A REFERENCE ARCHITECTURE FOR DIGITAL PRESERVATION

Gonçalo Antunes
INESC-ID
Rua Alves Redol 9, Apartado 13069, 1000-029 Lisboa, PORTUGAL

José Barateiro
LNEC
Av. Brasil 101, 1700-066 Lisboa, PORTUGAL

José Borbinha
INESC-ID
Rua Alves Redol 9, Apartado 13069, 1000-029 Lisboa, PORTUGAL

ABSTRACT

Apart from being a technological issue, digital preservation raises several organizational challenges. These challenges are starting to be addressed in the industrial design and e-Science domains, where emerging requirements cannot be addressed directly by OAIS. Thus, new approaches to design and assess digital preservation environments are required. We propose a Reference Architecture as a tool that can capture the essence of those emerging preservation environments and provide ways of developing and deploying preservation-enabled systems in organizations. This paper presents the main concepts from which a Reference Architecture for digital preservation can be built, along with an analysis of the environment surrounding a digital preservation system. We present a concrete Reference Architecture, consisting of a process to derive concrete digital preservation architectures, which is supported by an architecture framework for organizing architecture descriptions. In that way, organizations can be better prepared to cope with the present and future challenges of digital preservation.

1. INTRODUCTION

In order to achieve long-term digital preservation it is required to invest on a technical infrastructure for data storage, management, maintenance, etc. However, long-term digital preservation also raises several organizational challenges, since several business processes across the whole organization are affected by digital preservation.

Likewise, the complexity of long-term digital preservation increases with the fact that each type of business and specific organizations have their own particularities and special requirements, which makes the digital preservation business processes strongly dependent on their surrounding environment. For instance, the preservation policies depend on the type of data, its value for the organization, etc. As an example, the preservation of audio files requires recording information about compression and encoding/decoding which is not needed in the preservation of, for example, uncompressed XML files.

Concerning the organization type, memory institutions have several years of experience in dealing with the preservation of tangible objects. Additionally, the definition of preservation processes and policies concerning digital materials are common practices for these institutions. Usually, in the domain of memory institutions, technological solutions adopt the Reference Model for an Open Archival Information System (OAIS) [7], which provides a "framework for understanding significant relationships among the entities" involved in digital preservation. Actually, a framework can be described as "a set of assumptions, concepts, values, and practices that constitute a way of viewing the current environment" [12]. 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 behavioral issues.

Although the OAIS reference model has been widely adopted by memory institutions, it might not be suitable for scenarios with emergent digital preservation requirements, like industrial design. The OAIS reference model is definitely relevant for scenarios where the problem is to develop systems specifically for digital preservation, but it might not be appropriate for scenarios where the problem is to develop systems where digital preservation is a relevant property.

As a matter of fact, organizations with industrial design responsibilities produce a large amount of Computer-Aided Design (CAD) digital information within well-defined product lifecycles that cannot be aligned with the OAIS preservation processes and packages. Also, the collaborative environment of the scientific community, and associated services and infrastructures, usually known as e-Science (or enhanced Science) [11], involves digital preservation requirements. Actually, long-term digital preservation can be thought as a required property for future science and engineering, to assure that information that is understood today is transmitted to an unknown system in the future.

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 already addressed by the Enterprise Architecture [1].

According to [6], a Reference Architecture "captures the essence of existing architectures and the vision of future needs and evolution to provide guidance to assist in developing new system architectures". In that sense, we intend to demonstrate that a Reference Architecture should not be an artifact, but a process from which multiple architectural artifacts can result and be governed throughout their lifecycle. Based on that, we propose a Reference Architecture for digital preservation, capturing the essence of preservation architectures so that system architectures that are preservation-enabled can be developed and deployed in organizations.

The motivation for this work comes from the national funded project GRITO[1] and the European funded project SHAMAN[2], where requirements for digital preservation in e-Science and Industrial Design are being addressed.

This paper is organized as follows. First, Section 2 describes the concepts of architecture, reference architecture, stakeholder, view, viewpoint and enterprise architecture. Second, Section 3 describes the digital preservation environment where a preservation system inhabits. Next, Section 4 presents a framework to support the Reference Architecture. Section 5 presents the Reference Architecture which consists of a process for the development of concrete preservation-enabled architectures. Finally, Section 6 presents the main conclusions and future work.

