents, etc. are stored; the act of storing records.
Archived information	EIA-649A		Information that has been retained for historical purposes that can be retrieved and is usable.
Attribute			Information related to a particular data element
Baseline	EIA-649A		Agreed to information that identifies and establishes the attributes of a product at a point in time, which serves as basis for defining change.
Bill of information		BOI	The bill of information consists of all the information and relationships to completely document the entire life cycle of a product--including the associated project information, (administrative, contractual, technical, and financial data) and it's location.
Business context			The whole situation, background; or environment relevant to a particular organization or business.
Business rules			The policies, practices, and procedures that drive day-to-day business activity and define a way of doing business.
Change authority			Designated entity with the authority to approve, disapprove, or otherwise disposition a change request and direct its implementation.
Configuration change	EIA-649A		An alteration to a product and/or its product configuration information.

Term	Source	Acronym	Definition
Configuration management	EIA-649A		A process that establishes and maintains consistency of a product's attributes with its requirements and product configuration information throughout the product's life cycle.
Contract			Any formal agreement between two companies, a government agency and a company, interdepartmental work authorizations within a company, memorandum of agreement, and any other form of agreement.
Controlled vocabulary			A limited set of consistently used and carefully defined terms.
Customer satisfaction			The fulfillment of the requirements, conditions, needs, expectations, wishes, or desires for any person with whom one has dealings
Copyright			A form of protection provided to the authors of "original works of authorship". In the U.S., this is provided by U.S. Code, Title 17.
Data	EIA-649 (V.4); MIL-STD-2549		Recorded information of any nature (including administrative, managerial, financial, and technical), regardless of medium or characteristics.
Data acceptance rate			Percentage of submittals approved and disapproved by the customer on the first submission of the data item
Data element	MIL-STD-974		A basic unit of information representing attributes of a data item instance.
Data format			The desired organization, structure, or arrangement of the content of the data product described by the DID or other data tasking document. This term relates to the shape, sized, makeup, style, physical organization, or arrangement of the data product described in the DID or other tasking document.
Data item			A document, drawing, report, manual, technical order, revision or other submission.
Data item description		DID	A standardization document that defines the data content, preparation instructions, format requirements, and intended use of data required of a contractor for a specific data product.
Data management		DM	The process of applying policies, systems, and procedures for identification and control of data requirements; for the timely and economical acquisition of such data; for assuring the adequacy of data; for the distribution or communication of the data to the point of use; and for analysis of data use.
Data manager			An individual designated to apply Data Management disciplines and who is responsible for ensuring compliance with organizational policy.
Data schedule status			Summarizes deliver status; contains number delivered early, on time and late by month.

Term	Source	Acronym	Definition
Data view			Presentation and organization of the information for the user. How the user "sees" or uses the data will be unique to individual need. The data could be viewed for a multitude of purposes: record keeping, decision-making, information analysis, etc. Data can be viewed either electronically or manually.
Data delivery			Data can be "delivered" or provided to a user in either an electronic or hard copy format, depending on the needs of the user. Delivery can be for review and/or use, or for retention and on-going maintenance. If data is prepared and maintained on an electronic system, it can be considered to have been delivered when it is available on the system, and the user has been notified that it is available.
Data products			A basic unit of recorded information of any nature, regardless of media or characteristics.
Data type	EIA-836		A delineation of the essential property of an element or attribute, such as date, string, number, currency, enumeration, etc.
Deferred delivery			A method of delaying the delivery times for specified data.
Deferred ordering			A method to establish the right to obtain data that may be needed in the future but for which a specific requirement is not identified at the current time.
Deliverable data			Information that is given over or transferred to another party.
Disapproval	EIA-649A		Conclusion by the appropriate authority that a product, a process or information is incomplete or unsuitable for it's intended use.
Distribution			Data exchange between a data source and a data recipient regardless of media used.
Document			A self-contained body of information or data that can be packaged for delivery on a single medium. Information relating to the design, procurement, manufacture, test or acceptance of an item or service, such as specification, drawing, list, standard, pamphlets, report, or any printed, typed or written item.
eXtensible Markup Language	EIA-836	XML	A markup language that provides a strict set of standards for document syntax while allowing developers, organizations and communities to define their own vocabularies.

