International Planetary Data Alliance

System Architecture Specification

System Architecture Working Group

April, 2009

Version 0.1

CHANGE LOG

Revision	Date	Description	Author
0.1	April 2009	Initial draft based on information collected by the System Architecture Working Group.	D. Crichton, R. Joyner

TABLE OF CONTENTS

1.0 EXECUTIVE SUMMARY

The multi-disciplinary nature of planetary science and the increasing number of national space agencies involved in planetary exploration suggest the need for a common System Architecture to ease discovery, access and use of planetary data by world-wide scientists regardless of which agency is collecting and distributing the data and to ensure access to and exchange of high quality planetary science data products across international boundaries.

1.1 PURPOSE

The purpose of this document is to initiate the development of the System Architecture for the International Planetary Data Alliance (IPDA) which will convey the resulting architecture in a manner that is understandable to the broad spectrum of IPDA stakeholders.

- System Architecture (applies to the IPDA system as a whole):

    1. A formal description of a system, or a detailed plan of the system at component level to guide its implementation [5].
    2. The structure of components, their interrelationships, and the principles and guidelines governing their design and evolution over time. [6].

The IPDA, in its level 1 and 2 requirements, identified the following requirement which is the driver for the architecture

> *2.1 IPDA will develop recommendations for interoperability within a federation of international planetary data archive systems.*

Note: This document is currently in draft form and contains sparse content in certain sections as well as transitional verbiage to aide in the flow of the document. These issues and others will be addressed in subsequent versions of the document.

1.2 APPLICABLE DOCUMENTS

1) IPDA Information Level 1 and 2 Requirements, January 2008, http://planetarydata.org/standards/ipda-requirements-20080122.pdf.

2) Developing a Core Set of Data Standards for the IPDA, Concept White Paper, January 2007, http://planetarydata.org/documents/white-paper-wp/ipda-wp-001_1_0_2007feb07-ipda-developing-a-core-set-of-data-standards-for-the-ipda

3) NASA-PDS/ESA-PSA Planetary Data Interoperability, July 2005, http://planetarydata.org/documents/white-paper-wp/IPDA-STC-WP-001_1_0-2005JUL01-NASA%20ESA%20Interoperability.pdf.

4) Planetary Data System (PDS) Standards Reference, March 20, 2006, Version 3.7, JPL D-7669, Part 2.

5) The Open Group Architecture Framework (TOGAF), Version 8.1.1, Enterprise Edition, April 2007.

6) ISO/IEC 42010 (IEEE 1471-2000), Systems and software engineering - Recommended practice for architectural description of software-intensive systems, July 15, 2007.

7) *International Planetary Data Alliance (IPDA), Information Model, June 29, 2007.*

8) CCSDS Reference Architecture for Space Information Management, CCSDS 312.0-G-1, June 2006.

2.0 ARCHITECTURAL ELEMENTS

There are three areas of architecture that are commonly accepted as subsets of an overall software system architecture:

