Archiving Sensor Data

Applied to Dam Safety Information

José Barateiro[1,2], Gonçalo Antunes[1], Hugo Manguinhas[1], José Borbinha[1]
[1] INESC-ID, Information Systems Group, Lisbon, Portugal
[2] National Laboratory for Civil Engineering, Lisbon, Portugal
{jose.barateiro,goncalo.antunes,hugo.manguinhas,jlb}@ist.utl.pt

Abstract—**The consequences of structural failures in large civil engineering structures are potentially catastrophic, varying from high economic impacts to unrecoverable environmental damage or loss of life. To prevent that, these structures can be continuously monitored, therefore the management and preservation of the resulting data is crucial to support decisions concerning structural safety. However, preserving data also entails several risks and threats, comprising strong safety requirements. This paper analyzes the scenario of civil engineering safety, presenting the current systems used at the Portuguese National Laboratory for Civil Engineering to manage and preserve sensor data. The main risks that can impede the digital preservation of data are discussed and a solution is proposed where sensor data is objectively described and packaged in order to be reused in the future. This includes controlling the extraction of data from the operational systems, describing the representation of data through a Metadata Registry, and package the context information using a METS aggregator.**

Keywords— Sensor Data; Digital Preservation; Risk Management; Information Management; Workflow.

I. INTRODUCTION

The safety of large civil engineering structures like dams, bridges or nuclear facilities require a comprehensive set of efforts, which must consider the structural safety, the structural monitoring, the operational safety and maintenance, and the emergency planning [1]. The consequences of failure of one of these structures may be catastrophic in many areas, such as: loss of life (minimizing human casualties is the top priority of emergency planning), environmental damage, property damage (e.g., dam flood plain), damage of other infrastructures, energy power loss, socio-economic impact, etc.

The risks associated with these scenarios can be mitigated by a number of structural and non-structural preventive measures, essentially to try to detect in advance any signs of abnormal behavior, allowing the execution of corrective actions in time. The structural measures are mainly related to the physical safety of the structures, while the non-structural measures can comprise a broad set of concerns, such as operation guidelines, emergency action plans, alarm systems, insurance coverage, etc.

In order to improve the structural safety of large civil engineering structures, a substantial technical effort has been made to implement or improve automatic data acquisition systems able to perform real-time monitoring and trigger automatic alarms. This paradigm creates an imminent deluge of data captured by automatic monitoring systems (sensors), along with data generated by large mathematical simulations (theoretical models). Besides the fact that these monitoring systems can save lives and protect goods, they can also prevent costly repairs and help to save money in maintenance. In scenarios like this, it is crucial to provide solutions that support interoperability (i.e., the ability of two or more systems to exchange information and to use the information that has been exchanged [2]), including the concept of temporal interoperability (i.e., long-term preservation).

This paper focuses on the digital preservation dimension of interoperability, which aims to ensure that digital data remain authentic, accessible and understandable over a long period of time. As a first assumption, one can consider that the main reason to preserve data is to preserve its value, as an asset. Consequently, it does not make sense to preserve valueless data. However, to determine and assess the value of data is a difficult and error-prone task. On the other hand, it could be an error to consider that data that cannot be used today will have no value in the future. For instance, today's technology allows the simulation of mathematical models with a much higher resolution and volume of simulated data that was not possible a decade ago.

From this perspective, we assume that the preservation of data concerning the safety of large civil engineering structures is crucial, since: (*i*) observational data is unique and impossible to recreate, (*ii*) complies with legal requirements or contracts established with third-parties, (*iii*) allows the re-use of data for new research, and (*iv*) reduces costs (e.g., the retention of expensively generated data is cheaper to maintain than to re-generate) [3].

The work presented here was developed in the scope of the SHAMAN project[1], which has the aim of developing digital preservation techniques and tools. We analyze the scenario of monitoring dams to assure their structural safety. We show that the digital preservation of sensor data has to deal with the requirements of managing dynamic data, as sensors are continuously capturing data; and heterogeneous and potential large set of representation schemas. Finally, we present an approach, based on the Open Archival Information System (OAIS) Reference Model [4], and a working technical solution

---

[1] *http://shaman-ip.eu*

implemented specifically to address the challenges of dam sensor data.

