SQUARED Methods

Enterprise Information Architecture Vision

White Paper Series

Table of Contents

1

Introduction

Architectural decisions identify the system's key structural elements—data, Metadata, Information Flow, flexibility, reliability, security, identity management, and industry standards and Best Practices, the externally visible properties of these elements, and their relationships. The goals of the decisions are to maintain system integrity across the enterprise through a unified overall design, form or structure. Without a holistic approach to designing and building Information Systems, we will continue to make decisions optimized for local solutions, in many cases pushing enterprise-wide needs to the back seat.

What's important to recognize is that Information architectural decisions apply to structural units—applications, services, systems and data—that are to be used more broadly within the enterprise, and therefore, not to designs that effect only local solutions.

A well designed and implemented Information Architecture not only helps typical unit level project implementations, but it also keeps all the initiatives in the context of an Enterprise, a key concept in building an Enterprise Information Architecture. Components of an Information Architecture leverage existing applications, systems, interfaces, databases and processes to establish a set of practices for capturing and maintaining informational artifacts in an Enterprise context.

In this document, we provide a view of where we believe Organization needs to be in terms of Information Architecture. Starting with Vision, we provide a set of driving factors and rationale

for why these aspirations are meaningful. This document also presents a conceptual view of Information Architecture for Organization to ensure that all the system development initiatives and existing data and application systems can be leveraged to make key strategic business decisions based on accurate, timely and reliable data even though it is scattered across the organization on diverse technology platforms.

2

Enterprise Information Architecture Vision & Goals

Vision Statement:

To unlock Enterprise Wide Information and align Data & Interface assets with Business Processes; enabling Organization to harness full power of timely, reliable and accurate information and gain complete visibility into enterprise information assets

Our goal with Enterprise Information Architecture is to provide Organization an insight into enterprise information that enables our business to remain agile and flexible. The vision for Information Architecture is to go beyond IT functions, and in fact, align Organization's IT assets to its business processes in a centralized & standardized manner. Enterprise Information Architecture aspires to provide our business the ability to maintain and gather business information that allows us to make better-informed decisions.

How we develop and implement Information Architecture is a large contributor towards realizing our vision. Developing and implementing Information Architecture & Practices is a large contributor to realizing the true power of Information we envision here at Organization.

At the highest level, there are six key goals that shape our desired state.

Agility

As a corporation, we need to be able to adapt to change quickly within a structured framework as newer systems are implemented and existing systems are updated. This is critical as changes introduce risk if the business is not aware of the low level system changes. The goal is to create an Information Architecture that ensures alignment of Organization's business processes with our Data, Applications and Interfaces in a way that allows us to understand and adapt to changes and its impact easily and in a standardized way

Interoperability

As a corporation, where possible, we need our business processes to align with common industry practices so that we can leverage standard ways to interact with applications and interfaces. We need to be able to leverage applications from different parts of the enterprise to fulfill specific business functions. Therefore, we want to be in a state where our applications and systems (databases, third-party software, Home grown applications, interfaces and unstructured data such as documents, emails etc) can easily exchange information and use that information for processing.

Trusted Point of Truth

At Organization, we need the capability to make informed business decisions, and correctly process business critical information based on trusted information. Therefore, we want to have reliable and trusted knowledge of our data sources so that we can have accurate and reliable information at our disposal, from a trustworthy source. As a large and diverse corporation, we have multiple representations and copies of similar data spread throughout the enterprise. Data sources may have been created to serve the needs of local business efforts, making it expensive, difficult, and time consuming to form a common view of the information embodied within the data. We also need to have the capability to understand how data has evolved over time and be able to trace data back to its origin.

Reliability

To make timely, accurate business decisions, we need our information systems to present reliable information to the decision makers. Reliability of information is a major step that can be achieved by ensuring we

understand the data, the way applications and interfaces use this data and the way our Business views and creates this data. This knowledge provides a critical link between all components in multiple business units and presents us with a capability to leverage data in a reliable way.