2. MAIN CONCEPTS

This section describes the main concepts of concerning Reference Architectures. These concepts have been derived from international standards and related models of the area.

2.1. About Architecture

According to the IEEE Std. 1471-2000[3], architecture is "the fundamental organization of a system, embodied in its components, their relationships to each other and the environment, and the principles governing its design and evolution" [8].

The standard describes that a system (which has a mission) inhabits an environment which influences it. The system has an architecture which is described by an architecture description, providing a rationale for the architecture. The architecture description identifies the stakeholders of the system, which have concerns about the system. For its turn, an architecture description may be composed of several views (which might include several models of the architecture), which are according to the viewpoint of the stakeholder (which is used to cover the concerns of the stakeholder). The viewpoints might originate from a viewpoint library. The concepts of Stakeholder, Viewpoint, and View will be described in the following sub-sections.

2.2. About Reference Architecture

A reference architecture [8] is a way of documenting good architectural design practices to address a commonly occurring problem. It is way of recording a specific body of knowledge, with the purpose of making it available for further practical reuse.

A relevant source to better explain and understand these concepts is the work of the Service Oriented Architecture (SOA) Technical Group from the Organization for the Advancement of Structured Information Standards (OASIS). According to their SOA Reference Model [12], "Concrete architectures arise from a combination of reference architectures, architectural patterns and additional requirements, including those imposed by technology environments".

Reference architectures can be used to derive concrete architectures to a specific problem scenario, providing a basis from which those solutions can be derived. Architecture “must account for the goals, motivation, and requirements that define the actual problems being addressed” [12]. It is developed in an environment where some of the context is pre-defined (e.g., specific protocols, profiles, specifications, and standards).

In that sense, reference architectures can capture the essence of concrete architectures and relevant context and support the development of specific concrete architectures.

2.3. About Stakeholders

A successful architecture has to reflect the concerns and interests of the stakeholders. In [13], architecture is described as "a vehicle for communication and negotiation among stakeholders". Taking that into account, the architecture must also reflect the different viewpoints of all the interested parts, so that it can be communicated efficiently.

Also in [13], a stakeholder is defined as a viewer that perceives and conceives the universe, using his/her senses, in order to produce conceptions resulting from the interpretation of what is observed. A viewer can form a representation of the conceptions he/she makes using a determined language to express himself. When observing the universe, a viewer will be interested only in a specific subset of that universe, which is called a concern. The conceptualization of that subset of the universe is called a domain.

The process of abstracting a domain in a model is called modeling. In order to start a modeling process, a viewer must first construct a meta-model, comprising the

[1] http://grito.intraneia.com/ (FCT, GRID/GRI/ 81872/2006)

[2] http://shaman-ip.eu/ (European Commission, ICT-216736)

[3] IEEE Std. 1471-2000 consists in a standard for the architectural description and design of systems, recommended by the IEEE Computer Society. http://www.computer.org/standards

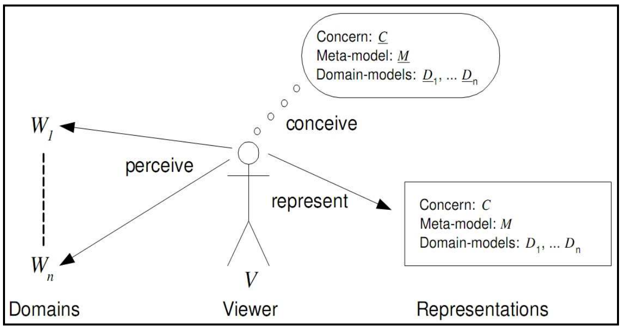

Figure 1. Viewing domains from a particular concern and meta-model [13].

meta-concepts and modeling approach, when modeling a domain. Figure 1 depicts a generic situation where a viewer with a determined concern and meta-model conceives and represents models for several domains.

Concluding, the concept of stakeholder has a crucial role in the development of an architecture since in order to be complete, an architecture should represent the different conceptions of the system through the use of models developed according to each of the relevant classes of stakeholders.

