Title: Aligning OAIS with the Enterprise Architecture

José BARATEIRO[1,2], Gonçalo ANTUNES[1], José BORBINHA[1]

[1]INESC-ID: Information Systems Group, Lisbon, Portugal

[2]National Laboratory for Civil Engineering, Lisbon, Portugal

Abstract

The OAIS reference model is a well established "framework to understand significant relationships among the entities" involved in digital preservation, which aims to provide an abstract quality measure for the design of archival systems or repositories. However, it also details on structural and behavioural issues. In order to better decouple fundamental processes (required to be OAIS compliant) from optional processes, and avoid the risk of subjective interpretations on non-formal models, we propose to align OAIS with the Enterprise Architecture approach, providing a formalization of the OAIS "Functional Model" using the Business Process Modelling Notation. Finally, we claim that this approach is an effective advance to better assess OAIS and its relevance in the development of digital preservation solutions.

Introduction

The Reference Model for an Open Archival Information System (OAIS) provides a "framework for understanding significant relationships among the entities" involved in digital preservation. Moreover, as a reference model it intends to be based on a "small number of unifying concepts" [1, 5].

Actually, a framework can be described as "a set of assumptions, concepts, values, and practices that constitute a way of viewing the current environment" [2]. Reference frameworks can be used as basic conceptual structures to solve complex issues, providing a starting point to develop solutions concerning the targeted environment. Probably with the intention to support that, OAIS goes much further than providing just a high level reference model, detailing also on structural and behavioural issues.

This would not be forcibly negative in a reference model, but the problem with OAIS is that such details found in the standard are neither always properly aligned with the high-level concepts or its relevance is not evident. Probably it is, as a consequence of this, that only the high-level and more generic OAIS concepts are usually taken in consideration to deploy digital preservation environments, while the more detailed levels are not considered. In order to contribute to clarify this, we need to better decouple the generic from the specific concepts of OAIS. For example, the OAIS "Functional Model" describes processes that manipulate entities described in the "Information Model". Even though conformance to OAIS does not require adherence to the "Functional Model" [4], mandatory responsibilities are performed by "Functional Entities". We need to decouple fundamental from optional digital preservation processes described in the "Functional Model".

This motivates us to re-address the problem of what really should be part of a reference model for digital preservation systems, assess how actually OAIS really accomplishes that and, based on that, how could the model be improved.

The motivation for this work comes from the national funded project GRITO[1] and the European funded project SHAMAN[2], where the inaccuracies and incoherencies of the OAIS model have been identified and faced in the challenges posed by the multiple domains to be addressed (see Figure 1).