Term	Source	Acronym	Definition
Intellectual property		IP	Products of the human intellect that the legal system is willing to protect against unauthorized use by others. Examples: inventions, discoveries, compositions and compilations. Different types of Intellectual Property also termed technology assets are: Trade Secrets Copyrights Trademarks and Service Marks Patents Maskworks Intellectual property must be documented to capture or record the idea or concept.
Intellectual property rights			The right to (or not to) use, make, have made, sell, lease, dispose of, modify, translate, copy, exhibit, perform, practice, etc., products or services embodying the Intellectual Property, and allow or not allow others to do so.
Interface	EIA-649A		The product attributes that exist at a common boundary of two or more products.
Legacy data			Data that is not in current standard digital format of is residing in older databases maintained with obsolete or inefficient technology. Legacy data can be in hard copy or digital format. The ability to use legacy data within an Integrated Digital Environment is severely limited unless the data is converted to standard digital format. Cost-benefit analyses are required to determine which legacy data should be converted.
License			A grant of the right to do something, which, if done without permission, would be illegal.
Life cycle	EIA-649A		A generic term for the phases in the life of a product from concept to disposal.
Metadata			Data about data. Properties used to identify or define a data item. This could include a title, document number, creation date, etc.
Ordering data			The act of contractually requiring access to, or delivery of, data in accordance with a data requirement that defines content, format, schedule or price.
Patent			A statutory monopoly on the use and commercial exploitation of an invention. Patent categories in the United States include: Design patents, which cover new, original and ornamental designs for articles of manufacture; Plant patents, which cover asexually reproduced varieties of plants; Utility patents, which are the most common type of patent.

Term	Source	Acronym	Definition
Product	EIA-649A		Something that is used or produced to satisfy a need or is the result of a process (e.g., documents, facilities, firmware, hardware, materials, processes, services, software systems).
Release			Authorization for dissemination of approved information and/or products subject to change management.
Requirement	EIA-649A		(1) Need or expectation that is stated and obligatory, (2) specified value for an essential product attribute.
Revision	EIA-649A		The result of revising a product or product configuration information (also see version).
Schema	EIA-836		A set of rules describing a document structure; used herein in the generic sense of data relationships and provide context for data element and attribute definitions.
Specification	EIA-649A		Information that explicitly states the requirements for product attributes
Style sheet	EIA-836		A list of specifications describing how to present a document in a particular medium.
Subcontractor			An organization or activity that provides goods or services to a contractor.
Technical data			Scientific or technical information recorded in any form or media necessary to operate and maintain a system. The term does not include computer software or data incidental to contract administration.
Validation	EIA-649A		Authentication that the requirements for a specific intended use or application have been fulfilled.
Verification	EIA-649A		Confirmation that a specified requirement has been fulfilled by the product.
Version	EIA-649A		A specific configuration of a product which varies from other configurations of the product (see revision).

