



United States Department of the Interior

OFFICE OF THE SECRETARY
Washington, DC 20240

AUG 22 2008



OCIO DIRECTIVE 2008-020

To: Heads of Bureaus and Offices

From: Michael Howell
Chief Information Officer

Subject: Designation, Management, and Enforcement of Authoritative Data Sources

Purpose:

This directive establishes the policy for:

- officially designating, maintaining, and using authoritative data sources (ADS); and
- maximizing re-use of a designated ADS in architecture and capital planning and investment control (CPIC) processes.

This policy is to be integrated with the CPIC process as outlined in the Department of Interior (DOI) CPIC Guide. The Investment Review Board (IRB) will consider the use and applicability of designated ADSs as part of the CPIC Pre-select and Select processes.

Policy:

1. The policy for official ADS designation and management:

1.1 Overview of the ADS Designation and Management Method Phases

A three-phased approach to officially designate and implement an ADS for the DOI enterprise shall be followed as illustrated in Figure 1 on the following page and described below.

Phase 1 - ADS Candidate Designation. The methodology used for blueprint or segment architecture development will have a core team or designated working group that conducts a comparative analysis of all systems using data of the same or similar theme or information class to ensure information requirements from known consumers or potential consumers are or can be accommodated in the designated candidate ADS. They will document the results of this analysis in the ADS Assessment and Candidate Designation Report. The results will reflect a recommendation to the blueprint or segment architecture core team for approval of the data source to be designated as a candidate ADS.

Phase 2 - ADS Acceptance. The bureau/office owning the candidate ADS will have 60 to 90 days to review the core responsibilities, identify gaps, define a funding strategy, establish performance metrics, develop a roadmap for transition, and submit this information in a Bureau ADS Acceptance Report. The bureau/office will submit their report to the blueprint or segment architecture core team who in turn will seek initial approval from the Data Advisory Committee (DAC) on the designated candidate ADS. The E-Gov Team will provide the decision to accept or reject the approval of the DAC, and the IRB will ratify the E-Gov approval decision.

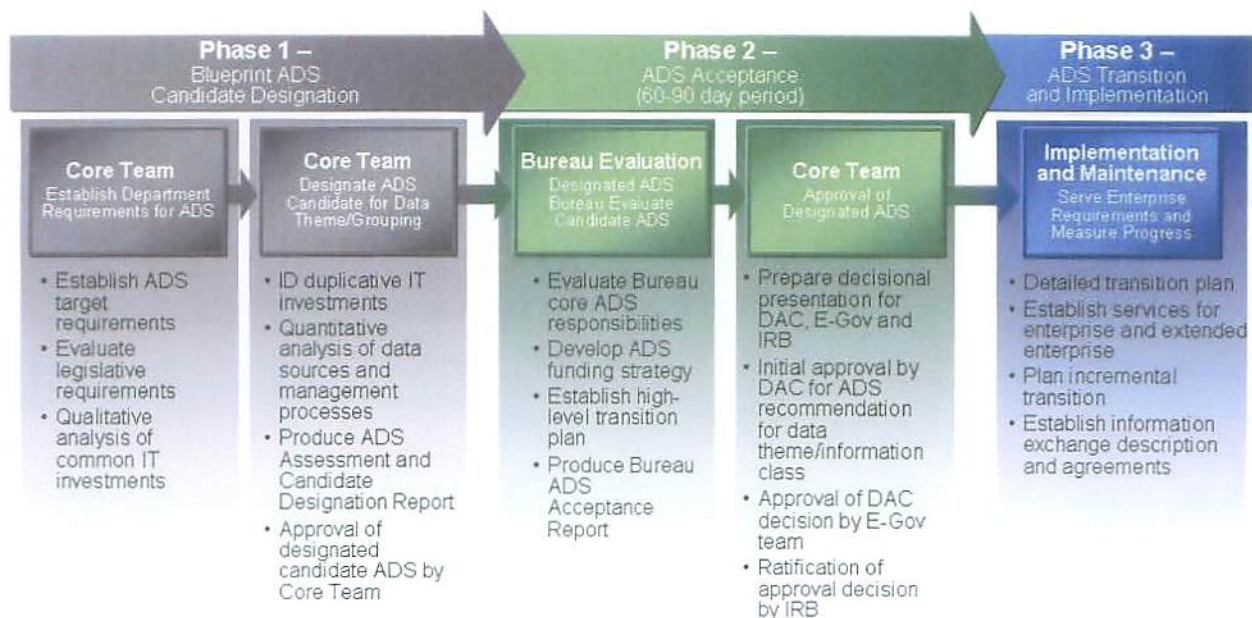


Figure 1: ADS Designation and Management Three-Phase Approach

Phase 3 - ADS Transition and Implementation. The bureau/office owning the candidate ADS will develop a detailed transition plan that describes how legacy data sources will be migrated to the officially designated ADS. The owning bureau/office will establish shared enterprise information sharing services, and describe information exchange requirements and service level agreements (SLAs) for internal and external information sharing partners.

