An Information Architecture Strategy

John Wunder and Dr. Will Tracz
Lockheed Martin Mission Systems

Information architectures play a strategic role in facilitating the efficient and effective storage, retrieval, and analysis of data in enterprise information systems. An information architecture consists of more than just data and the commercial off-the-shelf products in which it is stored. To provide the right information, in the right format, to the right person, at the right time, and to protect that information from unauthorized access, a robust information architecture strategy is necessary. This is also essential to establish data stewards, governance boards, and metrics to support the processes that describe sequences of information services that access the stored data.

This article describes an information architecture strategy[1], which by definition focuses on information, not data. Therefore, the scope of the architecture strategy extends beyond the physical architecture consisting of data, databases, data warehouses, and data marts, to the operational architecture focused on the processes that turn the data into information and facilitate the generation and manipulation of that information. The overarching goal of this strategy is to provide the end user with access to timely, accurate, and trusted information.

In the case of the Global Combat Support System-Air Force (GCSS-AF), the information architecture provides the tools to turn data from a myriad of systems into information, from information into knowledge, and from knowledge into power. This results in information superiority and reduced decision cycles across all Air Force echelons and operational theaters.

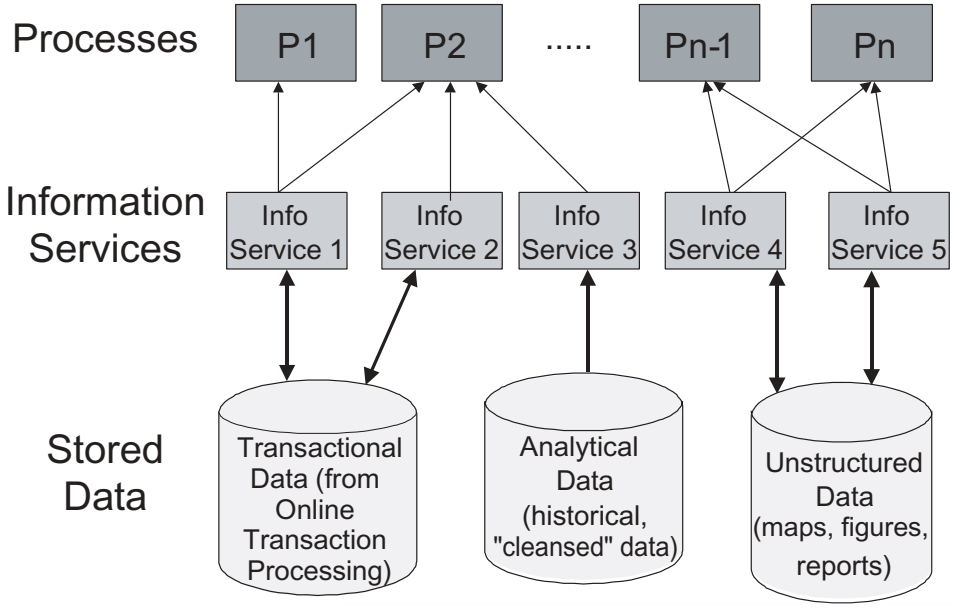
An information architecture provides easy storage, access, and retrieval of information. For example, it enables authorized airmen to access, from any location, such diverse data as the budget information for Wright-Patterson Air Force Base, the engine status at Lakenheath Air Base, or all the personnel deployed in an aerospace expeditionary force, all through a Web browser connected to the Department of Defense infrastructure.

This article first describes the key terms *process* and *information services* and then defines the kinds of data that are stored in the various components that make up the information architecture. Next is an overview of components found in information architecture with special focus on the repository and integration services. Finally, the roles and responsibilities of those who govern the information architecture are described.

Terminology

Operationally, from an information technology (IT) perspective, an information architecture supports *processes* that describe sequences of *information services* that access *stored data*. This relationship is illustrated in Figure 1.

Figure 1: *Relationship Between Processes, Information Services, and Stored Data*



Processes are sequences of operations. For example, at a data/information level, processes might include Find, Fix, Target, Engage, and Assess. At the decision-support level, processes might include Readiness, Crisis Action Planning, Deployment, and Employment/Sustainment. Processes invoke information services in a specific order (i.e., in a sequence that becomes an operational flow).

