Example Data Specifications & Information Requirements Framework

FRAMEWORK OVERVIEW

Alan D. Duncan

http://informationaction.blogspot.com

Tw: @Alan_D_Duncan

Information Strategy | Data Governance | Analytics | Better Business Outcomes

1 Purpose of the Framework

This example framework offers a set of outline principles, standards and guidelines to describe and clarify the semantic meaning of data terms in support of an Information Requirements Management process.

It provides template guidance to Information Management, Data Governance and Business Intelligence practitioners for such circumstances that need clear, unambiguous and reliable understanding of the context, semantic meaning and intended usages for data:

- Business usage of data in context;
- Evidence-based decision-making;
- Data quality requirements;
- Data sharing and publication expectations;
- Guidance from latest industry practices.

This will include data systems with both (near)real-time event-based processing, periodic batch processing requirements, and for scenarios such as:

- Enterprise-wide operational and transactional business systems;
- Locally operated faculty and school databases;
- Data warehouses and data marts;
- Peer-to-peer data integration;
- Enterprise Service Bus (ESB) and Service Oriented Architecture (SOA);
- Master Data Management (MDM);
- Extraction, Transformation and Loading (ETL/ELT);
- Data streaming;
- Business Intelligence and Analytics;
- Self-hosted, managed service and Cloud-based information repositories.

This Framework is prepared for use by both data-related projects and during operational service delivery:

- To ensure that due consideration is given to data semantic and quality meaning during planning, design and implementation stages;
- To help teams achieve improved consistency, integrity, repeatability and inter-operability of solutions with a strong data-related aspect;
- To ensure that delivered solutions meet minimum expectations in relation to the integration, storage, treatment and provenance of an organisation's data holdings;
- To be complementary to other delivery protocols (e.g PRINCE2, Agile methods & SDLC).

This document is best read in the context of broader Information Management & Data Governance practices, and with respect to the ISO8000 principles for Data Quality. (See Appendix A for further details).

DISCLAIMER: this generic framework is offered for guidance only. It is not intended to be comprehensive or to cover all possible options related to specific information requirements. The author accepts no responsibility for specific situations or circumstances where the framework is applied.

2 Framework Components

The framework comprises the following component templates:

Component	Purpose
Data Solution Outline & High Level Requirements Template (Project Mandate)	A template for capturing the overall high-level business requirements and expectations for business solutions with a significant impact on or requirement for data. (cf. the "Project Mandate" document in PRINCE2).
Business Data Requirements Template	A template for the clear and unambiguous definition of business data and information requirements. (cf. "Business Requirements Document", "Functional Specification" or similar from standard SDLC processes). As such, the contents will typically form the basis for population and publication of a business glossary of information terms.
Logical Data Definition Template	A template to define an outline structure for the clear and unambiguous definition of the discreet component data elements (atomic items of Entity/Attribute/Relationship/Rule) within the Logical layer of an Enterprise Information Model (a.k.a. Canonical Model).
Physical Data Specification Template	A template defining an outline structure for the clear and unambiguous definition of the discreet data elements (tables, columns, fields) within the physical data management layers of the required data solution.
Transformation Logic Template (Source to Target Mapping)	This document template defines an outline structure for the clear and unambiguous definition of transmission of data between one data storage location to another.
Analytics & Reporting Requirements Template	This document template defines an outline structure for the clear and unambiguous definition of analytics & reporting outputs (including standard reports, ad hoc queries, Business Intelligence, analytical models etc).

Please see each individual template for further details.

Appendix A: ISO8000 Principles

International Standard ISO8000 defines a broad set of principles and practices in relation to the semantic meaning of data and good management of data quality. It proposes that all data must demonstrate:

- Ability to freely resolve the metadata associated with data element values to a meaningful definition
- Ability to track the origin in time of the data element values
- Ability to track the organization that created or validated the data element values
- Ability to freely access the syntax
- Existence of a minimum specified set of data element values to satisfy user's requirement

Key to this is the explicit definition of semantic meaning for all data terms, together with the lineage and traceability of those definitions, as articulated by three core principles:

- **Syntax**: Each data set shall contain a reference to the syntax to which the data set complies. The reference shall be resolvable to the specification of the syntax through a mechanism that is publicly available, free of charge.
- **Semantic encoding**: Each data element value shall reference all concepts necessary to unambiguously define its meaning. Each reference shall be to a concept dictionary entry contained in a concept dictionary that supports an interface for free and anonymous resolution of a concept identifier.
- **Conformance to requirements**: Each data set shall contain a reference to the data requirements statement to which the data set complies. The reference shall be a globally unambiguous identifier that was used to encode the data set. The reference shall be resolvable to the data requirements statement through a mechanism that is publicly available, free of charge. The data requirements statement shall be publicly available.

About the author

Alan D. Duncan *is an evangelist for information and analytics as enablers of better business outcomes, and a member of the Advisory Board for QFire Software.*

An executive-level leader in the field of Information and Data Management Strategy, Governance and Business Analytics, he has over 20 years of international business experience, working with blue-chip companies in a range of industry sectors. Alan was named by Information-Management.com in their 2012 list of "Top 12 Data Governance gurus you should be following on Twitter".

Twitter: @Alan_D_Duncan

Blog: http://informationaction.blogspot.com.au

Intellectual curiosity

Skeptical scrutiny

Critical thinking


http://www.informationaction.blogspot.com.au/


@Alan_D_Duncan


http://www.linkedin.com/in/alandduncan