Annex C
Data Management Function Table

	Skill/ Function	Clerical	Budgeting	Cost/benefit analysis	Strategic planning and management	Program management	Contracting	Legal implications	Technical Library management	Configuration Management	Database management	Metadata management	Process design and development	Software engineering	Knowledge management		
Strategy and architecture																	
1	Development of enterprise DM strategies		X	X	X	X	X	X	X	X	X	X	X	X	X		
2	Development of DM plans		X	X	X	X	X	X	X	X	X	X	X	X	X		
3	Development of DM policies			X	X		X	X	X	X							
4	Development of IP strategies			X	X	X	X	X		X			X		X		
5	Integration of DM and knowledge management								X		X	X	X		X		
6	Resourcing of DM requirements		X	X	X	X											
Process and infrastructure design																	
1	Design of data access provisions						X	X		X	X	X	X	X			
2	Development of paper data formats	X							X				X				
3	Development of electronic data formats										X	X	X	X			
4	Design of DM processes			X		X	X	X	X	X	X	X	X				
5	Design and development of data envirionments			X	X	X	X	X	X	X	X	X	X	X	X		
6	Development of provisions for interoperability and interchange			X	X		X	X		X	X	X	X	X	X		
7	Development of training syllabi and courses	X	X	X	X	X	X	X	X	X	X	X	X	X	X		
8	Development and management of meta data							X	X	X	X	X	X	X			
9	Design of data products and views						X	X	X	X	X	X	X		X		
EXecution																	
1	*Requirements identification and definition*			X		X	X	X		X			X	X	X		
2	DM risk assessments		X	X	X	X	X	X									
3	Prioritization of data requirements		X	X		X	X			X				X			
4	*Control of data requirements*					X	X			X							
5	*Control of deliverables received*	X					X			X							

Annex C
Data Management Function Table

	Skill/ Function	Clerical	Budgeting	Cost/benefit analysis	Strategic planning and management	Program management	Contracting	Legal implications	Technical Library management	Configuration Management	Database management	Metadata management	Process design and development	Software engineering	Knowledge management		
6	*Oversight of data preparation*					X				X							
7	*Data marking*	X					X	X	X	X		X					
8	*Import/eXport control*	X				X	X	X	X	X		X			X		
9	*Preparation and maintenance of inventory master lists*	X				X			X	X		X					
10	Conversion from paper to electronic	X	X	X	X	X		X	X	X		X	X				
11	Management of data collaboratively developed via IPTs or similar methods		X	X	X	X	X	X	X	X	X	X	X	X	X		
12	Management of intellectual property	X	X	X	X	X	X	XX	X	X	X	X	X	X	X		
13	Implementation of access provisions	X	X			X	X	X	X	X	X	X	X	X	X		
14	*Data archiving*	X	X	X	X	X	X	X	X	X	X	X	X	X	X		
15	*Data disposal*	X	X	X	X	X	X	X	X	X	X	X	X	X	X		
Process and infrastructure maintenance																	
1	Recurring DM training	X	X	X	X	X		X	X	X	X	X	X	X	X		
2	Management of electronic repositories		X	X	X	X		X	X	X	X	X	X	X	X		
3	Management of paper repositories	X	X	X	X	X		X	X	X		X	X		X		