2.4. About Viewpoints and Views

Fundamental to the development of an architecture, and therefore to any reference architecture, are the concepts of "viewpoint" and of "view". The concepts are distinct and the need for this distinction is justified since a viewpoint is a "formalization of groupings of models" through a template or pattern for representing a set of concerns of a stakeholder [8]. A view is the concrete representation of a entire system from the perspective of a viewpoint, through a set of models. The viewpoint provides the categorization and the view provides the models according to the categorization.

In order to be complete, an architecture description must be composed of multiple views, addressing the concerns of multiple stakeholders. About the use of multiple views, the standard considers the following [8]: "The use of multiple views to describe an architecture is therefore a fundamental element of this recommended practice. However, while the use of multiple views is widespread, authors differ on what views are needed and on appropriate methods for expressing each view". Although the standard does not prescribe a set of views or modeling techniques for developing views, the field of Enterprise Architecture provides some examples of the views that should be considered in an architecture description.

2.5. About Enterprise Architecture

Enterprise Architecture 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 [10]. 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 Enterprise Architecture frameworks was the Zachman framework [15], defined as "...a formal, highly structured, way of defining an enterprise's systems architecture. (...) to give an holistic view of the enterprise which is being modeled."

The Zachman framework is summarized in simple terms in Table 1, where each cell on the table can be related to a set of models, principles, services, standards, etc., whatever is needed to register and communicate its purpose.

The columns of the Zachman framework express the

Perspective / Role	DATA What	FUNCTION How	NETWORK Where	PEOPLE Who	TIME When	MOTIVATION Why
Planner (Objective/Scope - Contextual)	Things important for the business	Business Processes	Business Locations	Important Organizations	Events	Business Goals and Strategies
Owner (Enterprise Model – Conceptual)	Conceptual Data / Object Model	Business Process Model	Business Logistics System	Workflow Model	Master Schedule	Business Plan
Designer (System Model –Logical)	Logical Data Model	System Architecture Model	Distributed Systems Architecture	Human Interface Architecture	Processing Structure	Business Rule Model
Builder (Technology Model – Physical)	Physical Data/Class Model	Technology Design Architecture	Technology Architecture	Presentation Architecture	Control Structure	Rule Design
Programmer (Detailed Representation – Out of Context)	Data Definition	Program	Network Architecture	Security Architecture	Timing Definition	Rule Speculation
User (Functioning Enterprise)	Usable Data	Working Definition	Usable Network	Functioning Organization	Implemented Schedule	Working Strategy

Table 1. The Zachman Framework

viewpoints relevant for this scope: the "What" refers to the system's content, or data; the "How" refers to the usage and functioning of the system, including processes and flows of control; the "Where" refers to the spatial elements and their relationships; the "Who" refers to the actors interacting with the system; the "When" represents the timing of the processes; and the "Why" represents the overall motivation, with the option to express rules for constraints where important for the final purpose.

The meaning of the rows are: "Scope" defines the business purpose and strategy; "Business Model" describes the organization, revealing which parts can be automated; "System Model" describes the outline of how the system will satisfy the organization's information needs, independently of any specific technology or production constraints; "Technology Model" tells how the system will be implemented, with the specific technology and ways to address production constraints; "Components" details each of the system elements that need clarification before production; and "Instances" give a view of the functioning system in its operational environment.

The Zachman framework influenced many other Enterprise Architecture frameworks [3].One of those frameworks is The Open Group Architecture Framework (TOGAF), which consists of a "detailed method and a set of supporting tools" [14]. It is divided in seven parts, the most relevant being 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 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, artifacts and building blocks. Deliverables represent the output of the projects and are contractually specified. Artifacts describe architecture from a specific

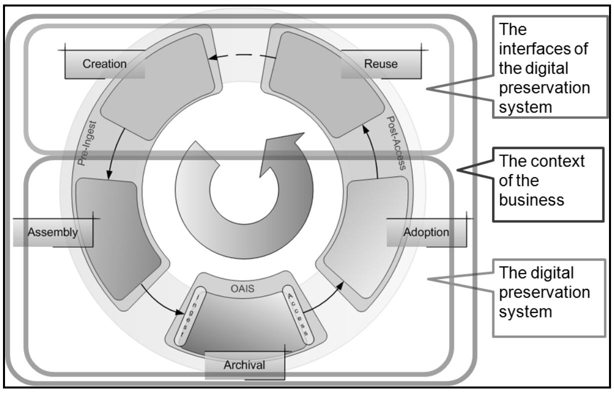

Figure 2. The Context of Digital Preservation in SHAMAN (adapted from [4]).