The remainder of this paper is organized as follows: Section II describes the scenario of monitoring concrete dams in the scope of the Portuguese National Laboratory for Civil Engineering[2] (LNEC). In section III, the main risks that can hamper the preservation of this information are discussed. Section IV describes the proposed solution to digitally preserve dam sensor data. Finally, Section V resumes the main conclusions and future work.

II. DAM SAFETY

The interpretation of the correlation of several parameters measured, in different physical locations of a structure, can be used to validate the current state of that structure and predict its future behavior under specific and controlled conditions [1]. This is a key factor to detect potential anomalies and to be able to make decisions on time, reducing the risk of failures with catastrophic consequences. In the case of concrete dams, for example, their behavior is continuously monitored by instruments (e.g., plumb lines, piezometers) installed in strategic points of the dam [5] [6], which can typically range from hundreds to few thousands of instruments or sensors.

The related raw data, usually known as "readings", is collected manually by human operators or collected automatically by sensors. These readings are transformed, by specific algorithms, into engineering quantities (physical actions that can be used to assess the behavior of the structure as, for example, a tension or a relative displacement). Actually, the term "reading" does not clearly correspond to raw data, since a reading is already a transformation from the raw data. For instance, an electrical instrument like an extensometer might provide raw data as a voltage (mV), which is then converted by a reading instrument (or by the sensor) into a resistance and a resistance relation, which are finally converted into an extension (engineering quantity). This monitoring information includes, essentially, instrument properties, readings and engineering quantities.

The Portuguese regulations [7] state that the National Laboratory for Civil Engineering is responsible for keeping an electronic archive of data concerning the dam safety. Thus, the preservation of this data is a legal obligation. Moreover, that obligation defines the duties of the different parties involved in dam safety, namely the dam owners, the dam safety authority and the dam engineers and builders. As a consequence, several entities are compelled to share data, and thus must face interoperability and preservation issues when dealing with heterogeneous sources of information [8].

Currently, LNEC uses a modular information system (*GestBarragens*) that provides components to manage dam observations, visual inspections, physical models and mathematical models. It also supports the management of technical documents and provides a set of exploitation tools, in the form of tabular and chart reports, graphical visualization of geo-referenced information, among others. However, the *GestBarragens* system was not designed for preservation purposes. Indeed, it supports the operational procedures to manage information concerning the dam safety, but does not assure the preservation of this information. It is a web-based system developed on the top of the .NET framework, where the underlying data is stored and managed in an Oracle 10g database. It uses a SOAP interface to provide and expose exploitation services as well as multiple ingest services.

TABLE I. summarizes an example of the data concerning the dam safety of a concrete dam. Currently, LNEC supports 32 different types of instruments with manual data acquisition and 25 different types of automatic monitoring instruments (implemented with sensors). Both the number and type of instruments installed in a specific structure depend on the stage of the structure's life and on a few hundred to thousand specific parameters that affect its behavior. Currently LNEC monitors about 80 concrete dams, generating an average of 264,000 records per day that have to be processed and preserved.

III. DIGITAL PRESERVATION RISKS

Although it is impossible to define all the requirements applicable for all digital preservation needs, a survey was made following a set of requirements based on the scenario presented in Section II.

First of all, digital preservation requires that a copy (or representation) of any preserved digital data survives over the actual system's lifetime, which is usually unknown, but may be as long as decades or even centuries (LNEC monitors concrete dams of more than 80 years old). This can be defined as a **reliability** requirement. Therefore, a digital preservation system must be designed to preserve data for an indefinite period of time without suffering any data losses.

Also, a future consumer should be able to decide if the accessed information is sufficiently trustworthy. Usually, this requires the assurance of the **authenticity** of digital data (which is already a common requirement for tangible objects), along with an accurate identification of their **provenance** (typically information about its creation, responsible entity, lineage, etc.). Moreover, it is crucial to assure the **integrity** of digital data, guaranteeing that their information content was not modified. Authenticity, provenance and integrity are thus crucial requirements for qualified specialists to trust and correctly approximate and estimate the behavior of large civil engineering structures.