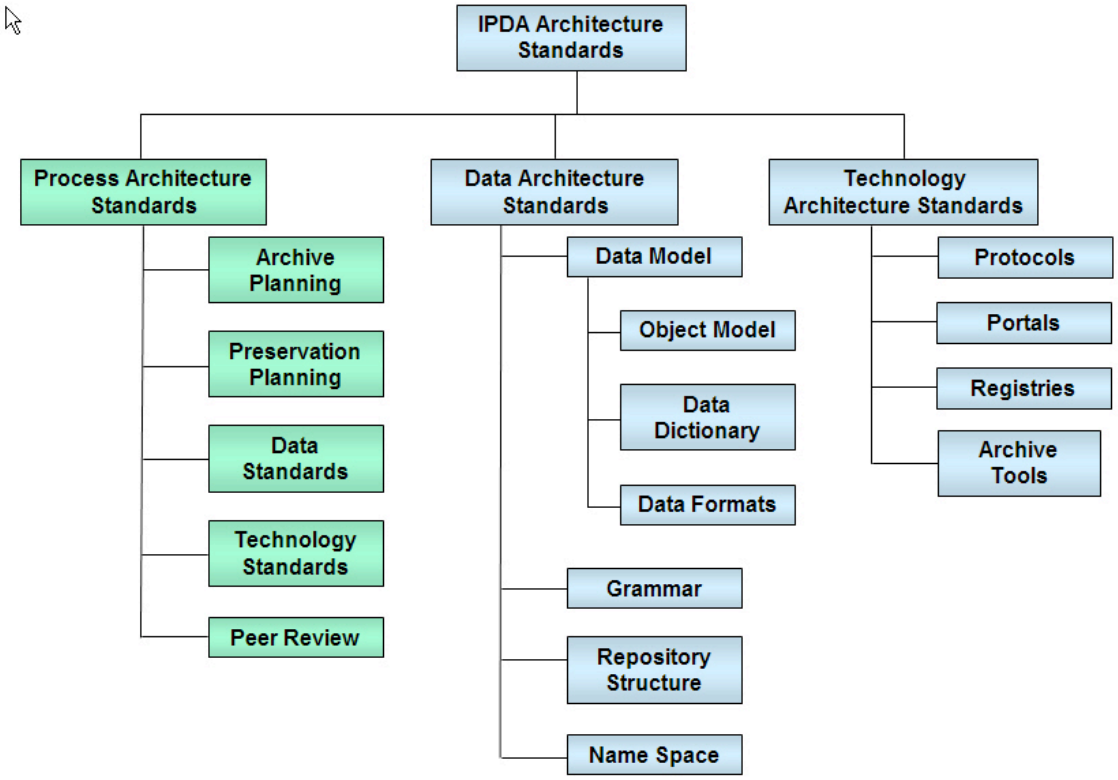
- Process Architecture
- Data Architecture
- Technology Architecture

The above three areas are defined in the following sections which specifically address the reference model elements in each of these architectural areas and specify where an element is derived from the IPDA Requirements [1].

2.1 PROCESS ARCHITECTURE

The Process Architecture defines the core processes of IPDA (e.g., archive management, preservation planning, peer review, standards management, etc). These elements are critical to the function of an archive system and the

processes can be shared across agencies. The highlighted portion of the architecture decomposition diagram below indicates the reference model elements associated with this portion of the system architecture:



The IPDA Process Architecture Standards promotes standards for collecting and describing planetary science data across the international community regardless of which agency is collecting and distributing the data. The IPDA Process Architecture Standards will be described in two documents:

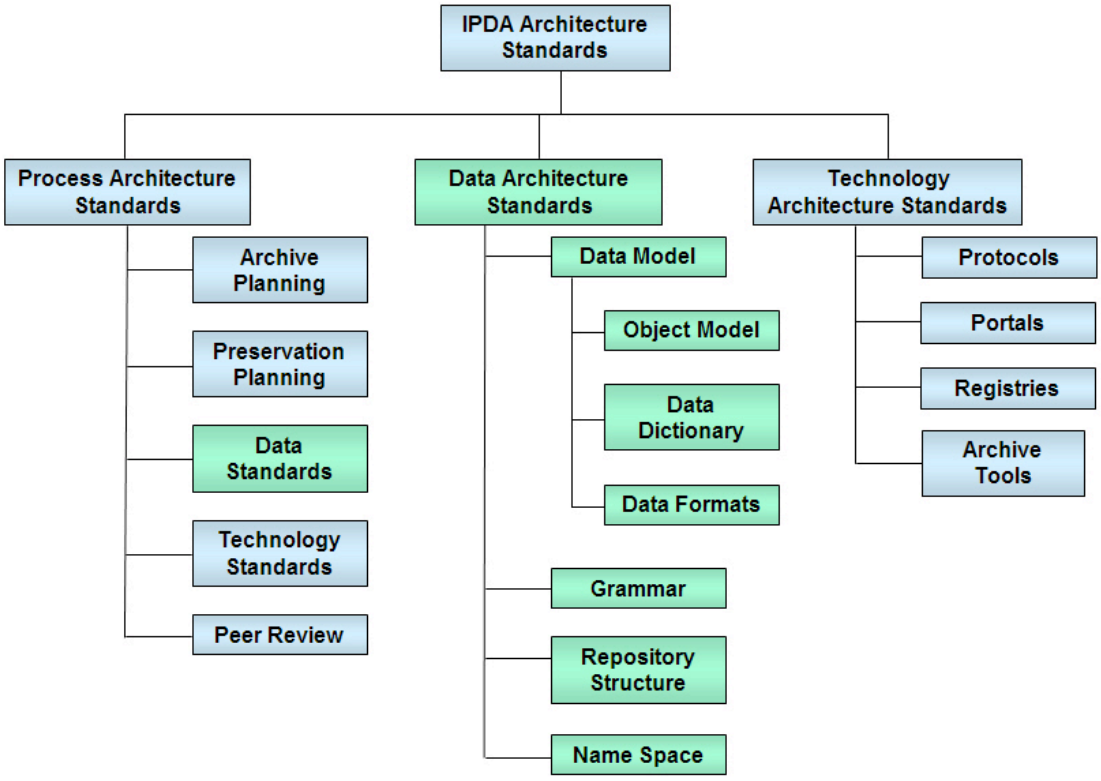
1. IPDA Standards Reference – identifies minimal standards to be IPDA compliant and includes examples

2. IPDA Archive Process Guide – identifies the common / standard processes for archiving data with the IPDA

Further detail can be found in section 4.1.

2.2 DATA ARCHITECTURE

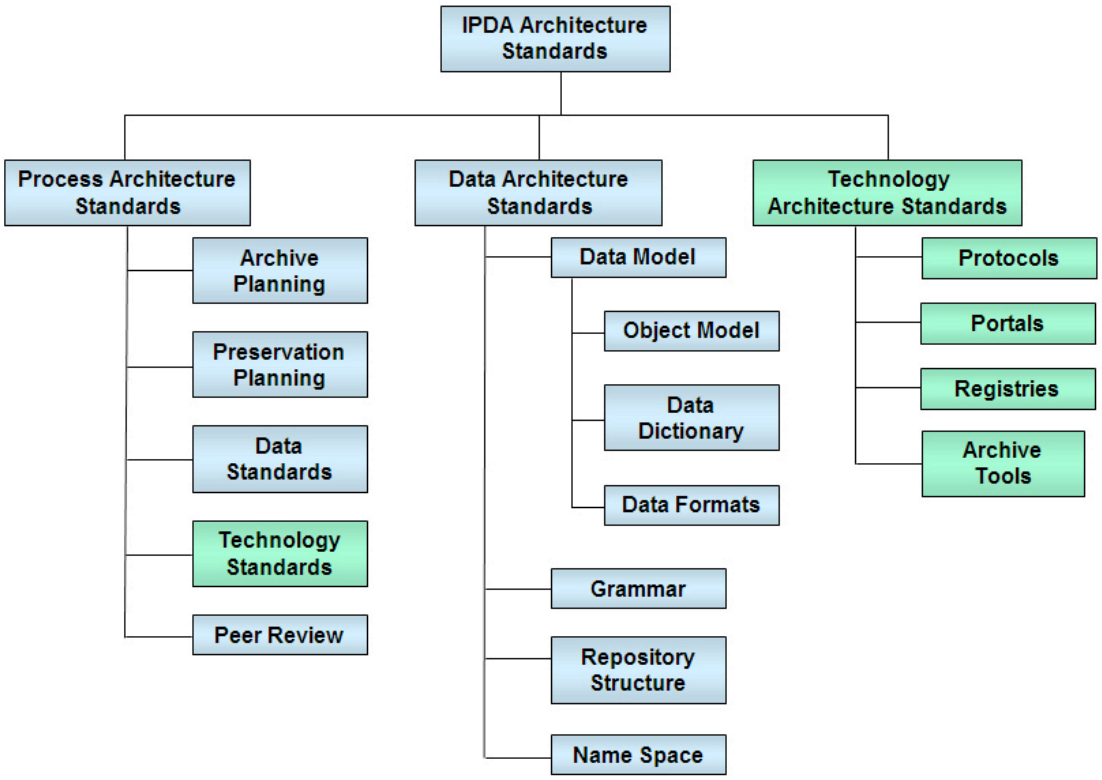
The Data Architecture is defined as the structure of an organization's logical and physical data assets and data management resources. The highlighted portion of the architecture decomposition diagram below indicates the reference model elements associated with this portion of the system architecture:



The bulk of the Data Architecture effort has been addressed in a separate document [7]. Refer to the "*International Planetary Data Alliance (IPDA), Information Model, June 29, 2007"* for more detailed information on the IPDA Data Architecture. Further detail is each of the elements can be found in section 4.2.

2.3 TECHNOLOGY ARCHITECTURE

The Technology Architecture describes the set of standards for enabling interoperability between planetary science data archive systems and the underlying services. The highlighted portion of the architecture decomposition diagram below indicates the elements associated with this portion of the system architecture:



Each of these elements are necessary to support development of compatible archives and data sharing among archive systems. Further detail can be found in section 4.3.

3.0 ARCHITECTURAL PRINCIPLES

Architectural principles are often used to form a general basis for decision making of architectural choices for a system. Large-scale software systems typically have explicit principles that are used to guide the evolution of components and the systems for large enterprise teams [6]. The principles detailed below are based on the architectural drivers and the initial set of principles identified by the IPDA Architecture Working Group, along with influences from the TOGAF example principles.

The principles are to be used as a guide to direct large-scale software systems. These principles can be used for guiding future decisions pertaining to upgrades and new technology choices. The principles are organized according to their most pertinent aspect of the system architecture (e.g., process/core, data, or technology). Descriptions for each of the principles include the following:

- **Statement** - Brief description of the principle.
- **Rationale** - Describes why the principle is important and any relationships to other principles.
- **Implications** - Lists any requirements or impacts this principle will have on the resulting architecture and system.

3.1 PROCESS / CORE PRINCIPLES

The process/core principles are as follows:

Common Use Software Tools

Statement: Shared tools across agencies are preferred over local tool development.

Rationale: Development and sharing of common software tools will both be cost effective and help to drive greater interoperability among agency planetary archive systems.

Implications:
- Both IPDA and local agencies will need to coordinate to identify common tools and to promote their use.

3.2 DATA PRINCIPLES

The data principles are as follows:

Model Driven

Statement: The IPDA will design and maintain a conceptual data model that is implementation independent.

Rationale: The data model is critical to defining and supporting interoperability among agency archive systems. This includes standard data objects, elements and values and their expressed relationships.

Implications:
- The data model should be defined using a formal data-modeling notation.
- Software developed for the system should be designed to evolve as the data model evolves.

Common Vocabulary and Data Definitions

Statement: Data are described consistently using a common vocabulary throughout IPDA archive systems. The definitions are understandable and available to all agencies and users across the IPDA.

Rationale: A common vocabulary, based on a data model, is an essential component for generating and maintaining quality metadata. It also aids in the understanding of the data from the users perspective. A common vocabulary is necessary when supporting cross-agency searching of federated archive systems.

Implications:
- An international data dictionary must be established and utilized uniformly throughout IPDA.
- Additions or modifications to the data dictionary must be coordinated across agencies.
- Agencies also need independence so the data dictionary must be concise and stable so there is little change at an international level.

Data are an Asset

Statement: IPDA data are an asset that has value to the international space agencies and the larger Planetary Science community and is managed accordingly.

Rationale: Data shared across agency boundaries should be available in a form that is IPDA compliant.