Security

Unauthorized modifications to data or systems can result in business decisions that do harm to the enterprise. Unauthorized use or disclosure of sensitive data can result in loss of proprietary information and reputation for the enterprise. Security is not an afterthought or add-on when building IT systems. Rather, security considerations must be an integral part of the overall information architecture design. We want to be in a state where we have applied consistent security practices to the data, applications and system that comprise our enterprise, and that we have applied those security practices as we design them. We also need to ensure that there is complete traceability of information to its source and a full auditing capability is available.

Suitability

As a global enterprise, Organization has a huge amount of data and information across a multitude of systems that needs to be incorporated into an overarching information vision. It is vital that we deliver the right information to the right people in the right way using the right tools. We need our information systems to present relevant data and information to the consumers in a way that engages rather than overwhelms them.

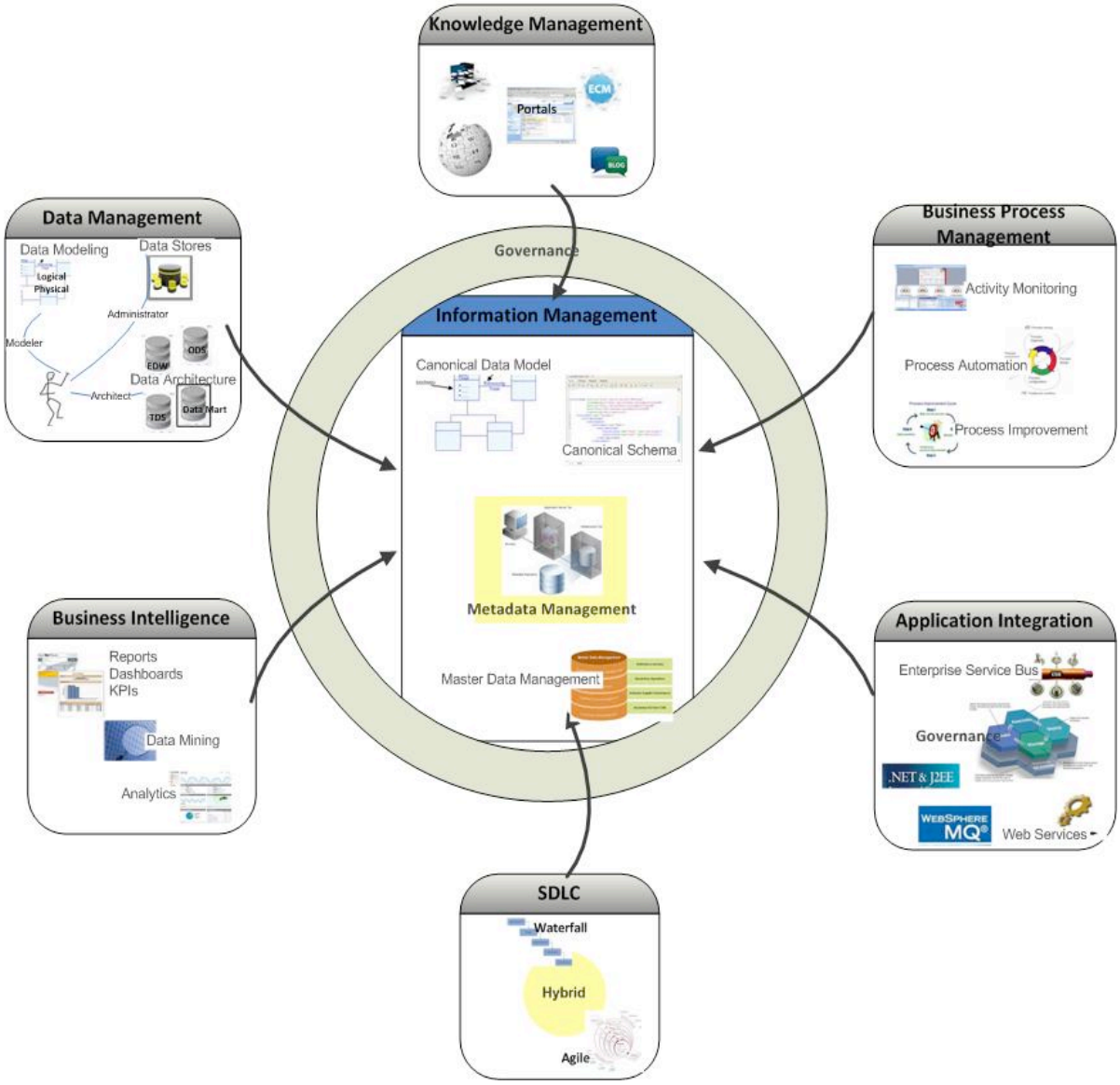
3

Enterprise Information Architecture – Conceptual View

The function of Information Architecture is to create a trusted alignment between our Business processes, Data and Applications & interfaces in a standardized way that can be easily shared and therefore leveraged across our enterprise. In order to accomplish this function, certain Software components need to be in place at Organization. This section presents a conceptual view of Information Architecture and desired Software components to support our vision for Information Architecture at Organization.

Conceptual View - Information Architecture

Shown below is a conceptual depiction of the desired Target State Information Architecture.



As shown in the above depiction, the Information Architecture addresses alignment between the following Key domains:

1. Data Management: This domain deals with the Data Modeling, Data Architectures, related standards and Best Practices, and corresponding Templates. The domain also addresses a central place to store Data Models to ensure accessibility of information. Typical systems in this domain are Transactional, Operational and Analytical (Data Warehouse & Data Marts) systems

2. Application Integration: This domain consists of Application Integration Architectures, including Service Oriented Architectures, Application Interfaces, J2EE & .NET components. Also available are Interface Governance practices and related standards and templates.