The provenance requirement is fundamental. Furthermore, the complex scientific computations that occur in the production workflow (e.g. calculation of engineering quantities from raw data, outlier's detection) also make it complex to manage. The production workflow can be seen as an example of a scientific workflow [9] where data transformations and analysis steps, as well as the mechanisms to carry them out, are captured and represented as a workflow [10]. An access and re-use scenario, which is common in the simulation of mathematical models should also be considered. In fact, a mathematical model consumes observational data (preserved in the archive) and produces new digital data that should be preserved.

Third, digital preservation requires that future consumers are able to obtain the preserved information as its creators

---

[2] *http://www.lnec.pt*

TABLE I. TYPICAL DATA REGISTERED FOR A REPRESENTATIVE CONCRETE DAM

Data Stage	*Description*	*# per day*	*Format*	*Notes*
Raw	Depend on the instrument type (e.g. voltage)	Currently discarded	Proprietary to the sensor	This information is currently discarded by sensors and not registered during manual acquisition
Processed readings	Transformed from raw data	Aprox. 3300 rows	.xls, .mdb, PDT, ascii	Sensors register data in .xls or .mdb and access a web service to send this information to LNEC. Manual acquisition can be registered into a PDT and automatically sent to LNEC or inserted via web interface or text file
Calculated engineering quantities	Calculated from readings	Aprox. 3250 rows	Oracle database	Algorithms to filter, clean and calculate engineering quantities are implemented as Oracle stored procedures (PL/SQL)
Analyzed	Tables, graphs, gis, mathematical simulations	Varies	.html, .xls, .pdf, .dxf (CAD), .xml	Uses several tools, including reporting tools and a geographic information system

intended, thus it must **deal with obsolescence** threats [11]. This requirement encloses several challenges, since digital data to be explored, require a technological context defined by specific software and, in some cases, even by specific hardware [12]. Moreover, in this special scenario, it is also crucial to preserve the processes involved in the creation of the preserved data. For instance, the scientific workflow for data acquisition must be preserved and linked with the generated data.

Finally, dynamic collections and environments for digital preservation require technical **scalability** to face technology evolution allowing, for instance, the addition of new components through incremental updates [13]. This also implies a requirement for supporting **heterogeneity** (which is reinforced by the requirements for scalability).

In previous work, a taxonomy for digital preservation risks (see TABLE II.) was proposed, which considers that a risk is the impact that occurs when an event (threat) is able to exploit a system vulnerability, affecting the achievement of the digital preservation requirements described above.

TABLE II. THREATS TO DIGITAL PRESERVATION [14]

Vulnerabilities	Data	Media faults Media obsolescence
	Process	Software faults Software obsolescence
	Infrastructure	Hardware faults Hardware obsolescence Communication faults Network service failures
Threats	Disasters	Natural disasters Human operational errors
	Attacks	Internal attack External attacks
	Management	Economic failures Organization failures
	Business	Legal requirements Stakeholders' requirements

Like common information system's architectures, this paper considers a preservation environment as the aggregation of different components, namely: (*i*) the information entities, including preserved digital data and metadata, (*ii*) processes controlling the information entities (can be supported by computational services), and (*iii*) the technological infrastructure that supports the preservation environment.

Each of these components may present several vulnerabilities, which we classify as: (*i*) data vulnerabilities, affecting the information entities, (*ii*) process vulnerabilities, affecting the execution of processes (manual or supported by computational services) that control information entities, and (*iii*) infrastructure vulnerabilities, enclosing the technical problems in the infrastructure's components.

A classification of threats to digital preservation is also proposed which distinguishes threats into four categories: disasters, attacks, management and business. Disasters and attacks correspond, respectively, to non-deliberate and deliberate actions affecting the system or its components. Management failures are the consequences of wrong decisions that produce threats to the preservation environment. Finally, business threats depend on a specific business context and occur when new or updated legislation, as well as new or updated requirements defined by related stakeholders concerned with the business, can produce an impact on the achievement of digital preservation requirements.