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 artifacts 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 Organization-Specific Solutions.

The Reference Architecture presented in this paper is largely inspired by TOGAF. It comprises an architectural framework and a process for the development of preservation architectures.

3. DIGITAL PRESERVATION ENVIRONMENT

As referred in Section 2.1, a "System inhabits an environment" which, for its turn, "influences the system".

Research undertaken in the SHAMAN project reached the conclusion that a bigger understanding of the environment where the preservation system operates is required [4]. A way of understanding the implications of the context of a digital object is through the analysis of its lifecycle. OAIS restricts itself to the "inner walls" of the archive, which may be insufficient in terms of the additional information required to preserve the object. A broader notion of the object lifecycle is needed, so that all the knowledge necessary to reuse the objects in the future is also preserved. The lifecycle of the digital object is represented in Figure 2.

The **Archival** phase spans the OAIS scope. **Creation** is the initial phase during which new information comes into existence. **Assembly** denotes appraisal of objects relevant for archival and all processing and enrichment for compiling the complete information set to be sent into the future, meeting the presumed needs of the designated community. It requires deep knowledge about the designated community in order to determine objects relevant for long-term preservation together with the information about the objects required for identification and their reuse some time later in the future. **Adoption** encompasses all processes by which information provided by the Archive is screened, examined, adapted, and integrated for Reuse. This phase might comprise transformations, aggregations, contextualization, and other processing required for repurposing of data. **Reuse** means the exploitation of information in the interests of the consumer and other processing required for repurposing of data.

Taking all this into account, in the perspective of the SHAMAN project, the digital preservation system

Vulnerabilities	**Process**	Software faults	T	.	.
		Software obsolescence	T	.	.
	Data	Media faults	T	.	.
		Media obsolescence	T	.	.
	Infrastructure	Hardware faults	T	.	.
		Hardware obsolescence	T	o	.
		Communication faults	T	.	c
		Network service failures	T	o	.
Threats	**Disasters**	Natural Disasters	T	.	C
		Human operational errors	t	O	.
	Attacks	External attacks	t	o	C
		Internal attacks	t	O	c
	Management	Organizational failures	.	O	.
		Economic failures	.	O	c
	Business Requirements	Legal requirements	.	.	C
		Stakeholders' requirements	.	o	C

Table 2. Taxonomy of Threats and Vulnerabilities to Digital Preservation

Stakeholders	Decision Making	Designated Community	Requirements and Conformance	Preservation Strategic Planning	**Viewpoints**
		Regulator			
		Auditor			
		Preservation Manager	Business Governance		
		Organization Manager			
		Technology Manager			
	System Building and Operation	System Designer	System Building and Support		
		Technology Provider			
		Technology Operator			
		Preservation Operator	Acting and Operation		
		Producer			
		Consumer			

Table 3. The Reference Architecture Framework

encompasses the phases comprised in the OAIS specification in addition to the Assembly and the Adoption of digital objects.

Considering the lifecycle of digital objects, the environment of the preservation system can be determined to be all that is outside and interfaces with the preservation system. In other words, the environment of the preservation system corresponds to the preservation "business" which the preservation system is supposed to support.

Taking into consideration this context of the preservation business and using Risk Management terminology [9], a taxonomy of threats and vulnerabilities of digital preservation, which takes technological, organizational, and contextual issues, can be devised [2].

Table 2 presents the taxonomy along with a classification of the threats and vulnerabilities according to the issues that may cause them (the capital characters represent bigger impact of a determined issue). The Reference Architecture for digital preservation draws from this analysis and is presented in the next sections.