3. Information Management: This domain primarily deals with capture, maintenance and sharing of Information and related artifacts across Organization. This includes a Subject Area Model, Canonical Logical Data Model a Metadata Repository that acts as a central authoritative source for maintaining metadata and Master Data Management capabilities. A comprehensive Governance framework will be established to ensure consistency of information across Organization and to enforce the Standards & Best Practices throughout the enterprise. Information Management also includes using standardized schemas (such as Canonical XML schema) to capture the information flow between applications and systems.

4. Business Process Management: This domain includes a business process view of Organization's activities along with related standards (such as BPEL) that allow interoperability between processes. Also in this domain are Business Vocabulary and Business Term definitions that form a critical reference point for Information Management. These Business definitions and terminologies present a consistent view of Organization across enterprise systems.

5. Business Intelligence: This domain is related to the Reporting and other forms of Decision Support capabilities such as Data Mining, Analytics, Dashboards and applications that present the enterprise wide data in the form of KPI's to the management to support business decisions.

6. Knowledge Management: This domain addresses the knowledge life cycle and communication and information sharing strategy to deliver structured and unstructured content using an enterprise wide platform (such as Universal Content Management (UCM) or SharePoint)

At a high level, this Future State Information Architecture presents some key benefits:

I. Business centric view of Organization, which becomes a critical reference point for enterprise wide integration and BI efforts

II. Central and trusted version of truth to ensure consistency of information

III. Complete life cycle view of information assets as data transitions from its System of Record through other systems and applications

IV. Insight into the data lineage and impact of data related changes on enterprise systems and applications

V. Shared information across Organization that can be leveraged across SDLCs to ensure consistency and accuracy of information.

The Metadata Repository (MDR) forms the key central repository which stores metadata for all the other contributing domains. All the Business Intelligence & Reporting initiatives will then leverage the business oriented view of Data, Applications, Interfaces and Services that is available using a Metadata driven repository.

The Subject Area Model, Entity Information Model (EIM) and Canonical Logical Data Model (CLDM) form business oriented reference points for all Applications, Interfaces and Databases. The alignment of all these artifacts is a critical content of Metadata management tools that tie all Enterprise wide systems to these business models and provides visibility into our information assets.