Some risks can remain unnoticed for a long period of time. For instance, a damaged hard disk sector can remain undetected until a data integrity validation or hard disk check is performed. Furthermore, one cannot assume threat independence, since a specific threat can generate other threats.

Considering the risks to digital preservation, this paper claims that a "digital preservation system" is itself an infrastructure in risk, comprising strong safety requirements, as happens in civil engineering structures. Moreover, since the safety of large civil engineering structures is directly dependent to the monitoring systems and the preservation of the associated data, the consequences of a failure in the preservation system can also produce catastrophic effects (e.g. loss of life, environmental damage, etc.).

IV. SOLUTION OVERVIEW

When addressing the problem of digital preservation for memory institutions (e.g., libraries, archives, museums) where the digital data to be preserved are typical static documents (e.g., images, text documents), it is a common accepted solution to apply the OAIS reference model, since the information package is composed by the digital data and a set of metadata associated with them. The ultimate objective of solutions based on the OAIS reference model is to mitigate the risks identified in Section III.

Research undertaken in the SHAMAN project determined that a bigger understanding of the context surrounding the production, preservation, and reuse of information (OAIS view) was needed in order to understand its implications on preservation. Thus, a model of the lifecycle of information was created and can be seen in Figure 1.

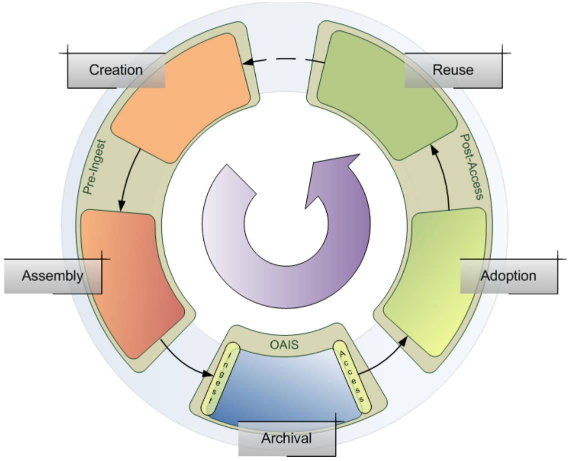

Figure 1. Information lifecycle [15]

During the *creation* phase new information comes into existence. Normally, information is not created for the purpose of archiving, thus archivable information can be the result of complex processes that involve a multitude of stakeholders. The *assembly* phase deals with the appraisal of information relevant for archival and all processing and enrichment for compiling the complete information to support future reuse. Normally, this compilation is called an *archival package*. The *archival* phase corresponds to the OAIS reference model and addresses the life-time of the digital data inside the archive, including the ingestion of and access to information. The *adoption* phase encompasses all processes by which information provided by the Archive is screened, examined, adapted, and integrated for the proper reuse. Finally, the *reuse* phase deals with the exploitation of information in the interests of the consumer.

The preservation of dam safety sensor data raises several challenges because of the data and process characteristics. First, data is not static (a data set is continuously increasing). Second, since new sensors (with different characteristics and results) have to be accommodated in the future, new data representations must be handled. Third, the representation of a dataset can evolve in the future (new devices can use different representations to store the same data), limiting the ability to understand the same type of data, as well as relating the same type of data when it was captured by devices using different data representations. Finally, the nature if complex and interlinked objects composed by datasets and their representation (an isolated dataset is useless to interpret the structural behaviour).

In order to control the complexity of data representations, some communities developed their own metadata initiatives as, for instance, the *Ecological Metadata Language (EML[3]),* or the *Federal Geographical Data Committee (FGDC[4]).* Yet, there will never be a unified metadata schema for all possible data. Thus, in a scenario that is not covered by current metadata initiatives, or when the information can be represented in heterogeneous schemas that can continuously change (like the sensors used in the civil engineering domain), the use of standard languages to describe data representations [16] is an expected solution.

The SHAMAN project developed an archival infrastructure that follows the OAIS reference model and uses the iRODS[5] data grid as storage substrate. The work presented here relies on this infrastructure to address the digital preservation risks related to media faults, process and infrastructure vulnerabilities, as well as the issues related to the volume of data and its imminent deluge. However, in the dam safety context, an information package is an interlinked object that must aggregate the sensor data, the information on the sensors that produced the data, as well as the description of the schemas used to encode them (it can include syntactic and semantic representation). These activities are part of the *creation* and *assembly* phases, which will influence the future adoption and reuse of this information.