Implications:
- IPDA member agencies should ensure that data ingested into IPDA conforms to the IPDA Standards pertaining to metadata content, data format and data quality.
- IPDA member agencies should ensure the safekeeping of its data holdings.

Data are Accessible

Statement: Data are accessible for users to perform their functions, regardless of where the data are located.

Rationale: In the age of the Internet, everything is expected to be online and downloadable at the click of a mouse. The data holdings of IPDA are no exception.

Implications:

- In order to enhance accessibility, the IPDA must offer search interfaces for discovering data within the holdings of each agency's archive system.
- The IPDA data holdings must be online and accessible via the Internet.

Data are Usable

Statement: Data are usable for users to perform science analysis appropriate for their domain.

Rationale: In order to make IPDA data usable, data should be made available to users in contemporary formats.

Implications:
- Data are available in ready-to-use formats or can be easily converted to those formats.
- Appropriate metadata is captured and made available to users in order to identify and analyze the usefulness of the data.

3.3 TECHNOLOGY PRINCIPLES

The technology principles are as follows:

Interoperability

Statement: IPDA software should conform to appropriate standards that promote interoperability, including IPDA standards for data sharing.

Rationale: IPDA should promote the use of common technology standards for sharing of data in its tools and services. This will help to increase sharing of data and software services across agencies.

Implications:
- Industry standards should be adhered to where appropriate. These standards should pertain to capabilities and interfaces as opposed to specific products.
- Services and interfaces must be well defined to promote interoperability.
- Member agencies should adopt IPDA standard, where possible.

Distribution

Statement: IPDA data and software services are distributed and maintained independently by member agencies

Rationale: IPDA is not a centralized system. Software services and data are governed by local agency policies and systems. This is one of the key drivers for interoperability.

Implications:
- Differences in quality of services, capabilities, etc across member agencies should be expected.

4.0 DETAILED ARCHITECTURAL VIEWS

This section specifically addresses each element in each of the three architectural areas (e.g., Process, Data, and Technology Architecture Standards).

4.1 PROCESS ARCHITECTURE ELEMENTS

The following elements are associated with the Process Architecture Standards area:

Archive Planning

This element covers the necessary standards and processes for archiving data that fits into the IPDA standards. It provides the necessary information that both a agency would need in order to develop their own processes for planning for the capture of archival data.

Preservation Planning

This element covers the necessary standards and processes for long-term preservation of data. It provides the necessary information to enable agencies to plan for long-term preservation.

Data Standards

This element covers the process for managing the IPDA Data Standards. The IPDA Data Standards consist of the core standards for defining archival data holdings based on the IPDA Data Architecture. This element is concerned with ensuring there is a process in place to coordinate, manage and communicate IPDA data standards.

Technology Standards

This element covers the process for managing the IPDA Technology Standards. The IPDA Technology Standards consist of the core standards for supporting discovery, access and sharing of data across agency archive systems. It is critical that IPDA have stable interfaces and ensure that changes to these interfaces are coordinated and communicated to software system developers that integrate with the IPDA.

Peer Review

This element covers the standards and processes for peer reviewing data submissions to the IPDA. While it is expected that each agency will have its own peer review process, this will provide a high level model for conducting peer reviews for use by agencies, as needed.

4.2 DATA ARCHITECTURE ELEMENTS

The following elements are associated with the Data Architecture Standards area:

Data Model

This element identifies the overall data model for IPDA including the objects, attributes and their relationships. The IPDA Data Model includes:

- Object Model – This sub-element of the Data Model identifies the collection of formalized objects and classes through which the IPDA examines and manipulates the specifics of the data elements identified within IPDA.

- Data Dictionary – This sub-element of the Data Model identifies the necessary metadata that contains definitions and representations for data elements identified within IPDA. It is expected that the data dictionary will define a general set of keywords for international use.

- Data Formats - This sub-element of the Data Model identifies the data formats needed for archiving data submissions to the IPDA.