1.2 ADS Core Responsibilities

Upon approval by the IRB at the end of Phase 2, each ADS shall have a steward appointed by the system owner to coordinate and/or perform the following responsibilities subject to existing federal and departmental policies and procedures:

- Managing the data contained in the ADS as an independent asset in accordance with DOI DM Manual Part 378, Data Resource Management policy and associated guidance documents.
- Coordinating, at minimum, with appropriate Principal Data Steward, Bureau Business Data Steward(s), Bureau Data Architects, and system owners.

- c. Managing the quality of data in accordance with the DOI Data Quality Management Guide. The use of the DOI Data Quality Management Guide to include the Guide's optional process phase for certification is mandatory for all officially designated ADSs.
- d. Planning, documenting, and executing the information sharing services in accordance with Chapter 5 of the Federal Enterprise Architecture (FEA) Data Reference Model (DRM) version 2.0.
- e. Determining the data and information appropriate for sharing and the desired level of discovery (e.g. discovery of a database or a record within a database; discovery of a document or a paragraph within a document).
- f. Determining appropriate levels of access controls.
- g. Enhancing sharing opportunities among information consumers by providing multiple formats, as requested.
- h. Ensuring data is timely and accessible to information consumers.
- i. Adopting additional technology and data standards, as appropriate.
- j. Coordinating with information consumers and system developers to collectively work to achieve optimum usability, understandability, and interoperability of the data without violating privacy, security, or other policies. ADS Stewards shall be responsive to priorities for data to be discovered and shared.
- k. Enabling information consumers to provide feedback to ADS Stewards who take appropriate action in response to consumer feedback.
- l. Developing metrics to measure consumer satisfaction. Outcome metrics should focus more on consumer value (e.g. "decrease in consumer decision-making time") and not counting metrics (e.g., "how many data sets were posted").
- m. Coordinating with the National Archives and Records Administration (NARA) to ensure appropriate management of records (Federal Records Act 1950). The context of an ADS is the lifecycle and management of data while the context of a system of record (Privacy Act 1974) is the lifecycle and management of the official record to include the management of personal identifiable information.

1.3 New ADS Recommendations and Approvals

The blueprint or segment architecture core team process for new ADS recommendation approvals shall include:

- a. Coordinating with the Principal Data Stewards, the business owner, the system owner, and other appropriate stakeholders and governance teams to make the initial ADS recommendation.
- b. Seeking initial DAC approval of the recommendation.
- c. Requesting an E-Gov Team decision to accept DAC approval of the recommendation, and
- d. Gaining ratification of the DAC approval decision by the IRB.

1.4 Governance for Officially Designated ADS

The data within an ADS are enterprise assets with value to multiple consumers. When a DOI ADS is established and matures within the organization, a governance structure will be required to systematically identify, evaluate, and prioritize the requirements. These requirements will drive the evolution of the data content in associated services for the ADS. This governance structure will provide access and influence for the business constituency and stakeholder communities to mature the ADS in an enterprise value-based management approach.

The governance structure provides the opportunity to identify common needs and investigate common solutions prior to initiating new investment development efforts. The ADS Steward, appointed by the system owner to perform the core responsibilities defined in Section 1.2 of this policy, and the owning bureau/office will be members of the governance team. The ADS governance activities should work in conjunction with existing Communities of Interest (COI) or develop them to improve communication and outreach and ensure access to subject matter experts. The governance model should include the Principal Data Stewards for subject areas and information classes under their charge while integrating the respective business data stewards from the bureaus. The DAC should provide support to the ADS governance efforts with analysis, methods, policy and standards guidance and support. Any resulting capital planning investments will be subject to existing policy and guidance.

2. The policy for ADS enforcement:

- a. Through ADS integration with the CPIC cycle, all new investments under CPIC will review the availability of current official ADSs that are published on the Interior Enterprise Architecture web site located at <http://www.doi.gov/ocio/architecture/index.html>.
- b. New investments shall not create data when that data currently exists and are available for sharing in an existing official ADS.
- c. When IT investments are prioritized through the DOI CPIC investment boards, the intention is to fund an officially designated ADS at a higher level of priority and divest in other redundant data sources.
- d. Approved blueprint system consolidation projects shall identify the disposition of legacy data impacted by consolidation recommendations. Legacy data will be analyzed, standardized, and migrated to the officially designated ADS as deemed appropriate.