MDR provides key capabilities including:

- Single & Trusted version of truth for metadata that captures an enterprise wide view of data & information assets
- Business Centric View of Organization using Business Dictionary
- Data Lineage & System of Record Identification
- Impact Analysis to plan changes to Business Processes
- Ability to share enterprise metadata using multiple access channels (portals, Reports, APIs etc.)
- Ability to store mappings of physical assets to information models across business units

4

Phases & Milestones

Realizing Organization's Enterprise Information Architecture vision is an enterprise wide effort, which is divided into four phases:

- Initiate
- Unlock
- Align
- Leverage

This section described the key milestones in each of these proposed phases:

Initiate

In Initiate phase, the goal is to establish the vision and get a buy-in from key business & IT stakeholders. This phase also sets a base foundation for the Information Governance and Best Practices to integrate SDLC with Information gathering initiatives Finally, Initiate phase sets the infrastructure required for Organization to create a long term sustainable Information Architecture.

The key deliverables for the Initiate phase are:

a) **Establish a Vision for Organization Information Architecture** & communicate

with IT executive management and Organization Themes management to get their buy in.

b) **Establish Organization's Information Governance Methodology** that will be used to capture, maintain and share Information across organization.

c) **Establish Best Practices and related Templates** to capture, maintain and publish Information related artifacts

d) **Establish base infrastructure needs** necessary to realize Organization's Information Architecture Vision (such as Metadata Repository, MDM platform)

Unlock

In Unlock phase, the goal is to capture information related artifacts and populate the key components of Information Architecture by importing models, metadata, relationships into a Central MDR in an Enterprise Scope.

The key deliverables for the Unlock phase are:

a) **Publish a Subject Area Model (SAM) and Entity Information Model (EIM) for Organization** to capture a Business Centric logical view of key subject areas within Organization and associated entities. These models can be leveraged for multiple efforts, such as Enterprise Data Warehouse (EDW), identifying Master Data (MDM) candidates, building an attribute level Canonical Logical Data Model, and mapping business requirements to enterprise perspective to gain visibility into the data. We anticipate that this model along with the Canonical Logical Data Model will be created in a phased manner aligned with ongoing initiatives.

b) **Populate Metadata Repository (MDR)** to import available data sources across Organization**.** The goal is to capture every available Data Source and its data model in the MDR. Metadata Repository provides a visibility into Enterprise wide data and promotes a common and shared understanding of information.

c) **Establish Publishing & Communication strategy for information** to share the data and information assets using collaboration platforms and to communicate the effective ways to leverage the data, information & metadata content for SDLC, BI and

Enterprise initiatives (Such as EDW)

Align

In Align phase, the goal is to align the Business Processes, Subject Area Model and the Canonical Logical Data Model (CLDM) with the enterprise data assets into a central repository such as MDR. Also part of this is alignment of SDLC with Information Architecture Best Practices to capture, manage and leverage Informational artifacts during implementation efforts. The goal is to link Business Processes with Service Interfaces, ETL processes, Workflows and Application Interfaces. At the end of this phase, Organization will have a solid centralized Repository that provides the key alignment between business and all Physical systems, data sources and interfaces.

The key deliverables for the Align phase are:

a) **Populate Central Repository (such as MDR)** to describe **ALL** data sources across Organization**.** The goal is to capture every available Data Source and its relationship with the Subject Area model & CLDM to link Physical data and information assets, Interfaces and Services to Business processes.

b) **Integrate Information Capture & Publishing Best Practices with SDLC** to ensure that the Information capture and publishing is enforced at appropriate phases during an SDLC implementation. The templates and guidelines will be leveraged from the Initiate phase.

Leverage

In Leverage phase, the goal is to fully utilize the information architecture, infrastructure and best practices captured in the central Repository to support enterprise wide initiatives such as Business Intelligence (BI), Decision Support Systems, Reporting needs and Application integration implementations.

The key deliverables for the Leverage phase are:

a) **Full BI capabilities** to support Data Lineage, Impact Analysis, System of Record Identification, Transformation needs for all Organization initiatives.

b) **Full integration of CLDM with SDLC** to Capture & maintain information in MDR current.

c) **Full utilization of information architecture, infrastructure & best practices for all Informational needs** to identify

System of Record, Business Terminology, Mapping of Data, Services and Processes to Business.

