DATA Act Information Model Schema (DAIMS)

Architecture

U.S. Department of the Treasury

October 2, 2018

Table of Contents

1. Introduction

The DATA Act Information Model Schema (DAIMS) standardizes federal spending data terminology, meaning, and format to drive transparent, consistent, reliable, and accurate data for public consumption. The adoption of the DAIMS recognizes spending data as a federal asset. It also requires the creation or refinement of agency data management policies and procedures for effective spending data governance and stewardship.

The U.S. Department of the Treasury issued data standards to provide prescriptive and descriptive guidance, instruction, and technical support for the DAIMS. The DAIMS contains the documents, schema, and other artifacts that express the data standards so that content consumers and content providers know what the data elements mean, how the data elements will be used, and why the data elements are included in the DAIMS. In effect, the DAIMS is the language of the DATA Act.

The DAIMS incorporates feedback from the federal communities and external stakeholders. The consensus building process include working group meetings, feedback periods, and weekly meetings with agency representatives. As the data standard of the DATA Act, the DAIMS is in alignment with the ISO Directive 2 definition for a data standard, "document, established by consensus and approved by a recognized body, that provides, for common and repeated use, rules, guidelines or characteristics for activities or their results, aimed at the achievement of the optimum degree of order in a given context."[1] As a U.S. federal government standard that is in alignment with industry data standardization best practices, the DAIMS delivers:

- Standardized data elements that are complete and reflect DATA Act requirements
- Data elements and structures that are documented consistently, clearly, and accurately
- Extensibility to support future communities of interest, financially oriented domains, and emerging technologies
- Business comprehension of data element content

DAIMS leverages and aligns with the following federal guidance and architectures:

- eXtensible Business Reporting Language (XBRL) – an open international standard for digital business reporting. XBRL enables business reporting to move between organizations in an accurate and digital manner
- U.S. Federal Enterprise Architecture (FEA)[2] – a U.S. Office of Management and Budget (OMB), Office of E-Government and IT initiative that provides an approach for integrating the

---

[1] ISO Directives Part 2, https://www.iso.org/directives-and-policies.html
[2] The Common Approach to Federal Enterprise Architecture,
http://obamawhitehouse.archives.gov/sites/default/files/omb/assets/egov_docs/common_approach_to_federal_ea.pdf

strategic, business and technology management as part of organizational design and performance improvement

- OMB Circular No. A-119 Revised, Federal Participation in the Development and Use of Voluntary Consensus Standards in Conformity Assessment Activities[3] – establishes policies to improve the internal management of the Executive Branch with respect to the U.S. Government's role in the development and use of standards and conformity assessment
- OMB Memorandum M-13-13, Open Data Policy Managing Information as an Asset[4] – a framework for managing information at each stage of its lifecycle to promote interoperability and openness

The U.S. Department of the Treasury established guiding principles to produce holistic and integrated DAIMS artifacts. The guiding principles provide the foundation for the decisions made to craft, assemble, and deliver the DAIMS artifacts. As is the case with industry standards' specifications, the guiding principles don't constrain the development of a data standard. Instead, they help to explain the motivation, judgement, and factors considered in delivering the data standard. The DAIMS guiding principles include:

- Accuracy – Provide consistent naming and precise definition and format
- Adaptability – Enable flexibility to accommodate ad hoc federal spending requirements and unanticipated future requirements
- Completeness – Be comprehensive and consistent; the standard could include all data elements necessary to meet federal spending reporting requirements
- Ease of Use - Design for usability and understandability