4. REFERENCE ARCHITECTURE FRAMEWORK

An architecture description identifies the stakeholders of the system and is composed of several viewpoints that reflect the concerns of the stakeholders [8]. In this section, we present a framework for architecture descriptions to support the Reference Architecture.

Following the guidelines of the IEEE Std. 1471-2000, the stakeholder identification should take into account [8]: (*i*) the users of the system; (*ii*) those responsible for the acquisition and governance of the system; (*iii*) the developers and providers of the system's technology; and (*iv*) the maintainers of the system as a technical operational entity.

4.1. Stakeholders

The classes of stakeholders identified upon to this moment are: (i) Designated Community - As stated in OAIS, this is "an identified group of potential consumers who should be able to understand a particular set of information. The Designated Community may be composed multiple user communities". It may affect the design and development of the preservation system, since the system should satisfy their requirements; (ii) Preservation Manager - The person responsible for the definition and management of preservation policies (but that does not operate with the system, as that is the role of the Preservation Operator); (iii) Regulator - The person responsible for any external imposing rules concerning the preservation business, such as legislation, standards, etc. Those can apply to the organization, the technology, or the systems' usage; (iv) Auditor - The person responsible for the auditing and certification of the organization compliance with the established standards, rules and regulations; (v) Organization Manager - The top of the organizational structure with the main responsibility of defining the overall business objectives and strategy. It is typically a Chief Executive Officer, but it also might be a committee; (vi) Technology Manager - The person responsible for the definition of the overall technological strategy (software, hardware and infrastructure in general). It is typically called a Chief Information Officer, but it also might be a committee; (vii) Consumer - Represents the user accessing to the preserved objects, with a potential interest in its reuse; (viii) Producer - The person responsible for the ingestion of the objects to be preserved (the owner of the object, but it also can be any other entity entitled for that); (ix) Preservation Operator - The business worker responsible for the operation of the system. It may be aware of the details of the design and deployment of the system, but its main concern must be to assure the direct support to the business; (x) System Designer - The person responsible for the design and update of the architecture of the system, aligned with the business objectives; (xi) Technology Provider - The person responsible for the implementation and deployment of the architecture of the system or only its components; and (xii) Technology Operator - The person responsible for the regular operation and maintenance of the technological infrastructure (user accounts, replacement of damaged components, etc.).

4.2. Viewpoints

After the analysis of the stakeholders and their concerns, the viewpoints listed in Table 3 were derived. The main source used for that was the *Trustworthy Repositories Audit and Certification: Criteria and Checklist* (TRAC) [5], due to its wide scope view.

These viewpoints are: (*i*) **Preservation Strategic Planning** - Deals with the organization process of defining the digital preservation mission, vision and strategy in the context of the organization-wide mission, vision and strategy. It defines the direction of the organization concerning preservation. Although generally elaborated by the top-level management, it concerns all the stakeholders; (*ii*) **Requirements and Conformance** - Deals with the extra-organizational context that influences the adoption or operation of the system. It might be at the level of requirements of potential users or at the level of the legal framework that regulates preservation activities, also including the auditing of the system and involved processes; (*iii*) **Business Governance** - Deals with the high-level management of the preservation infrastructure, in terms of regulation, policies, best-practices, etc. It comprises three level: organizational, preservation and technological; (*iv*) **Acting and Operation** - Deals with the usage of the system and all the administrating and operational tasks related to preservation; and (*v*) **System Building and Support** - Deals with the technical analysis, design, implementation, and deliver of the system or of its components, including the related infrastructure.

The viewpoints can be further divided in sub-viewpoints which will correspond to models of the architecture. Each of these sub-viewpoints will correspond to a model which can be developed using the Unified Modeling Language (UML), or other formal or informal representation technique. For example, a sub-viewpoint of the Preservation Strategic Planning