ANNEX D
NON-COMMERCIAL PRACTICES FOR DATA MANAGEMENT

Table of Contents

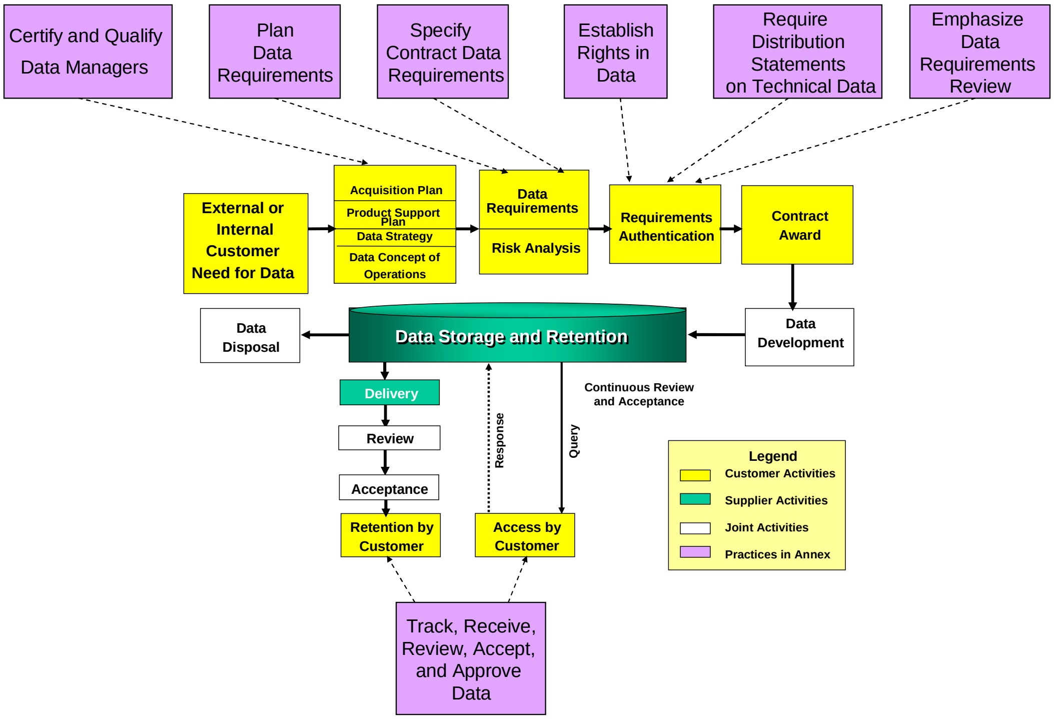
Introduction

The purpose of this Annex is to identify Department of Defense (DoD) data management (DM) procedures and relate the existing practices to GEIA 859 where applicable. This Annex is intended to identify data procurement, management, and administration practices but not to direct or limit the use of any particular practice or business strategy.

The procedures within this Annex describe business practices that DoD performs as a potential data customer and why the DoD performs differently from common industry data customers. Although Public Law and other regulatory requirements often govern the business practice for data, only a few specific regulatory references are cited. The Federal Acquisition Regulation (FAR) and the Defense FAR Supplement (DFARS) are two of the regulatory requirements.

The principles in GEIA 859 focus on the creation and management of data with a lesser emphasis on the acquisition of data and little emphasis on obtaining data rights. The DoD is interested: in the acquisition of data access, and data delivery; and obtaining data rights in an evolving collaborative development and management environment. The DM model shown in the GEIA 859 shows the activities performed by the customer, supplier, and joint activities. The model follows the spectrum of the entire life cycle beginning with customer need and ending with data disposal. The DM practices identified in the Annex have a heavy emphasis on activities prior to acquisition or contract award. The practices identified in the Annex are overlaid in the following figure to show how they relate to the GEIA 859 DM model.

Figure: GEIA-859 Data Management Model & Non-Commercial Practices


Certify and Qualify Data Managers
Plan Data Requirements
Specify Contract Data Requirements
Establish Rights in Data
Require Distribution Statements on Technical Data
Emphasize Data Requirements Review
External or Internal Customer Need for Data
Acquisition Plan
Product Support Plan
Data Strategy
Data Concept of Operations
Data Requirements
Risk Analysis
Requirements Authentication
Contract Award
Data Development
Data Storage and Retention
Data Disposal
Delivery
Review
Acceptance
Retention by Customer
Response
Query
Continuous Review and Acceptance
Access by Customer
Track, Receive, Review, Accept, and Approve Data
Legend
Customer Activities
Supplier Activities
Joint Activities
Practices in Annex

Practice 1—Certify and Qualify Data Managers

This practice introduces the concepts of qualification and certification for DoD Data Managers. DoD Data Managers meeting the qualification requirements (training and experience) set forth by Service or Agency requirements perform the DM function in support of DoD activities. Certification is not a requirement of all Services, but is considered a best practice. DoD Data Manager certification does not normally require attainment of an industry certification, but rather formally recognizes fulfillment of specified prerequisites and is the final step in the qualification process. Qualified and certified individuals facilitate defining program data requirements.

Reference Principle 8—*Continuously Improve Data Management*

Practice 2—Plan Data Requirements

This practice focuses on the DoD perspective as a data customer and identifies the process involved in establishing program data requirements that consider all life cycle phase aspects of a program. Initial data requirement planning is an integral part of the program planning process and is free of restrictions and the considerations, which must be made during subsequent phases of data acquisition to narrow the data requirements.