Information services provide intuitive access to information within the information architecture such as domain-relevant information on such things as aircraft, munitions, personnel, or bases. Information services also provide information on domain-specific capabilities such as sortie generation capacity, fuel consumption, skill training and certifications, or runway capacities. Information services apply domain rules, semantics, and syntax to convert raw data into information.

Two information services examples follow. The first illustrates the difference between raw *data* and decision-support *information* provided through an information service. The second highlights the data cleansing or scrubbing capabilities of the information architecture.

Example 1: Munitions

In a combat situation, many users (e.g., from the joint commander, to the unit level commander, to the munitions supply clerk) need to know the *capacity* of an ammo dump. However, for safety reasons munitions are not active in the ammo dump but are built when needed. If direct access to munitions *data* were granted to end users, then the lack of munitions domain knowledge could lead to confusion rather than useful, decisional knowledge. Therefore, munitions information service could provide accurate *information* to authorized users

by applying the appropriate domain rules and semantics associated with the munitions data resulting, for example, in the required information about the munitions *capacity*.

Example 2: General Ledger

All program managers need access to current expenditures. Accounting data is generated by many sources, potentially in many different formats. A general ledger information service assures that only complete, accurate, and audited expenditures of every budget line item are available to the user by eliminating conflicting or incomplete data (i.e., via data cleansing or scrubbing technology).

Information services accesses *stored data* to generate information. Generally, there are three types of data stored in information architecture. These types of data, as illustrated using the previous example in the general ledger information service, are as follows:

- **Transactional data.** These are day-to-day operations recorded as they occur. In the general ledger example, this is the recording of the thousands of day-to-day expenditures required to run an enterprise.
- **Analytical data.** These are historic transactional data that may be time-stamped and stored immediately after an event occurs, or are recorded later. Furthermore, the format of the data may be different than its originally recorded format. Analytical data are used to determine trends and make predictions. In the general ledger example, the previous year's expenditures would need to be recorded for later analysis to determine future budget requirements.
- **Unstructured information.** These are stand-alone documents that provide direct guidance/information for a task. These data include maps, weather reports, pictures or, for the general ledger example, expenditure guidance from Congress or the Department of Defense.

Information Architecture in Context

Advances in IT in the areas of workflow and design notation/specification standards (e.g., Business Process Modeling Language and Unified Modeling Language) have reached the point where it is possible for mission processes to be composed, automated, and executed by end users without requiring IT professionals or traditional application development. In addition, IT Web services[2] [1] have simplified the creation, invocation, and aggregation of domain-specific information services (e.g., aircraft, munitions, inventory, or budget).

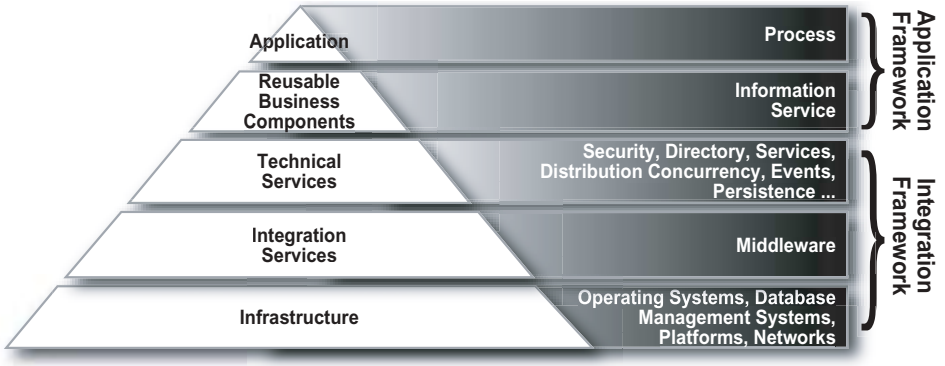


Figure 2: *GCSS-AF Architecture, Frameworks (2), Layers (5), and Example Technologies*

In the Air Force, these advances allow warfighters to sequence and aggregate processes and information services to resolve unique, unanticipated situations as they occur. For example, the Regional Support Squadron (RSS) at Air Mobility Command may determine that they are having an unacceptable number of tire failures; however, no application was written to determine whether the tire failure was due to a particular lot of tires, type of aircraft, or runway. The information architecture allows the RSS member to easily use the Supplier, Aircraft, and Runway and Tire Failure Information Services to isolate the root cause of the failures and execute appropriate corrective actions without any new system or code development.