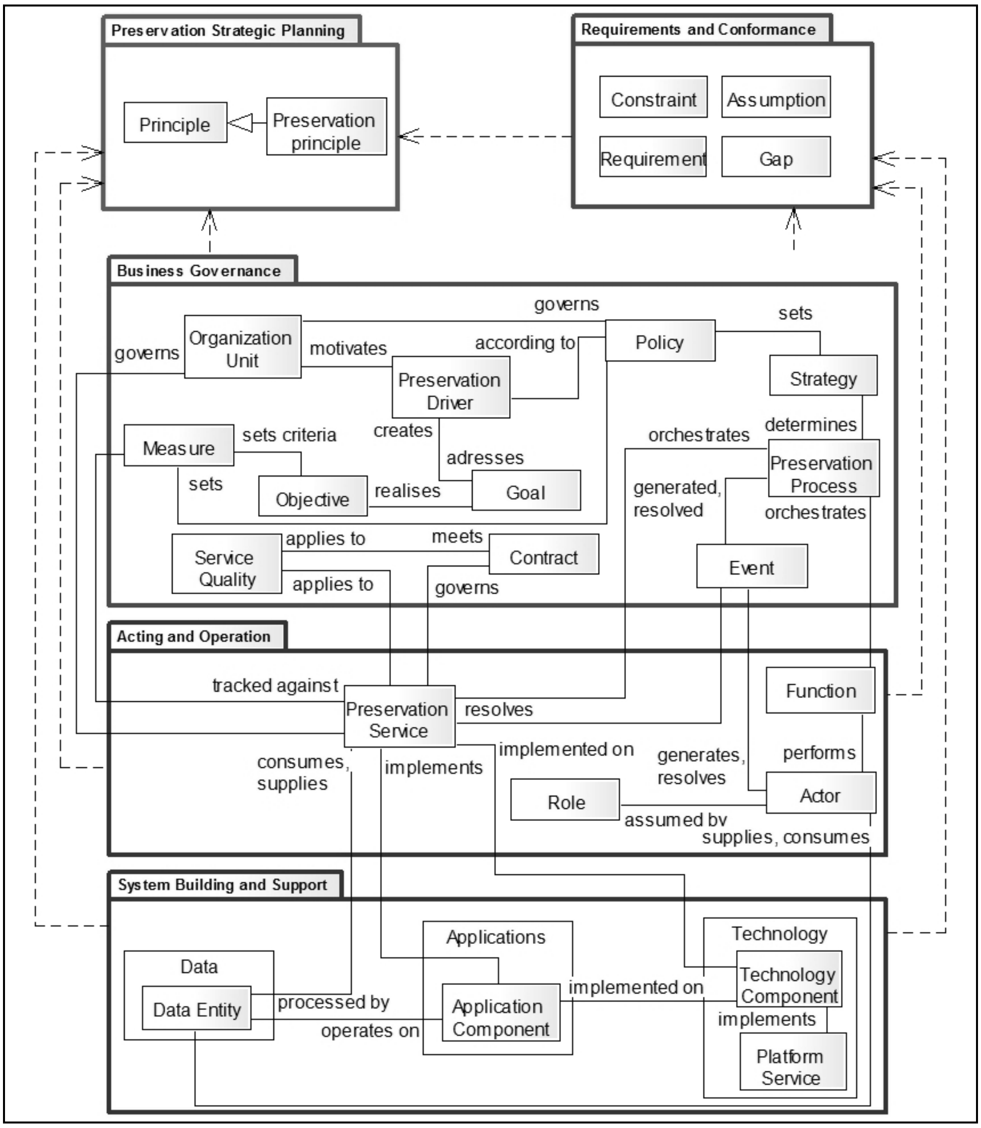

Figure 3. Reference Architecture Meta-model

viewpoint is the Preservation Principles Catalog, which contains a list of all the Preservation Principles that the architecture must comply with. The representation of this sub-viewpoint can be made through a table or a list.

4.3. Architecture Meta-model

The Architecture Meta-model provides a set of entities of the digital preservation domain, including the relationships between them. Those entities provide a common language for the domain which should be used on the development of the viewpoints of the architecture, when instantiating concrete architectures derived from the reference architecture. The meta-model enables the tracing between the different entities of the domain on the models of the architecture that result from the application of the Reference Architecture, enhancing the alignment between different viewpoints.

The meta-model is based in the TOGAF Content Meta-model of the Content Framework [14]. Figure 3 represents the entities of the digital preservation domain and relationships between the entities of the meta-model and also the relations between the viewpoints of the meta-model, using the Unified Modeling Language (UML).