The proposed solution elaborates on the creation and description of information packages to control the media obsolescence vulnerabilities that occur when the representation format becomes obsolete and unable to be rendered, even if the "bit stream" survives over time. Since the information package is composed by sensor data (from distinct types of sensors), along with their contextual and representation information, a network of object have to be aggregated to create a meaningful object in the context of civil engineering. The Metadata Encoding and Transmission Standard (METS[6]) is a widespread metadata representation to encode structural metadata in XML. The use of METS provides and extensible way to represent the aggregations required by the illustrated scenario.

On the other hand, to address the management of schema representations (including the definition of sensors, raw data, processed readings, etc.) and their dynamic nature (new or updated schemas to represent the same information), it is critical to manage metadata that describes the information representation. This is not a new requirement in the community, where, for instance, previous work developed the *Metacat* framework [17], which is able to store, retrieve and transform XML documents managed stored in a relational database. In this paper, we use the concept of Metadata

---

[3] *http://knb.ecoinformatics.org/software/eml*

[4] *http://www.fgdc.gov*

[5] *https://www.irods.org*

[6] *http://www.loc.gov/standards/mets*

Registry (MDR), which was conceived to represent a system that allows the management of multiple schemas (not limited to XML) and the export of information about the schema. It also supports the creation and management of mappings between different schemas. This concept is formalized by the ISO 11179[7] series of standards. Accordingly, the MDR can be used to address the challenges of representing the encoding of sensor data (including the definition of sensors and the data stages listed in TABLE I. also supporting the future migration of information packages.

For demonstrating the preservation of data in this particular scenario, we developed a *Service Oriented Architecture (SOA)* solution, as shown in Figure 2. Our proposal comprises services for: acquiring data stored in the *GestBarragens* information system, acquiring a description of the schema representation, packaging the data together, and ingesting the data package into the archival system. Such a solution is controlled by a service orchestrator (Service Orchestration component) parameterized in *Business Process Execution Language (BPEL)* and executed by a *GlassFish Open ESB BPEL* engine. This way, the BPEL file representing the creation/assembly process of the information package is itself part of the package, which is critical for provenance purposes.

The following components implement *Java Web Services* that are orchestrated by the Service Orchestration component:

- **Data Extractor**: Extracts data from the *GestBarragens* system, according to the parameters defined by the *Assembly Orchestration.* To support the dynamic nature of sensor data, it has the option to define the time window for data extraction, full extraction, incremental extraction and the list of dams to extract. The recursive use of full data sets uses more space, while incremental data sets require the recomposition of data sets on access.
- **Metadata Registry** (MDR): Supports the registration and management of multiple data schemas, addressing "the semantics of data", "the representation of data", and "the registration of the descriptions of that data". It also supports the creation and management of the mappings between data schemas, as well as the export of both schema and mapping information.
- **Data Aggregator**: the METS schema is used to "wrap" all the information, acting as structural metadata. The information that is aggregated by the METS includes: (*i*) data about the characteristics of the sensor which produced the readings (e.g., calibration constants, validation intervals, etc.), (*ii*) schema information of the data containing the characteristics of sensors, (*iii*) observational data, (*iv*) schema information of the observational data, (*v*) BPEL file representing the assembly process, and (*vi*) generated HTML files to facilitate human navigation under the METS components.

The *Assembly Orchestration* component starts by (1) acquiring observational data and sensor information from the *GestBarragens* through the Data Extractor component, specifying both the type of export (time window, full, or incremental from last export) and the dam or list of dams, and (2) acquiring the related schema information by requesting it from the MDR, depending on the type of data exported in the previous step. It continues, by (3) requesting the generation of a METS file to package the dam information and (4) submitting the package into the archive (using the *ingest Web Service* of the SHAMAN archive). When the submitted package enters the SHAMAN archive, it is then managed as common information packages, as those constructed for typical data objects like images or text. In the case of the SHAMAN archive, an information package is encoded in plain zip and includes an OAI-ORE manifest[8] to aggregate resources contained in the information packages (e.g., information content, preservation metadata).