Practice 2.1—Determine risk, maturity of systems, and life-cycle phase.

Acquisition strategy (AS) and Acquisition Plan (AP) are reviewed to determine the elements of risk, maturity of systems and equipment, and life-cycle phase of materiel development. When determining the data requirements for a program, the AS is examined to determine the intended methods of initial acquisition, operation and use, life cycle support, and disposal. The AP for each procurement and contractor logistics support, drive the initial set of contract data requirements and determine when the data are needed.

Practice 2.1.1—Acquire minimum essential data.

Justification of the need for all data requirements is established to ensure the acquisition of minimum essential data. Each data requirement and justification is reviewed and an analysis is performed to determine the minimum essential data needed. For follow-on contracts in a mature program, a review of the various program plans is performed for all functional areas (e.g., Integrated Logistic Support, Configuration Management, Reliability & Maintainability), which provides identification of specific data drivers that have become part of the program.

Practice 2.1.2—Generate Data Requirements from the work tasks.

Specific data requirements are generated from the work tasks in the Statement of Work (SOW). Requirements for work statements and any related attachments or special contract requirements, in coordination with program participants and functional staffs, are typically reviewed. Work requirement statements are scoped in accordance with the AP and program phase. A Statement of Objectives (SOO) may be used in lieu of SOW requirements in the request for proposal (RFP). These work statements are the basis for data requirements.

Practice 2.2—Identify Data Requirements from Data Calls.

A "data call" or alternative (e.g., an integrated product team) is conducted to identify the overall requirements for contractor-prepared data for a specified program or procurement. The purpose of the data call is to afford all program participants the opportunity to identify individual data requirements. The results of the data call confirm and adjust the data requirements planning effort.

Practice 2.3—Review, select, and consolidate Data Requirements.

The minimum data requirements for administration, management, and technical aspects of the acquisition are determined. The review, selection and consolidation of data requirements are conducted to assemble the data procurement package. Reviews also indicate possible deferred delivery and deferred ordering candidates.

Reference Principle 2—*Plan for, Acquire and Provide Data Responsive to Customer Requirements*

Practice 3—Specify Contract Data Requirements

This practice defines the DoD-unique actions involved in properly specifying data requirements prior to the award of a contract, from the perspective of the DoD as a data customer. This practice includes matching the work tasks in the SOW and pre-award contract elements to an appropriate Data Item Description (DID), which may be tailored to delete unnecessary requirements. Typical pre-award contract elements include, SOOs, SOWs, DIDs, attachments to the contract (e.g., Technical Manual Contract Requirements (TMCR), SOW attachments), special contract provisions, technical instructions, reference documents cited for compliance (e.g., specifications, plans, etc), and an Integrated Logistic Support Plan (ILSP). Each DID is subjected to a complete examination to ensure that the required data product is explicitly described. Where an existing DID does not provide sufficient content requirement or a new requirement is identified, a one-time DID is recommended for use. The Contract Data Requirements List (CDRL, DD Form 1423) is the mechanism used to place data and delivery requirements on contract and is the bridge between the SOW and the DID.

Reference Principle 2—*Plan for, Acquire and Provide Data Responsive to Customer Requirements*

Practice 4—Establish Rights in Data

This practice addresses data rights requirements for contractor data, how the DoD establishes or obtains data rights, and the considerations and clarifications between the contracting parties that are necessary for identification of proprietary data claims. The typical rights for the DoD are established in the DFARS Part 227. The rights are Unlimited, Limited, Restricted (Specifically Negotiated License Rights), and Government Purpose License. A data rights category is assigned based upon the type of data being generated by the contractor and acquired by the DoD. The DoD ascertains the total data requirements of a program and balances competing interests in rights to and value of data. The application of business decisions supports the acquisition and use of the data. Each data element is identified and the contractor is notified of the intent for the Government to require delivery. The contractor is allowed to price the data and reach agreements that streamline the collection, maintenance, and delivery process in advance of the contract initiation. This yields a clear understanding between the parties regarding contractor rights, the value of the data, and resulting legal responsibilities.