5. REFERENCE ARCHITECTURE FOR DIGITAL PRESERVATION

A Reference Architecture "provides guidance to assist in developing new systems architecture"[6]. In that sense, should be a process which origins and governs the lifecycle of architecture artifacts, supported by a framework, which was presented in the previous section.

The IEEE Std. 1471-2000 does not provide or recommends a methodology for architecture development [8]. In other hand, the TOGAF specification [14], which is aligned with the IEEE Std. 1471-2000, provides a solid and detailed method for the development of architectures. Therefore, it was decided to base the SHAMAN architecture development process in the principles of the TOGAF Architecture Development Method (ADM). The result was the SHAMAN Architecture Development Method (SHAMAN-ADM).

The SHAMAN-ADM comprises six different phases (Figure 4), which are in line with the architecture viewpoints of the reference architecture framework presented in Section 4.

The Preservation Strategic Planning phase deals with the initiation of the architectural activities, comprising the definition of the enterprise scope of the architecture, the existing organizational context, (preservation) business requirements, the architecture principles, the identification of the relationships between the architectural framework and other governance frameworks, evaluating the maturity of the architecture, and developing an Architecture Vision that provides guidance throughout the development of the architecture.

The Business Governance phase is concerned with the development of a business governance architecture for digital preservation that supports the Architecture Vision. The Acting and Operation phase determines the requirements and functions required by the actors of the system, supporting the Architecture Vision.

The System Building and Support is divided in three sub-phases. The Data Architecture phase determines the data needed to support the effective preservation of digital objects. Also, data migration requirements should be supported by the data architecture resulting from this phase. The Applications Architecture phase defines the applications needed to support the data and business of digital preservation. The Technology Architecture determines the technology components needed to support the application components defined in the previous phase. Finally, the Architecture Realization phase is concerned with the architecture implementation process.

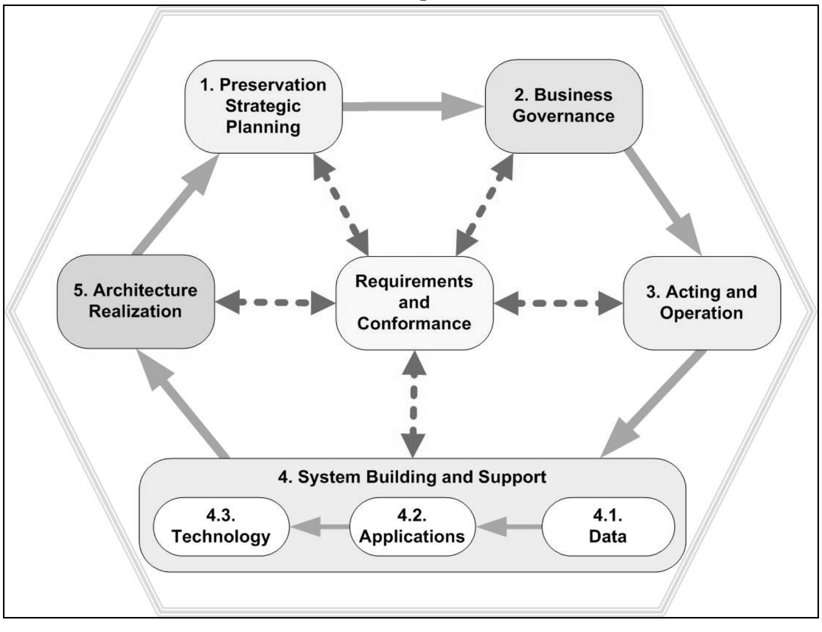

Figure 4. The Reference Architecture Development Method (SHAMAN-ADM).

The Requirements and Conformance should be a continuous practice throughout the application of the ADM. The management of requirements should be dynamic and preservation requirements at all levels shall be identified and stored, fed into and out of all the phases of the development cycle.

The application of this process in conjunction with the Reference Architecture framework should result in a architecture with preservation properties and in conformance with the requirements of the preservation stakeholders.