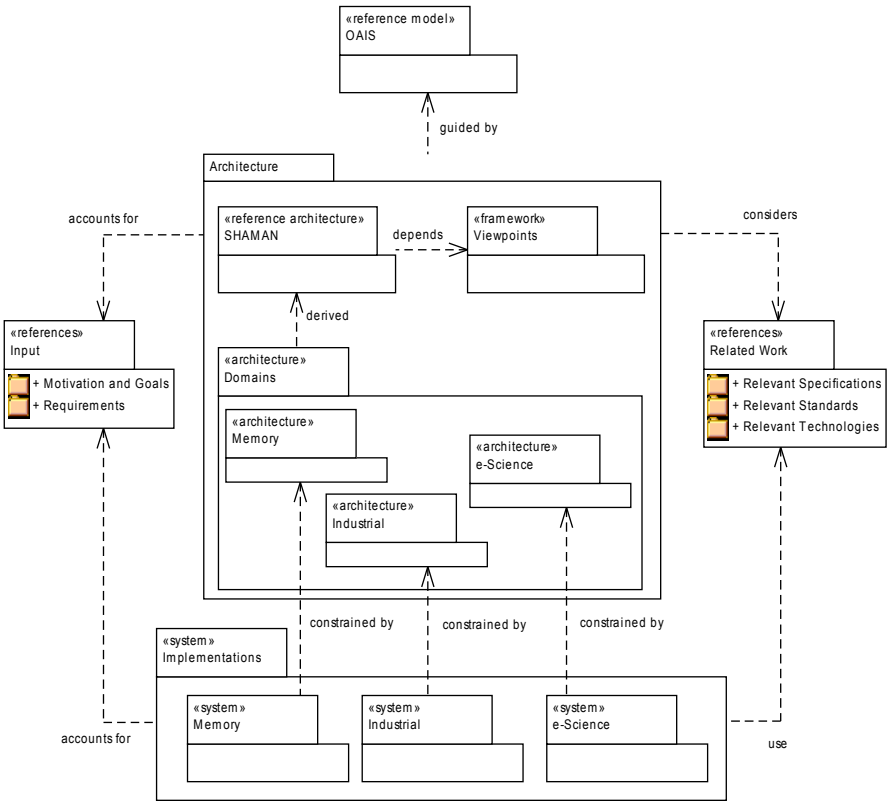

Figure 1 – The SHAMAN architectural model

OAIS and the Enterprise Architecture

In fact, we should recognize that, in the scope of digital preservation, it is crucial to better consolidate the perspective of the engineer (responsible for specific design and deployment of technological systems) to the perspective of the business architect (responsible by the business specifications, considering the related multiple systems, processes, and roles). Those concerns are addressed by the Enterprise Architecture [3] approach, which is defined as a coherent whole of principles, methods, and models that are used in the design and realization of an enterprise's organizational structure, business processes, information systems, and infrastructure [6]. An Enterprise Architecture framework is a communication tool to support the Enterprise Architecture process. It consists of a set of concepts that must be used as a guide during that process.

One of the first and most comprehensive Enterprise Architecture tools is the Zachman framework[3], defined as "…a formal, highly structured, way of defining an enterprise's systems architecture. (…) to give a holistic view of the enterprise which is being modelled".

Table 1 provides an overview of this framework, where each cell represents a view of the business, which can be defined by a set of models, services, principles, etc.

[1] http://grito.intraneia.com/ (FCT, GRID/GRI/ 81872/2006)
[2] http://shaman-ip.eu/ (European Commission, ICT-216736)
[3] http://www.zifa.com

The vertical dimension of the Zachman framework clearly separates the business (in this case, digital preservation) from architectural designs and implementations (as intended by OAIS). Using the Zachman framework as a reference, we can conclude that OAIS mainly stresses the Business Model level, willing to provide the unifying concepts and common understanding of the digital preservation business.

View	What (Data)	How (Function)	Where (Network)	Who (People)	When (Time)	Why (Motivation)
Scope	Things important to business	Processes the business performs	Locations the business operates in	Organizations important to the business	Events significant to the business	Business goals/ strategies
Business Model	e.g. Semantic model	e.g. Business process model	e.g. Business logistic system	e.g. Workflow model	e.g. Master schedule	e.g. Business plan
System Model	e.g. Logical data model	e.g. Application architecture	e.g. Distributed system architecture	e.g. Human interface architecture	e.g. Processing structure	e.g. Business rule model
Technology Model	e.g. Physical data model	e.g. System design	e.g. Technology architecture	e.g. Presentation architecture	e.g. Control structure	e.g. Rule design
Components	e.g. Data definition	e.g. Program	e.g. Network architecture	e.g. Security architecture	e.g. Timing definition	e.g. Rule specification
Instances	e.g. Data	e.g. Function	e.g. Network	e.g. Organization	e.g. Schedule	e.g. Strategy

Table 1 - The Zachman Framework

Another important reference in the Enterprise Architecture domain is The Open Group Architecture Framework (TOGAF), which consists of a ”detailed method and a set of supporting tools” [7]. It is divided in seven parts, though the most relevant are the Architecture Development Method (ADM), the Architecture Content Framework, and the Enterprise Continuum and Tools.

The ADM is defined as the core of TOGAF. It consists of a cyclical process divided in nine phases, which begins with the elaboration of the architecture principles and vision and goes through the elaboration of the concrete architectures and consequent implementation.

The Architecture Content Framework is the TOGAF alternative to the use of the Zachman framework or any other architecture framework. The Content framework divides the types of architecture products in deliverables, artefacts and building blocks. Deliverables represent the output of the projects and are contractually specified. Artefacts describe architecture from a specific viewpoint, an example being a diagram. Building blocks are reusable components of business, IT, or architectural capability which can be combined to deliver architectures and solutions. Deliverables are composed of artefacts which for its turn describe building blocks.

The Enterprise Continuum classifies the assets that may influence the development of concrete architectures. It contains two specializations, the Architecture Continuum and the Solutions Continuum. The Architecture Continuum classifies the architectures in Foundation Architectures, Common Systems Architectures, Industry

Architectures, and Organization-Specific Architectures. These can be used to guide and support the development of Solutions, which the Solution Continuum classifies as Foundation Solutions, Common Systems Solutions, Industry Solutions, and Organisation-Specific Solutions.

Formal Modelling of OAIS

Hence, in this paper we propose to revisit the OAIS "Functional Model", identifying the fundamental digital preservation processes, that is, the processes that are fundamental in digital preservation environments.

Furthermore, the OAIS "Functional Model" does not use any standard modeling notation, and ordinary language is used to describe the model, which makes it subject to subjective interpretation. For instance, a "Functional Entity" is a vague and subjective term, which is both used to refer structure and behavior.