2. Conceptual Information Model

A conceptual information model describes information representing an enterprise or a system. Applied to the DAIMS, the conceptual information model is a high-level diagram and supporting descriptive information used to represent and communicate the DAIMS architecture to business users. This section maps the multi-level view of the DAIMS architecture which includes Domains, Components, Elements and Metadata. The relationship between the different levels and the model are visualized in Figure 1.
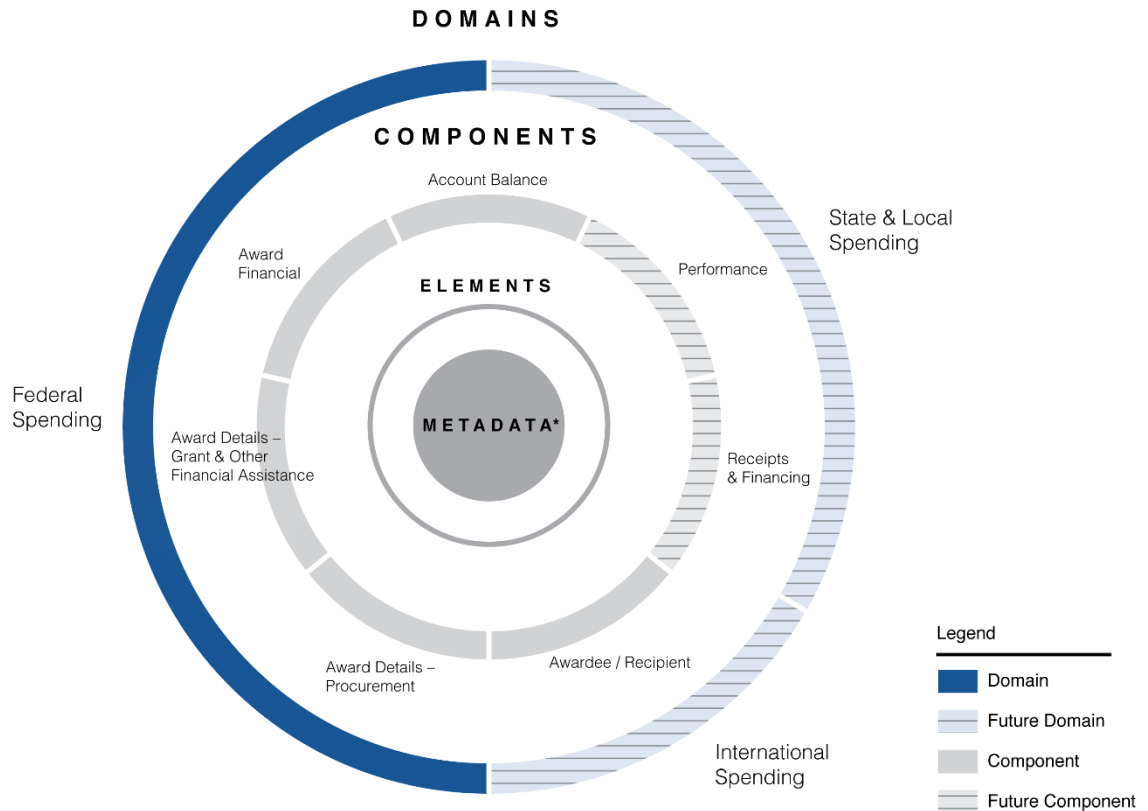
Level 1: Domain Level

Currently, DAIMS is scoped to support financial and spending transparency for the U.S. federal government; the content describes the 360-degree lifecycle of federal spending. DAIMS' adaptability can, dependent upon business requirements, support the reporting of spending data for communities of interest such as state and local governments. DAIMS has the potential to support:

---

[3] https://www.nist.gov/sites/default/files/revised_circular_a-119_as_of_01-22-2016.pdf
[4] https://obamawhitehouse.archives.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf

- State and Local Spending: Data elements and relationships that are unique to state and local jurisdictions, or shared with federal spending

DOMAINS

COMPONENTS

Account Balance

Award Financial

ELEMENTS

State & Local Spending

Performance

Federal Spending

METADATA*

Award Details – Grant & Other Financial Assistance

Receipts & Financing

Award Details – Procurement

Awardee / Recipient

International Spending

Legend
- Domain
- Future Domain
- Component
- Future Component

* *DAIMS metadata leverages ISO/IEC 11179 standard for metadata registry*

Figure 1. DAIMS Conceptual Information Model

- International Spending: Data elements and relationships that are unique to cross-boundary international spending, e.g., U.S. Agency for International Development (USAID), The World Bank, and financial sector data standards