The following table captures the phases, some example deliverables and key milestones for the above phases:

Phase	Milestone	Milestone & Deliverables Description	Key Considerations
Initiate	Organization Information Architecture Vision	A complete description of Organization Information Architecture Vision and Goals along with target Implementation Phases & timeline	None
	Information Governance Methodology	IA Governance Methodology including: ▪ Standards & Best Practices with Templates ▪ Review & Approval policies ▪ SDLC Integration policies ▪ Communication & Sharing Plan ▪ Access Control policies	✓ Approved Organization Information Architecture Vision Required
	Base Information Infrastructure	Base Infrastructure to: ▪ Capture Metadata ▪ Create Subject Area Model & Entity Model ▪ Share information	✓ Approved Organization Information Architecture Vision Required
Unlock	Create Subject Area Model (SAM) Release	Complete Subject Area Model for Organization that captures enterprise view of	✓ Subject Area Model (SAM) aligned with

	1.0	critical areas	ongoing initiatives
	Create Entity Level Information Model (EIM) Release 1.0	Entity level model expanded from the Subject Area Model (SAM)	✓ Need a complete draft of the Subject Area Model (SAM) to build EIM
	Populate Metadata for Data Models	Import Logical/Physical Data Models into the MDR	✓ Base information Infrastructure required
	Capture Enterprise Resource Metadata	Populate Metadata for available BI resources, KPIs, System Interfaces and CLDM	✓ Base information Infrastructure required
Align	Mapping	Map imported data models, system interface metadata (such as WSDL), BI resources with Subject Area Model & EIM	✓ Aligned with ongoing implementation and integration initiatives
	Full SDLC Integration	Complete Integration with SDLC to capture, manage and leverage information related artifacts	✓ Approved Information Governance Methodology required
Leverage	BI Integration, Full SDLC Integration, Ongoing Metadata Capture	Leveraging full Informational capabilities in BI/Reporting initiatives	✓ For full leverage of capabilities, a fully released CLDM and fully populated MDR is required
	Attribute Level Canonical Logical Data Model (CLDM)	Ongoing effort to build and update the Canonical Logical Data Model and Canonical XML Schema for	

	and Canonical XML Schema	facilitating standards based integration across systems and applications	

5

Leveraging & Maintaining Information Architecture

Our goal is to make Information Architecture Organization's core strength. In order to realize that goal, our applications and systems need to better align with business changes. On the contrary, our applications must enable the business to such a degree that they provide information in a consistent and accurate manner to enable informed decision making for strategic long term competitive advantage.

Data Management

Data Architecture

Data Architecture practices influence creation of data systems for supporting transactional systems, Business Intelligence applications, Reporting solutions and analytical systems. Critical element of these systems is reliable data sources, correct understanding the data lineage to ensure right data is used to support applications and the right interpretations for data elements are used in order to make accurate business decisions. Well designed Information Architecture ensures that Data Architectures are based on accurate, reliable enterprise data.

Data Governance

Data governance is about ensuring high quality for enterprise data assets and establishing clear ownership and processes for managing these enterprise data

assets. Information Architecture is a critical component of Data Governance and ensures consistent data definitions are shared across the enterprise and all business functions understand the role of their data in an enterprise context. Effective data governance is critical to ensure alignment between business processes and IT systems and Information Architecture helps to accomplish that through use of enterprise repositories such as Metadata Repository, Master Data Management etc.

Master Data Management

Master Data Management is probably the single most important component of the Information Architecture. It is critical for our systems to have a clear authoritative source when it comes to critical enterprise data areas such as customers, products & parts, employees etc. Since the data for each of these subject areas is scattered across multiple systems, an MDM platform is a key technology component that enables high data quality by identifying the master data attributes and ties clear ownership and data sources to each of the subject areas.