Grammar

This element identifies the necessary grammar(s) for describing IPDA data. A grammar is a common syntax and structure for capturing the metadata that is used to describe IPDA data.

Repository Structure

This element identifies a recommendation for the overall organizational structure for an IPDA archive. Common structures, whether physical or virtual, will be necessary to ensure consistent access across member agency archive systems.

Name Space

This element identifies the name space structure for the IPDA standards. The namespace will be critical to identifying the management of “keywords” and “objects” within the model. This will allow for separation of international portions of the data model from local agency portions.

4.3 TECHNOLOGY ARCHITECTURE ELEMENTS

The following elements are associated with the Technology Architecture Standards area:

Protocols

This element describes the common access protocols used for accessing IPDA information and data. For example, a critical function under development called “PDAP” (Planetary Data Access Protocol) is required in order to provide remote access and discovery services to an online registry of data products and data sets.

Portals

The IPDA provides an online “portal” for accessing information regarding IPDA standards and events. This portal is currently housed at http://www.planetarydata.org .

Registries

This element describes the shared registry services which are used by multiple IPDA members and institution. Registries are catalogs of IPDA service offerings and standard data values that are necessary to enable interoperability. For example, a registry may contain information about services offered within the IPDA (e.g., various access points for getting planetary data from an agency) or it may provide standard data values such as mission names, etc so they are used consistently across agencies.

Archive Tools

This element describes a set of archive tools for generating and validating data by data providers. These archive tools are shared across member agencies and are necessary to improve the quality of the archive. In addition, if tools are shared, it will help to ensure that data is captured in a consistent manner.

5.0 HIGH LEVEL ARCHITECTURE CONCEPTS

5.1 TECHNICAL ARCHITECTURE

This section provides a set of high level architecture diagrams that depict the mapping of both the technical and data architecture views to conceptual models for their use. Details are provided in the sections below.

5.1.1 IPDA SUPPORT FOR THE DATA SYSTEM FLOW

The data system flow shows how tools and software services from the IPDA reference architecture map to a model that supports both the acceptance/ingest of data through to distribution to users. *Archive Tools*, for example, are anticipated to be tools that can support both the generation and validation of data during acceptance and ingest from data producers. *Shared services,* are services that support the sharing of data between IPDA archive systems and to users. Section 5.2.2 shows more detail on data sharing.

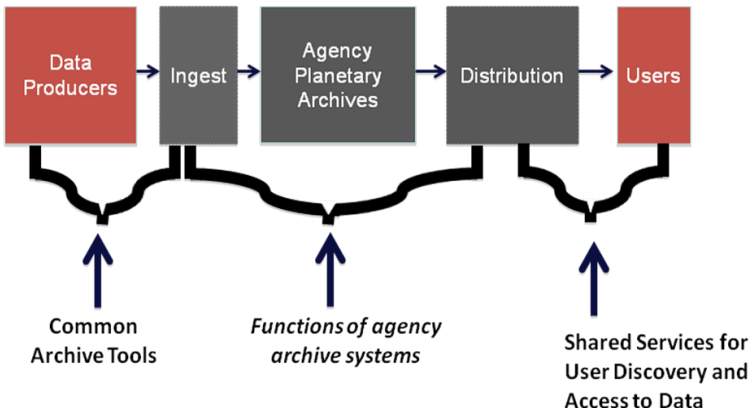


Figure 1: Tools and Services to Support Interoperability Across the Archive Lifecycle

5.1.2 TECHNICAL ARCHITECTURE FOR DATA SHARING

Sharing data across agency boundaries is one of the critical functions of the IPDA. Figure 2 below shows the relationship between agency data systems and the necessary functional capabilities required to share data. The blue circles in the figure indicate *protocols* that are required to support data discovery, access and exchange. The IPDA portal, as discussed earlier, serves as an entry point to link to agency data systems as well as disseminate information (standards, events, etc). The *IPDA Service and Data Registries* refers to registered services and common information that needs to be shared by all member agencies.