In order to avoid the potential subjective interpretations of the "Functional Model", we present a set of BPMN – Business Processes Modeling Notation[4] diagrams, which make it possible to formally detail business processes by representing sets of activities and their contexts of execution, making it possible to describe behavior, abstracting from the systems' architecture. Making the connection with the Zachman framework, those diagrams fit the "How" cell, formally specifying the fundamental digital preservation business processes (required to be OAIS compliant).

Figure 2 presents a BPMN diagram[5] example of the OAIS core environment, detailing the relationships and information flows between the OAIS Functional Entities: Ingest, Preservation Planning, Data Management, Archival Storage, Administration and Access. Using the BPMN notation, these entities are Business Processes or simply processes. Thus, if one refers the Ingest process, it only deals with behaviour (not structure).

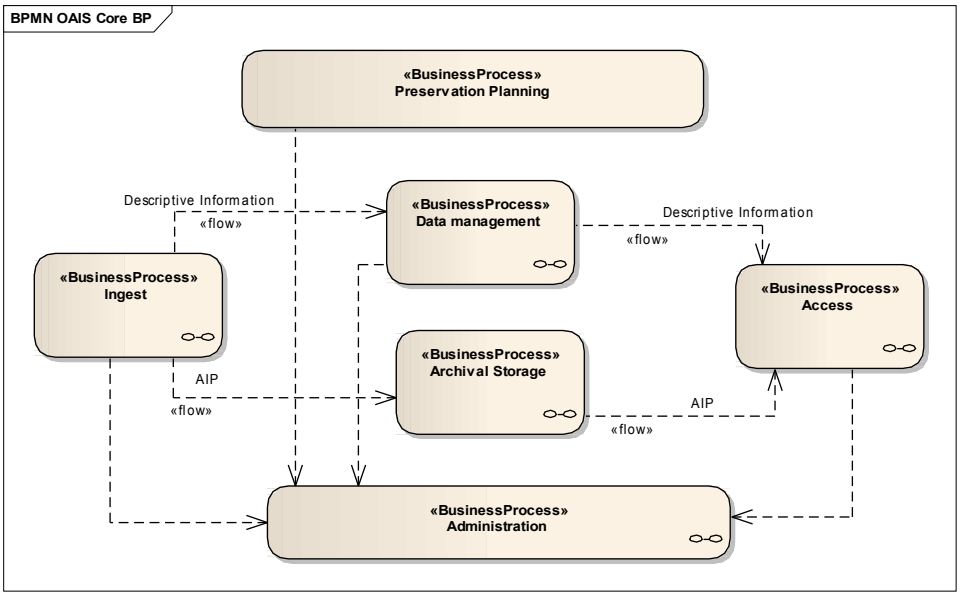

Figure 2 – BPMN diagram of the OAIS core environment

From this high-level business processes' diagram, it is relevant to detail the activities involved in each process. For instance, Figure 3 details the Ingest business process.

[4] http://www.bpmn.org/

[5] BPMN diagrams created from the OAIS diagrams and textual descriptions.

The use of the BPMN formalism avoids any kind of ambiguous interpretations. Moreover, it clearly separates all the structure elements (e.g., SIP) from behaviour elements (e.g., Generate SIP errors). It is also possible, among other things, to clearly know mandatory processes (or activities), mandatory flows, alternative flows, conditions, etc.

Another important "feature" of the BPMN formalism is possibility to generate BPEL – Business Processes Execution Language[6] files, representing the business specifications provided by the BPMN specifications.

As a consequence, if BPEL is used to "orchestrate" digital preservation services, it is possible to assure the alignment between services (implemented at the system level and supported by technology, manual processes, mechanical, etc.) and the digital preservation business.