Metadata Management

Metadata is another important component of the Information Architecture and ties all aspects of applications, databases, system of records together via custom metadata at multiple levels. Metadata Repository must be leveraged right from requirements analysis through final deployment to ensure consistency of enterprise data

System Integration Efforts

A well defined Information Architecture and Practice will support Interoperability, which is the ability of different systems to communicate with each other in a standardized way. In a large, heterogeneous systems environment like Organization, achieving this standardization is difficult. But with a concept of Canonical Logical Data Model and by creating a Canonical XML Schema, systems that need to exchange information and integrate with other systems can leverage these artifacts as logical reference points to define and derive data exchange needs.

Knowledge Management & Collaboration

Knowledge is built upon the foundation of data and information. In order to ensure knowledge is identified and stored effectively for enterprise distribution and application, data consistency is critical. Information Architecture provides a consistent

view of data and puts it in context of Organization's business processes to support effective knowledge management practices.

Success of Knowledge Management program depends on ability to effectively share, distribute and interpret information with right alignment with business processes. Information Architecture infrastructure must be leveraged by the collaboration platforms to ensure that knowledge is current, consistent and represents the current state of Organization's information systems and data assets.

Information Access Control & Security

As we design and build systems that automate our business processes, we are in many cases building systems that perform critical functions for our business. In such cases, we have invested significant effort to ensure that these systems produce the appropriate and expected results.

Audit tracking and managing access rights to systems are one aspect of security. But the key take-away point is that information security is about protecting our data and systems from unauthorized access—to create, read, delete, transform, etc. And the fact that security is important, means that it can't be an afterthought when we build IT systems. Rather, security must be part of the design from the beginning. Security requirements should be gathered with other functional and non-functional requirements during the requirements phase of the project. The facility to be able to audit and monitor security is vital to the success of this.

SDLC

SDLC phases are critical from Information Architecture standpoint. In order to ensure that the information architecture is current, and accurately represents the current state of data and information assets, all SDLC implementations must feed appropriate information to the Information Management program.

For example;

Business Requirements Analysis phase should leverage the information architecture to define appropriate data requirements, appropriate systems to integrate with and rules and transformations that must be performed on enterprise data to satisfy the business needs.

System Design phase must ensure that the conceptual, logical and physical data models are aligned with the business oriented Subject Area model and the Canonical Data Model; and interfaces are designed

using the Canonical Schemas and appropriate mappings are maintained to ensure consistency of information

System Development phase must update the Information Infrastructure (such as Metadata Repository) with the most current data models, business processes, interface mappings to keep it aligned with the enterprise context

During Review and Close Out phase, appropriate artifacts must be reviewed to ensure all information related to integration efforts are up to date in the Information Repositories, so they can be leveraged across Organization

Glossary of Terms

Term	Definition
API: Application Programming Interface	A set of rules and specifications to allow software to communicate and interact
BI: Business Intelligence	A set of technologies, processes and applications to help identify and analyse business data
BPEL: Business Process Execution Language	Standard executable language for specifying actions within business processes via web services
CLDM: Canonical Logical Data Model	Provides a detailed data-driven view of the enterprise
CWM: Common Warehouse Metamodel	Specification for modelling metatdata within a data warehousing environment
EDW: Enterprise Data Warehouse	Enterprise wide repository of data for the purposes of reporting and analysis using BI tools
EIM: Enterprise Information Management	A holistic approach to using information across the enterprise
ETL: Extract, Transform and Load	Process to extract data from source systems, clean and deliver to a target systems (often an EDW)
IA: Information Architecture	Structural design of shared information environments
KPI: Key Performance Indicator	A set of key metrics defined and used to measure business performance
MDR: Metadata Repository	Centralised storage repository of information about data; usage, relationship with other data, origin, format. Also known as a Data Dictionary
SAM: Subject Area Model	Definition of entities and relationships for a specific subject area
SDLC: Software Development Lifecycle	Structures for the development process of software
SOA: Service-Oriented Architecture	Provides functionality as a set of interoperable services to facilitate implementation within multiple systems
XMI: XML Metadata Interchange	Standard for exchanging metadata information via XML
XML: Extensible Markup Language	Set of rules and standards for encoding documents in machine-readable format
WSDL: Web Services Description Language	XML-based language that provides a model for describing web services

Index