Finally, current IT standards and mechanisms for representing data patterns allow a clear separation between information creation and usage and data storage and retrieval. This allows implementers of core information services to use common, centrally administered and managed commercial tools to store and retrieve the data. These are the cornerstone concepts of information architecture strategy and enable the delivery of accurate, timely, and trusted information to the warfighter.

Information architecture is part of an enterprise information system's physical architecture. For example as shown in Figure 2, the GCSS-AF architecture consists of two frameworks comprising five layers. While Application Framework contains Air Force Information Services, the Integration Framework consists of information technology (commercial products) that enables integration and supports the capabilities in the Application Framework. Thus, in this case, the focus of the information architecture is primarily on the Application Framework. Processes form the top layer of the Application Framework and invoke information services that reside in the second layer. The data store services are found in the Integration Framework (i.e., in the data warehouse or operational databases).

Key Building Blocks

As shown in Figure 3 (see page 6), information architecture consists of the following components:

- A technology infrastructure that is comprised of the following:
  - Processes targeted to support the enterprise business processes (or in the case of the Air Force, Air Force missions).
  - Information services developed to support business-/mission-specific processes.
  - Data store services that are part of the enterprise infrastructure and consist of the following:
    - A repository that is a directory for looking up and calling processes and information services via the Web.
    - Online transaction processing (OLTP) databases for recording day-to-day transactions.
    - An enterprise data warehouse (EDW) for storing data that can be used for trend analysis and prediction.
    - Extraction, translation, and load (ETL) technology to move bulk data from databases to the EDW.
    - Enterprise application and integration (EAI) technology to provide data conversion services (wrappers) for current applications into information services and processes.
    - A portal to present browser-accessible, mission-focused information.
- An organizational infrastructure comprised of leaders (who define strategy/vision and performance metrics), stewards (who ensure information accuracy), and operational experts (who define process requirements and needs).

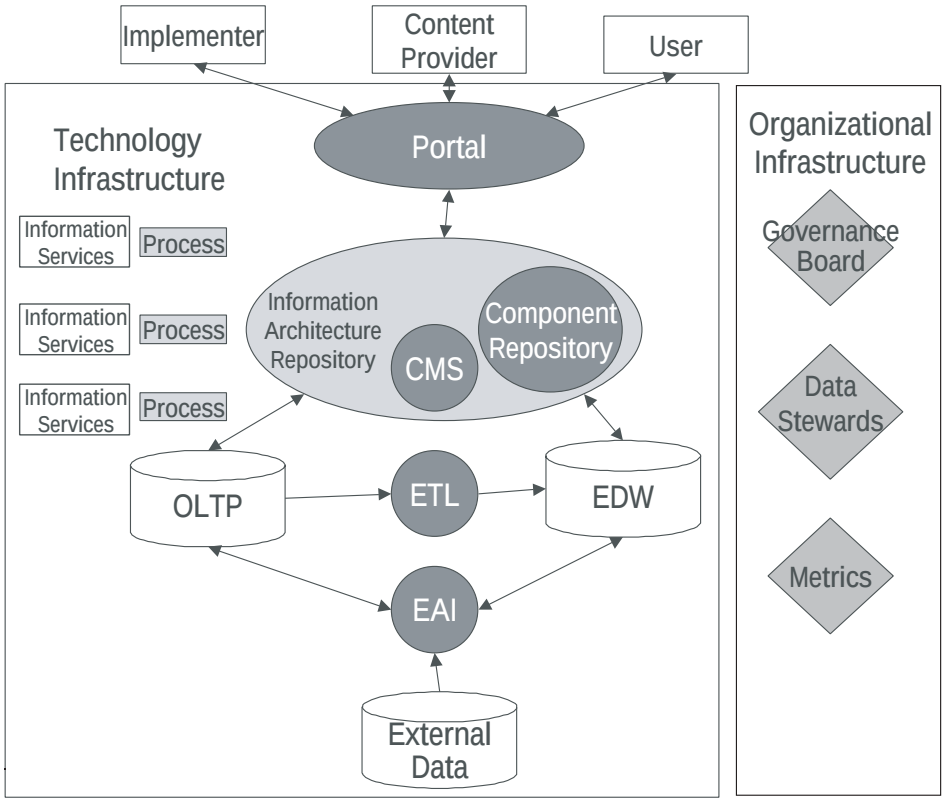


Figure 3: *Information Architecture Strategy Components*

For the Air Force, the information strategy leverages the GCSS-AF integration framework technologies while providing well-defined governance for the application framework implementation.