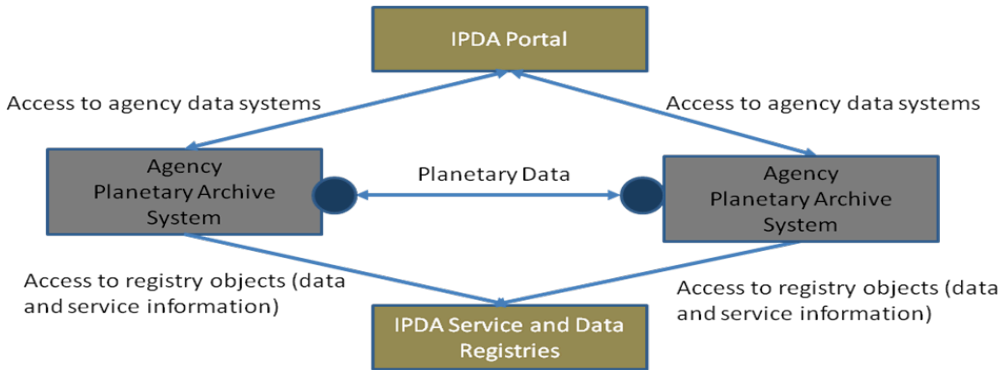


Figure 2: IPDA Technical Architecture for Data Sharing

5.2 DATA ARCHITECTURE CONCEPT

The data architecture concept is built around the notion of an *information object.* The below figure, defined in the CCSDS Reference Architecture for Space Information Management (RASIM) [8], provides a high level view that an information object is composed of a data object and a metadata object. The metadata object is used to describe the data object. This concept is currently used in the planetary community and defined in the PDS Standards Reference [4]) where a PDS “label” is used to describe a set of science data.

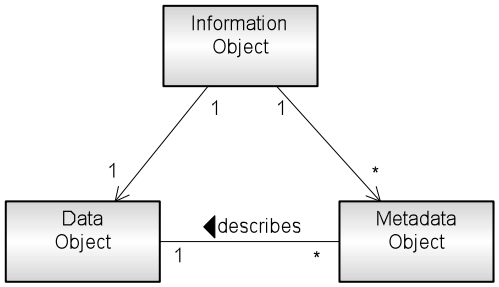


Figure 3: Information Object

An information object is then governed by a domain model. In the case of the IPDA, the domain model is the IPDA Data Model which contains definition of standard data objects as a model (or schema) and the associated attributes captured in a data dictionary. This is depicted in figure 4 below. In practical terms, the instance of a data products (e.g., a PDS Image) is defined by the domain model. The meta model further constrains domain models so that they are built by more general models and standards themselves.

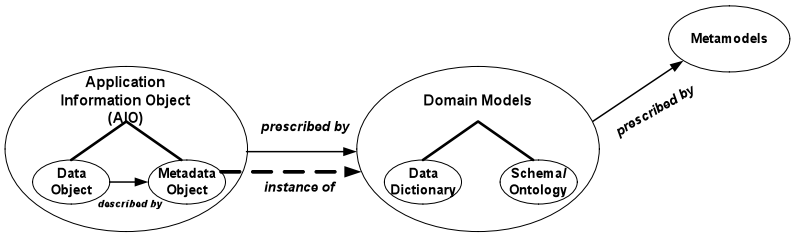


Figure 4: Relationship between models and information objects

6.0 RECOMMENDED IPDA PROJECTS

Project	Requirement	Recommendation
Shared Tools	4.2	Develop a project that identifies tools developed by IPDA member agencies that can be shared. Post results to the website.
IPDA Standards Website	1.2	Develop a set of standards documents/online guides to navigate IPDA standards and processes. Post results to the website
Shared Registry Plan	2.3	Assemble a plan for sharing common data and information (e.g., missions, instrument information, etc) among member agencies. Present plan for

		data and services at the next Steering Committee.

APPENDIX A ACRONYMS

The following acronyms pertain to this document:

TBD