3. The policy for waiving the mandatory use of an ADS:

There may be instances when a new investment in the CPIC process might require the use of a data source that has not been designated as authoritative, even when an officially designated ADS for the required data theme or information class exists. The following actions shall be taken to adjudicate such instances:

- a. All new investments shall submit a waiver to the DAC for review and guidance when an official ADS exists but does not appear to satisfy the new investment's information requirements.
- b. The DAC will coordinate a collaborative review with appropriate governance bodies, the appropriate Principal Data Steward(s), the business owner, and the blueprint or segment architecture core team to determine if the new investment's waiver is justified, and will provide recommendations to the IRB to approve the waiver.

Scope:

This directive applies to and includes:

- all DOI bureaus and offices.
- all DOI LoB and E-Gov initiatives, and
- all DOI investments in information system components using shared information and data which are under CPIC review for planning, development, maintenance, and integration initiatives. These investments include all work performed by DOI bureaus/offices, contractors, universities, private sector organizations, and other government (federal, state, and local) agencies funded by DOI.

Background:

A key factor in furthering DOI's information sharing and interoperability goals is in the use of officially designated ADSs. An ADS is a single, officially designated automated source that provides trusted, timely, secure data on which lines of business (LoB) rely. Trusted data means that the data providers exercise management responsibility for appropriate practices, procedures, and processes to produce data that is within acceptable thresholds for quality, integrity, and security. These thresholds are typically determined by the principal and business data stewards. The intended outcome of using an officially designated ADS is to provide information that is visible, accessible, understandable, and credible to information consumers that include DOI business users, DOI information exchange partners, and IT applications and services. The assessment and official designation of ADSs is accomplished through analysis and recommendations that are documented in but not limited to Modernization Blueprint projects, business process reengineering projects, and E-Government (E-Gov) related projects.

The ADS three-phase designation should be part of the development of modernization blueprints, also known as segment architectures. When developing modernization blueprints, the ADS perspective considers analysis of data assets and management processes to ensure designated candidates:

- support business goals and intended outcomes that are enterprise-wide or cross program in scope;
- provide data quality, data stewardship, data standardization, and data re-usability;

- provide services and metadata required to be trusted by information consumers;
- provide SLAs, as appropriate, and
- obtain IRB ratification of ADS approval decisions.

The DOI intends to invest in any ADS that is officially designated and divest redundant data sources through the CPIC process.

The main objectives for the use of an officially designated ADS include:

- improving decision support for investment planning, programmatic direction, resource usage, and policy development;
- improving data re-use, exchange, and consolidation of redundant data;
- providing information consumers with the ability to discover, understand, and access data previously not available to them;
- providing a foundation for an integrated, business-driven information sharing solution that is tied to DOI's performance goals; and
- improving the quality of data and information used to support business processes and decision making.

The data in an officially designated ADS will continue to be managed by the appropriate steward in accordance with DOI DM Manual Part 378, Data Resource Management Policy and the OCIO Directive 2006-11 Data Standardization Procedures. Core responsibilities for managing an ADS are established in Section 1.2 of this directive. If the designated ADS is also a System of Record (Privacy Act of 1974), the responsibilities for records management (Federal Records Act 1950) are not superseded by the ADS designation.

Contact:

If there are any questions on this directive may be referred to Suzanne Acar, Office of the Chief Information Officer at 202-208-3216 or e-mail: suzanne_acar@ios.doi.gov.

Attachment: Glossary

cc: Bureau and Office Chief Information Officers
Interior Architecture Work Group
Data Advisory Committee

Glossary

Accessible — Data are made available such that information consumers and applications can access them, except when limited by policy, regulation, or security.

Authoritative Data Source (ADS) — An ADS is a single, automated data source that has been officially designated and ratified by the IRB to provision data or information for a DOI data theme or information class.

Community of Interest (COI) or Community of Practice (COP) — A COI is a collaborative group of users who require a shared vocabulary to exchange information in pursuit of common goals, interests, and business objectives. They may include Lines of Businesses (LOBs) within the government and external organizations that are dedicated to the support of business functions.

Cross-cutting — In the context of this policy, cross-cutting describes a behavior of data or information services in which the data span across or are used by multiple organizations or Lines of Business (LOBs).

Data — A value or set of values representing a specific concept or concepts. Data becomes “information” when analyzed and possibly combined with other data in order to extract meaning and to provide context.