Information Architecture Repository

The information architecture repository may be thought of as the *phone book* of the information architecture. For example, to make a connection to a process or an information service requires the following:

1. Look it up in the repository.
2. Use the resulting details for contacting the process or information service, including the information source, quality, format, and address.
3. Place the required calls to accomplish the mission.

Having the end user perform the above sequence of steps in his or her *operational* environment to create a new automated task removes the user's need to specify requirements for developers to implement the application in a traditional development environment. To grow these information architecture capabilities, end users or IT implementers must register new processes and information services in the repository providing all the characteristics required to enable automatic activation.

The information architecture repository is composed of two pieces: a content management system (CMS) for content providers to manage unstructured data, and a component repository for content providers and implementers to manage structured data. The information architecture enforces a common information model (common categories and data attributes) on the CMS and the component repository. Standards such as the Dublin Core Standard for Metadata [2] (data about data) and the Universal Description, Discovery, and Integration (UDDI), when coupled with the domain-specific data classification categories of information architecture form the baseline data characteristics and assure common and easy access.

This dual repository defines, enables, and enforces the information architecture. The repository is not just a design artifact but also the production information access tool. Through registration in the repository, processes and information services are made available to authorized users within and without the enterprise. The repository enables information discovery, exchange, and flow.

Integration Services

Figure 4 illustrates the flow of data through existing information architecture integration services (i.e., ETL and EAI) that support the composition of mission-specific information services.

ETL

Extract, transform, and load technologies are used to cleanse and normalize data. When stored in the data warehouse, these cleansed data allow for trending, aggregation across disparate sources, and general analysis of the data to support planned and unplanned investigations. The metadata-driven ability of ETL tools to enforce data definition is of particular importance to information architecture. Data extraction methods supported by the information architecture allow the timely extraction of newly created or updated data and can support scheduled as well as event-driven data-extraction operations. Efficient data extraction and filtering can be performed from databases and flat files without requiring access to application source code. This can also be performed without significantly affecting end-user response time.

EAI

Many large enterprises have an extensive and effective information infrastructure in place. Unfortunately in many cases, this legacy information infrastructure was optimized for point solutions for segregated users (i.e., what has been referred to as *stovepipes*). EAI allows an enterprise to leverage the value of the existing enterprise stovepipe applications. In particular, EAI can wrapper current systems to appear as information architecture processes and information services, thereby garnering the enterprise information benefits without rewriting the current systems.

EAI initially will publish event information to the information architecture making it available to any authorized subscriber. As the information architecture matures, the EAI capability can be used to drive mission execution processes. For example, an event such as the reassignment of airmen will trigger such processes as user account updates, base housing activation, command notifications, or chaplain visits.

EAI is well suited to the publish/subscribe philosophy where data is cleansed and transformed into enterprise information at the source, in near real-time, on an event basis. EAI tools can be coupled with transaction monitoring tools capable of generating alarms when data or operational problems occur. Near real time information transfers and performance monitoring are obvious advantages in environments where the reliable delivery of information is important.

Roles and Responsibilities

For any information architecture strategy to succeed, it must not only provide explicit guidance for the process and information service implementers as well as the content providers and end users, but also must pro-

vide oversight in establishing common goals and metrics, usually in the form of a governance body. In the Air Force's information architecture strategy, it is essential that the governance body consists of several *advocates* : 1) a doctrine operations advocate, 2) a technical systems advocate, and 3) an integration advocate, with doctrine advocacy taking precedence. The steward of the information must be the doctrinal owner of the operation. In order to institutionalize this process, an enterprise must support an information management board (IMB), which should comprise three members:

1. The chair is the transformational leader of the doctrinal operations and is responsible for all decisions, assuring that the operational architecture reflects current doctrine and concept of operations.
2. The information technology lead of the implementation (generally appointed by the chief information officer) is responsible for the system and technical architecture, assuring the operational architecture is implemented.
3. The integration scribe (appointed) is responsible for implementing and encoding the architecture in the GCSS-AF Integration Program.

The IMB is responsible for implementing the following:

1. The information architecture.
2. The definition of metrics and measures of effectiveness (MOE) of those metrics such as sortie generation capacity, mission capability, or cost per flying hour.
3. The documenting, prioritizing, allocating, authorizing, synchronizing, scheduling, and monitoring resolution of operational requirements.
4. The institution of lower-level IMBs to control information within doctrinal categories.