Level 2: Component Level
Within the federal spending domain, the DAIMS components contain a set of related data elements that represent a business function or an entity with a well-defined boundary. Components may share data elements and, dependent upon business requirements, may also be used across different state, local, or international domains. The DAIMS components are:

- Account Balance: Includes federal budgetary balances (obligations, outlays, and payments) with accounts, object class, program activity, and U.S. Standard General Ledger (USSGL)

- Award Financial: Includes federal award level transactions with accounts, object class, program activity, and award identification
- Award Details - Grants & Other Financial Assistance: Contains the award and awardee attributes for financial assistance, including federal sub-award attributes and linkage to prime awards
- Award Details - Procurement: Contains the federal award and awardee attributes for procurements, including federal sub-award attributes and linkage to prime awards
- Awardee/Recipient: Contains awardee and recipient details

Potential future DAIMS content includes:

- Receipts & Financing: Describes federal receipt and financing balances with accounts and sources
- Performance: Describes performance measures and outcomes linked to federal grants, awards or other financial assistance

Level 3: Element Level

A data element is the description of an atomic unit of data that has precise meaning and semantics. DAIMS defines data elements and their relationships, organized by domains and components.

Level 4: Metadata Level

Metadata is the data providing information about one or more aspects of the data or an individual data element. DAIMS metadata leverages ISO/IEC 11179 standard for the DAIMS metadata registry, data element naming and data element definitions. The metadata model defines concepts of the data elements, and provides the domain value enumerations and validation rules.

3. Metadata

Metadata's simplistic definition is data about the data. When applied to DAIMS, metadata represents a federal spending data element's name, definition, type, size and allowable values. Metadata provides the foundation for building consensus and establishing shared meaning. For example, it ensures that data elements representing a procurement award from one federal agency is easily understood, equivalent in meaning, and comparable to a procurement award from another federal agency.

The DAIMS metadata registry is an information system for the registration of metadata that defines and describes federal spending data elements. The metadata registry and its registered data aligns to the ISO/IEC 11179 Information Technology – Metadata Registries (MDR) standard. ISO/IEC 11179 provides guidance for standardizing data element names and descriptions to achieve

common understanding of the data across organizations and organizational boundaries. It also supports the ability to harmonize data elements with equivalent meaning and representation, and promotes data and data component re-use over time, space, systems and applications. ISO/IEC 11179 includes:

Standard	Description	DAIMS Applicability
ISO/IEC 11179-3[5]	Part 3: Registry metamodel and basic attributes - describes the attributes of data elements and associated metadata to be specified and registered as metadata items in a metadata registry (MDR)	DAIMS Metadata Registry
ISO/IEC 11179-4[6]	Part 4: Formulation of data definitions - specifies the requirements and recommendations to formulate data definitions that registered in a metadata registry	DAIMS Metadata Registry
ISO/IEC 11179-5[7]	Part 5: Naming principles - describe and specify common features of naming conventions	DAIMS Metadata Registry

The additional metadata attributes captured by the registry includes:

- Domain: Identifies the domain / business line where the element applies
- Data Element Label: A unique label for each element
- Data Type: String, date, or number
- Max Element Length: Indicates the maximum length of the data element value
- Documentation: Contains the business definition of the data element
- Element Use: Indicates whether a data element is an extracted value, required, optional, or conditionally required per the validation rules
- Element Number: A unique identifier for each element
- Enumerations/Domain Value: Provides information on potential input values for each data element
- Example Value: Demonstrates an acceptable data entry value
- Submission Instructions: Additional information for reporting the element
- Validation Rule: Includes validation applied to the relevant element when data is submitted

---

[5] http://standards.iso.org/ittf/PubliclyAvailableStandards/c050340_ISO_IEC_11179-3_2013.zip
[6] http://standards.iso.org/ittf/PubliclyAvailableStandards/c035346_ISO_IEC_11179-4_2004(E).zip
[7] http://standards.iso.org/ittf/PubliclyAvailableStandards/c060341_ISO_IEC_11179-5_2015.zip

4. Specifications

Normative artifacts define the prescriptive requirements that are necessary to conform to the DAIMS. Informative artifacts assist in providing an understanding of the DAIMS and how it is leveraged to meet the intent of the DATA Act. Informative artifacts do not outline conformance requirements.