6. CONCLUSIONS AND FUTURE WORK

This paper presented a Reference Architecture for Digital Preservation. This work demonstrates a framework and a process from which concrete systems architectures with preservation properties can be derived, addressing particularly two digital preservation domains which introduced new and emergent requirements that cannot be addressed directly by OAIS: the Industrial Design and the e-Science domains.

We also presented the main concepts which form the background to the Reference Architecture, namely the concepts of Architecture, Reference Architecture, Stakeholder, Viewpoint and View, and Enterprise Architecture. Additionally, we motivated our approach through a general analysis of the digital preservation environment.

Future work will now focus on the application of the Reference Architecture to concrete cases to be explored on the scope of the SHAMAN project, which will result in the production of preservation-enabled architecture for specific cases. Another possible result may be a specialization of the reference architecture into the three domains of focus explored by the project, if irreconcilable differences are found between the domains.

7. ACKNOWLEDGEMENTS

The research reported was mainly supported by the project GRITO (a Grid for Digital Preservation), funded by the FCT (Portuguese Foundation for Science and Technology) under the contract GRID/GRI/81872/ 2006, and by the project SHAMAN (Sustaining Heritage Access through Multivalent Archiving), funded under 7th Framework Programme of the EU under the contract 216736.

8. REFERENCES

[1] Anaya, V., Ortiz, A. "How enterprise architectures can support integration", *In Proceedings of the First International Workshop on Interoperability of Heterogeneous Information Systems*, Germany, 2005.

[2] Barateiro, J., Antunes, G., Freitas, F., Borbinha, J. "Designing digital preservation solutions: a Risk Management based approach", *In the 5th International Digital Curation Conference - "Moving to Multi-Scale Science: Managing Complexity and Diversity"*, London, UK, 2009.

[3] Borbinha, J. "It is Time for the Digital Library to Meet the Enterprise Architecture", *In Proceedings from the 10th International Conference on Asian Digital Libraries*, Hanoi, Vietnam, 2007.

[4] Brocks, H., Kranstedt A., Jäschke, G., Hemmje, M. "Modeling Context for Digital Preservation", *Smart Information and Knowledge Management: Advances, Challenges, and Critical Issues (Springer Studies in Computational Intelligence)*, 2010.

[5] Center for Research Libraries. *Trustworthy Repositories Audit and Certification (TRAC) Criteria and Checklist.* Chicago, 2008.

[6] Cloutier, R., Muller, G., Verma, D., Nilchiani, R., Hole, E., Bone, M. "The Concept of Reference Architectures", Systems Engineering, 13 (1), pp. 14-27, Wiley Periodicals, 2010.

[7] Consultative Committee on Space Data Systems. *Reference model for an open archival information system.* ISO 14721:2003, 2003.

[8] IEEE Computer Society. *IEEE Std. 1471-2000: IEEE Recommended Practice for Architecture Description of Software-Intensive Systems*. IEEE, New York, 2000.

[9] International Organization for Standardization. *ISO 31000: Risk Management Principals and Guidelines*. ISO, Geneva, Switzerland, 2009.

[10] Lankhorst, M. *Enterprise Architecture at Work: Modelling, Communication, and Analysis*. Springer, Berlin/Heidelberg, 2005.

[11] Miles, S., Wong, S. C., Fang, W., Groth, P., Zauner, K.-P., Moreau, L. "Provenance based validation of e-science experiments", *Web Semant*, 5(1):28-38, 2007.

[12] Organization for the Advancement of Structured Information Standards. *Reference Model for Service Oriented Architecture version 1.0*. OASIS Standard, 2006.

[13] Proper, H. A., Verrijn-Stuart, A. A., Hoppenbrouwers, S. J. B. A. "On Utility-based Selection of Architecture-Modelling Concepts", *In Proceedings of the 2nd Asia-Pacific conference on Conceptual Modelling*, Newcastle, NSW, Australia, 2005.

[14] The Open Group. *TOGAF Version 9*. Van Haren Publishing, Zaltbommel, Netherlands, 2009.

[15] Zachman, J. "A Framework for Information Systems Architecture", *IBM Systems Journal*, 12 (6), pp. 276-292, 1987.