Summary

Information architecture is more than just data in a database. It is a hierarchy of products and services organized to assure the timely and accurate delivery and storage of data and information across the enterprise. This article has described the components of an information architecture strategy using examples from the GCSS-AF Enterprise Architecture, which is currently under development as part of the GCSS-AF program.◆

References

1. Web Services Interoperability Organization <www.ws-i.org>.
2. Weibel, S., J. Kunze, C. Lagoze, and M. Wolf. "RFC 2413 Dublin Core Metadata for Resource Discovery." Sept. 1998 <www.faqs.org/rfcs/rfc2413.html>.

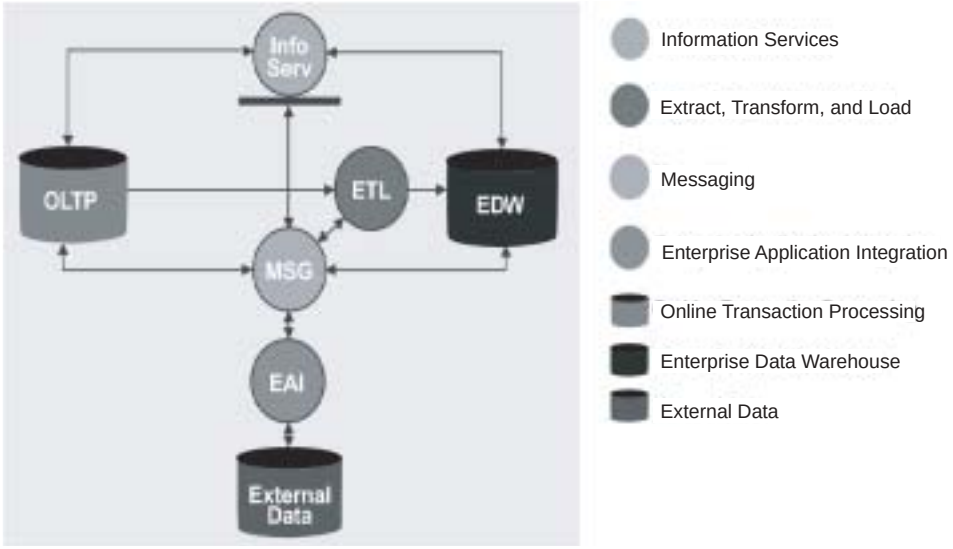


Figure 4: *An Information Architecture Data Flow*

Notes

1. This article is based on the "GCSS-AF Information Architecture Strategy" White Paper, document number PROJ-2003-GCSSAF-0449, and is the result of the efforts of several individuals on the GCSS-AF project.
2. A Web service is defined as a service that is available on the World Wide Web, is accessible via Web protocols, is self-defining for humans and machines, and is registered in a Web repository for easy lookup and activation.

About the Authors

John Wunder is a certified Lockheed Martin architect and has been the lead system architect/composer on the Global Combat Support System-Air Force since 1999. Prior to this, he was lead architect for Dow Chemical Process Control, and software architect for the U.S. Army battlefield digitization project. Wunder has been involved in information technology for more than 20 years.

Lockheed Martin Mission Systems
1801 State Route 17C, MD 0605
Owego, NY 13827
Phone: (607) 751-6096
Fax: (607) 751-2538
E-mail: john.wunder@lmco.com

Will Tracz, Ph.D., is a principal research scientist in the Global Combat Support System-Air Force (GCSS-AF) Enterprise Engineering department at Lockheed Martin Mission Systems. He is responsible for investigating innovative applications and extensions of the GCSS-AF Architecture Integration Framework. Tracz is a member of Lockheed Martin's Corporate Advanced Software Technology Focus Group as well as an ad hoc member of the 1999 Air Force Scientific Advisory Board Commercial Off-the-Shelf study. He is a member of the Rochester Institute of Technology Software Engineering Advisory Board, editor of the Association for Computing Machinery SIGSOFT Software Engineering Notes, and the author of more than 100 technical reports and books.

Lockheed Martin Mission Systems
1801 State Route 17C, MD 0605
Owego, NY 13827
Phone: (607) 751-2169
Fax: (607) 751-2538
E-mail: will.tracz@lmco.com