The DAIMS normative artifacts within are maintained by internal change control procedures and follow a standard change management process for all changes. Informative DAIMS documents may undergo change as the DAIMS normative artifacts evolve, but they do not necessary follow a standard change management process. Normative artifacts within the DAIMS architecture include:

1. Schema Specifications
    a. Reporting Submission Specifications (RSS): includes a listing of the data elements with specific instructions for federal agencies to submit content in the appropriate format. It is a human-readable version of the data standard for agency submitted content
    b. Interface Definition Documents (IDD): contains a listing of the elements, with supporting metadata to understand data pulled from government-wide systems for procurement and from agency financial assistance systems. It's a human-readable version of the data standard for content extracted from external data sources
2. Validation Rules: verify data elements based on their values, metadata, and related elements. There are severity levels associated with the validation rules, such as error and warning. An error invalidates the data record based on the schema. A warning indicates a potential data quality issue, even though the record is valid based on the schema
3. XBRL Taxonomy: a machine-readable version of the data standard consisting of:
    a. Packages: identifies collections of submitter and consumer data. A package defines pertinent elements
    b. Elements: individual data points expressed as XBRL facts and application programming interface (API) endpoint properties.
        - Element properties: names used to identify data, accuracy, data type and data type facets (e.g., length, pattern, enumeration restriction)
        - Labels: human-readable descriptive labels, code constrained-length labels, user interface specific labels and documentation definition label
        - Properties: implemented as Reference Parts. Have code-consumable breakdown of specification attributes, submission attributes, and API mapping

Informative artifacts within the DAIMS architecture are meant for informational purposes and include diagrams and sample files.

5. Conformance

DAIMS conformance validation is handled external to the DAIMS. DAIMS data validation is handled through an open source application called DATA Act broker (https://github.com/fedspendingtransparency/). The DAIMS XBRL schema also contains metadata that can be used for validating the agency submitted data.

6. Schema Versioning

DAIMS data standard will continue to evolve to support spending data reporting requirements, USAspending.gov website needs, and additional domains and components. Schema versioning is important to manage change, gain consensus, and support the systems using the current standard. DAIMS versioning includes major release, minor release, and maintenance release.

Major Release - Major releases are expected to be impactful to most or all of the users and typically necessitate all users to upgrade to the latest release. Major releases will follow a 1.x, 2.x, etc. naming convention. Major release will typically occur on a twelve-month cycle.

Minor Release – Minor releases only affect part of the user community and therefore perhaps not necessary for everyone to comply (e.g., a new feature or bug fix that is only appropriate for a limited number of use-case scenarios). Minor releases will follow a x.1, x.2. etc. naming convention. Minor release will typically occur on a six-month cycle, following a major release.

Maintenance Release – After finalization of a major or minor release, and in some cases before, there is an expected period of discovery where the exercise of a standard in practice uncovers issues with the documentation. Maintenance releases include errata, addendum, and appendix releases.

An errata releases is a set of changes to a release that are not substantive in nature. In practice, it means correcting errors that make the documentation more accurate, but do not materially change the intended or actual implementation of the standard. When more material changes are required or new capabilities/scope of the standard is desired, then the process cycle is initiated again for the next targeted major or minor release.

An addendum is used to provide incremental expansion (as opposed to revision) of a standard. Addendums are often used to allow for policy changes and the modular addition of ancillary or complimentary standards (such as the support for an alternate submission).

Appendices are typically additions to the standard, similar to an addendum, but generally are primarily intended to provide a deeper level of documentation support to the core of the standard. Appendices might provide super-user details of facilities that are already available but under-documented.

7. Conclusion

The DAIMS architecture is designed to meet the needs of business users and technical systems. It serves as the foundation for the DAIMS data standard. It is extensible through domain, component, element, and metadata. The architecture defines the DAIMS specifications as normative documents and how the schema will be versioned. Additional descriptive content on the DAIMS architecture will be released as necessary.