Data Asset — A managed container for data. Examples include a relational database, web site, document repository, directory, or data service.

Data Discovery — The process of discovering data that exists within an enterprise which includes locating a resource, and using a process that exploits metadata descriptions (such as a search engine) to obtain knowledge of information content.

Data Steward — A person or group that manages the development, approval, creation, and use of data within a specified subject area or information class, and ensures that data standards can be used to satisfy requirements throughout DOI. DOI Data Stewards include:

- The ***ADS Steward*** is appointed by the system owner to maintain an officially designated ADS. The ADS Steward is responsible for the content of the data asset and for performing the ADS core responsibilities defined within Section 1.2 of this policy.
- The ***Business Data Steward*** coordinates the implementation of new data standards with subject matter experts (SMEs) and database administrators (DBAs) in systems supporting a Line of Business (LoB), and verifies data security and data quality requirements have been met for each data standard.
- The ***Principal Data Steward*** coordinates the creation or review of proposed data standards with all Business Data Stewards and Bureau Data Architects for a LoB.

maintains current DOI data standards for their respective LoB, and submits proposed data standards to the DOI Data Architect for formal cross-functional review.

Data Theme — A collection of one or more data objects that describes a particular subject matter for a given area. For example, in a GIS there may be several separate data themes representing streams, land parcels, and so forth. Each theme may consist of many data objects representing individual geographic entities.

E-Government — The federal government's use of information technologies to exchange information and services with citizens, businesses, and other arms of government.

Enterprise — A complete business or organization consisting of functions, divisions, or other components used to accomplish specific objectives and defined goals.

Federal Enterprise Architecture (FEA) — A tool that enables the federal government to identify opportunities to leverage technology and alleviate redundancy or to highlight where agency overlap limits the value of IT investments. The FEA facilitates horizontal (cross-federal) and vertical (federal, state, and local government) integration of IT resources, and establishes the "line of sight" contribution of IT to mission and program performance. The outcome will be a more citizen-centered, customer-focused government that maximizes technology investments to better achieve mission outcomes.

Information — Data with context; the meaning that an information consumer assigns to data by means of the known conventions (e.g. metadata) to access and use data to perform business functions.

Information Class — An information class categorizes and provides context to a group of data entities that support a business functional area.

Information Consumer — Information consumers search catalogs of data to discover data assets that are available. They access data from shared spaces based on understanding of what data is available based on its context and structure as described using metadata.

Information Exchange — The process of sending and receiving data in such a manner that the information content or meaning assigned to the data is not altered during the transmission.

Interoperable — The capability for one system, and/or its components, to efficiently and effectively exchange information, data, and logic with one or more systems. Metadata is available to allow mediation or translation of data between interfaces, as needed.

Line of Business (LoB) — The middle tier of the Federal Enterprise Architecture (FEA) Business Reference Model (BRM). LoBs represent the internal operations of the federal government and its services for citizens, independent of the agencies, bureaus and offices that perform them.

Metadata — A term used to mean data that describes data. The term metadata is used to define all the dimensions of data required to build databases, services, and applications for information producers and consumers.

Modernization Blueprint — A transition plan for migrating from the as-is architecture to the target architecture typically specific to a business or functional area.

Owning Bureau/Program Office — The owning bureau/office is the organization within DOI that is responsible for the content of the ADS and for ensuring the ADS core responsibilities defined within Section 1.2 of this policy are performed.

Principal Data Steward — See Data Steward.

Qualitative Analysis — Subjective assessment of data gained by surveying information stakeholders and consumers.

Quantitative Analysis — Numerical measurement of data content based on various data quality dimensions (e.g. accuracy, timeliness, completeness).

Record of Decision (ROD) — A public document that explains the reasons for a project decision and summarizes any mitigation measures that will be incorporated in the project.

Source — In context of this policy, a source is an automated data source where data are stored and can be obtained, such as a data base or data file.

Subject Area — A collection of data classifications that represent broad categories of information. These major areas of information and data subjects support the business lines of an organization.

System of Records (SOR) — As stated in the Privacy Act of 1974, a system of records is “a group of any records under the control of any agency from which information is retrieved by the name of the individual or by some identifying number, symbol, or other identifying particular assigned to the individual.”

Trusted — Information consumers and applications can determine and access the authority of the source because the pedigree, security level, and access control level of each data asset is known and available.

Understandable — Users and applications can comprehend the data, both structurally and semantically, and readily determine how the data may be used for their specific needs.

Visible — Users and applications can discover the existence of data assets through catalogs, registries, and other search services. Data assets are advertised or “made visible” by providing metadata which describes the assets.