The DoD may declare unlimited rights to any technology and the associated data developed exclusively using Government funding when such action meets the minimum essential program requirements, and data items or data products are identified for delivery. Negotiations between the Government and the contractor, relating to data rights of specific data products, provide early identification of Government rights, procedures for the acquisition of technical data, and the rights to use, modify, reproduce, release, perform, display or disclose technical data, computer software, and software documentation.

The DoD can include a provision (clause) for pre-notification of rights in technical data in the contract. This clause requires the contractor to assert data rights position on each data product prior to contract award. Conflicting claims of data rights can be subject to negotiation with each side providing the legal basis for claims, and avoiding unexpected consequences for either party.

Data possession and data access do not automatically convey data rights. For example, the DoD may have specified, negotiated, and obtained unlimited rights to a data item, but may request that the preparing contractor maintain possession of the data item until the actual delivery of the data is required. The ramifications of so doing must be clearly understood by both parties, and responsibilities and expectations for maintenance of the data by the contractor must also be identified.

Reference Principle 6—*Establish and Maintain an Identification Process for Intellectual Property, Proprietary and Competition Sensitive Data*

Practice 5—Require Distribution Statements on Technical Data

This practice identifies the DoD-unique aspects of marking distribution statements on technical documentation. Distribution Statements denote the extent a technical document may be distributed, released and disclosed without additional approvals or authorizations. Distribution statements are required on all technical documents in the possession of, or controlled, by DoD components.

A Distribution Statement marking is distinct from, and in addition to, a security marking. DoD Directive 5230.25 sets forth policies, procedures, and responsibilities for withholding unclassified technical data with military or space application from public disclosure. DoD Directive 5230.24 establishes a distribution marking system for technical documents. Both directives implement the provisions of Public Law 98-94. The DoD identifies Distribution Statement requirements on the CDRL for each applicable data item.

Reference Principle 6 – *Establish and Maintain an Identification Process for the Intellectual Property, Proprietary, and Competition-sensitive Data*

Practice 6—Emphasize Data Requirements Review

This practice portrays the DoD perspective as a data customer and explains the DoD use of a Data Requirements Review Board (DRRB) or equivalent to review, evaluate, adjust and consolidate data requirements for acquisitions at certain funding thresholds and for critical technology development and sustainment.

DoD organizations establish DRRBs to emphasize that only essential data is acquired, that the data requirements are consistent with the phase of the program life cycle and conform to the overall policy of the FAR, DFARS, and organizational instructions, and that intended uses of the data are consistent with the justification.

The DRRB conducts a multi-functional review of the data requirements and is empowered to validate the need and applicability of Government requirements for cited deliverables. Other DRRB functions include validating adequate quality assurance provisions, data delivery dates, and that the stated distribution fulfills the essential requirements of a program. The DRRB fosters digital delivery of data, and determines whether deferred ordering or deferred delivery is considered to manage the expense associated with each data item.

Reference Principle 2—*Plan For Acquire and Provide Data responsive to Customer Requirements*

Practice 7—Track, Receive, Review, Accept and Approve Data

This practice addresses the DoD data customer role of receiving, reviewing and determining the acceptability of data deliverables.

There are several post contract award steps conducted to monitor that CDRL delivery requirements are accomplished. Organizational Data Manager may maintain a computer-based data file index of all data deliverables required for each awarded contract. The acquisition manager, Program Manager (PM) and Data Manager may utilize matrices to document the conversion of event-triggered CDRL actions into actual dates and to identify and control data review by the responsible office. The responsible office certifies data acceptability (inspection and acceptance) and validates distribution statements. The responsible office also surveys requiring and using activities to determine the validity of the data and the methods of use and application.

The acquisition manager, PM and Data Manager are advised of actions and all data deficiencies, respond to inquiries from contractors, and advise changes in data requirements.

Reference Principle 2—*Plan For Acquire and Provide Data responsive to Customer Requirements*