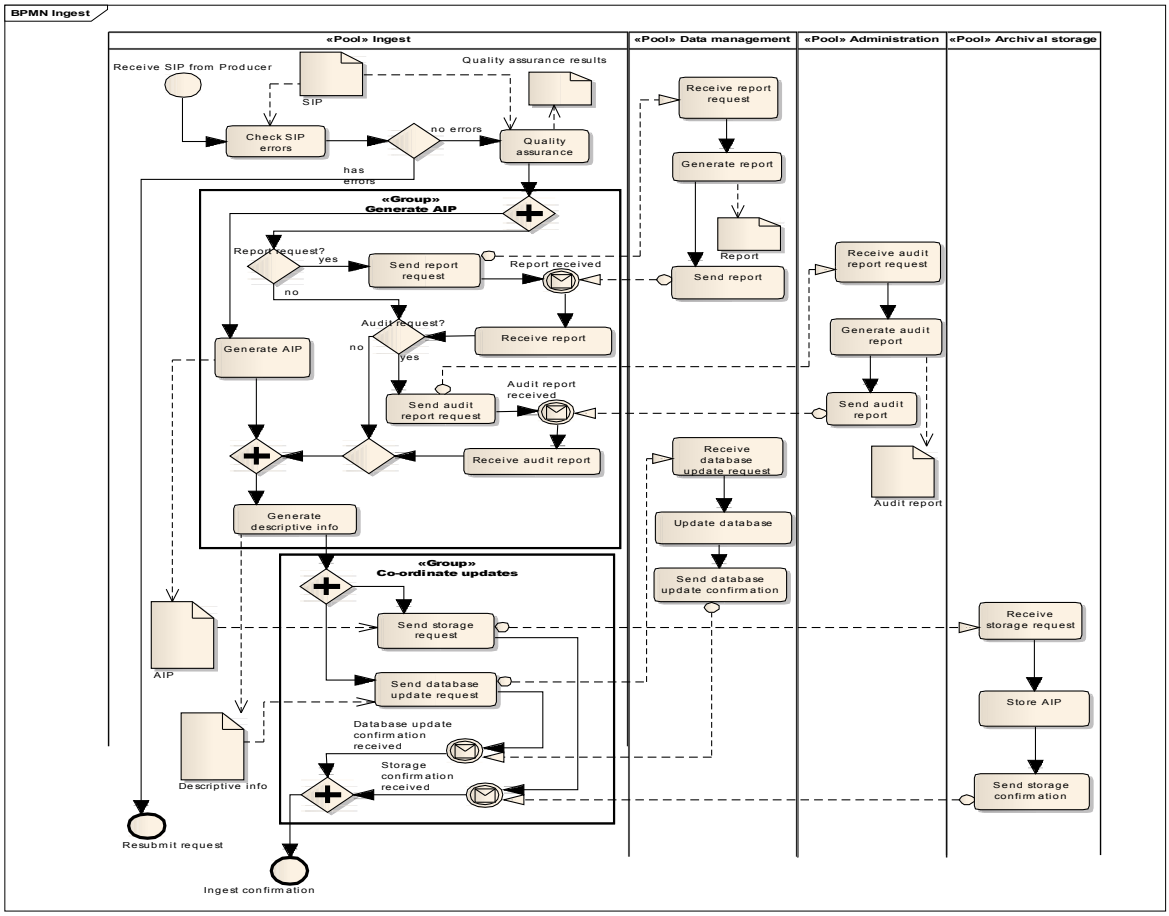

Figure 3 – BPMN diagram of the OAIS Ingest functional entity

Conclusion

This paper proposes to revisit the OAIS "Functional Model", identifying the fundamental digital preservation processes. In order to avoid the potential subjective interpretations of the "Functional Model", we present a set of BPMN – Business Processes Modelling Notation diagrams, which formally detail business processes (representing sets of activities and their contexts of execution), making it possible to describe behaviour, abstracting from the systems' architecture.

[6] http://www.oasis-open.org/committees/wsbpel/

The motivation for this work comes from the projects GRITO and SHAMAN, where OAIS is being faced to the challenges posed by multiple domains.

Concluding, this paper proposes two main contributions. First, to provide an exercise on the BPMN formalization of the OAIS “Functional Model”, focusing on the key requirements and related business processes, i.e., those process that deployed systems must adhere in order to be considered OAIS compliant. Second, we underline lessons from this exercise that will be relevant for a better assessment of OAIS and its future revision towards a more effective reference model for architects and engineers.

References

1. Consultative Committee on Space Data Systems, ISO 14721:2003 - *Reference Model for an Open Archival Information System (OAIS)*, 2003.
2. Organization for the Advancement of Structured Information Standards. Reference Model for Service Oriented Architecture. Committee Specification 1. August 2006.
3. V. Anaya and A. Ortiz. How enterprise architectures can support integration. In Proceedings of the First International Workshop on Interoperability of Heterogeneous Information Systems. Germany, 2005.
4. Julie Allinson. OAIS as a reference model for Repositories – an evaluation. 2006.
5. Lavoie, Brian. *The open archival information system reference model: introductory guide*. DPC Technology Watch Series Report 04-01, Digital Preservation Coalition and Online Computer Library Center, January 2004.
6. Lankhorst, M. Enterprise Architecture at Work: Modelling, Communication, and Analysis. Springer, Berlin/Heidelberg, 2005.
7. The Open Group. TOGAF Version 9. Van Haren Publishing, Zaltbommel, Netherlands, 2009.