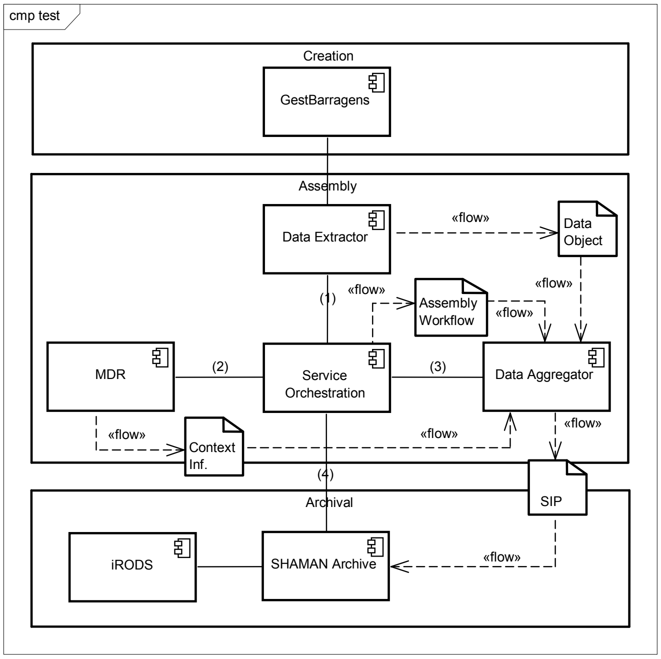

Figure 2. Overview of the proposed solution for digital preservation integrated with the GestBarragens information sytem.

When the dam safety data is accessed from the archive for future use (*adoption*), the information package is self-contained, in the sense that it includes, not only the preserved data, but also all the information required to render this data (structural information provided by the schema representation extracted from the MDR), in addition to the context information required to understand the data itself (context information like the type and characteristics of sensors, location, data units, etc.).

Finally, the integration of a MDR and the decoupling between data, its schema representation and the context information in the information packages, support the use of migration techniques inside the archive. In fact, migration is one of the most effective techniques used in digital preservation to avoid the obsolescence of data representations/formats. For observational data, mappings between schemas supported by the MDR provide an effective

---

[7] *http://metadata-standards.org/11179*

[8] *http://www.openarchives.org/ore*

tool to migrate from an obsolete schema representation to an updated representation. Note that migrations can be often lossy, making it critical to plan when, how and what to migrate [18].

The use of a SOA architecture, and the respective service independence, allows the adoption of this solution to several scenarios. From the proposed services, only the *Data Extractor* is scenario dependent. It is also independent from the archival solution. It only requires an archival service that can be accessed through a *Web Service* that can be configured and called by any BPEL engine.

V. CONCLUSIONS AND FUTURE WORK

This paper is motivated by the real case study of managing data concerning the safety of large civil engineering structures. It describes the technological solutions that are being used in LNEC and shows that these solutions were designed for operational purposes and do not address emergent digital preservation requirements. For instance, in this type of scenario, future research requires details about provenance and production workflows (e.g. conversion from raw data to engineering quantities) that are not currently handled. From the analysis of scenarios handled at LNEC, this paper motivates the need for digital preservation and surveys its main requirements and threats. This analysis shows that a preservation system is itself an infrastructure in risk, requiring continuous actions to be safe for a long period of time. Since the safety of civil engineering structures is directly connected to the underlying monitoring data, a failure in the preservation system can potentially produce catastrophic consequences.

In the scope of the SHAMAN project, a digital archive supported by an iRODS data grid infrastructure was developed. This archive adopts the widely accepted OAIS reference model, where digital data is packaged with metadata to be reused in the future. This paper motivates the need to extend this model in order to support complex and dynamic digital data, as those generated by sensors that continuously monitor the behavior of the dam structures. The proposed solution comprises components to manage the complexity of heterogeneous schemas, control the dynamic behavior of datasets, and aggregate context information in a meaningful way. The assembly of this information is also a complex process, which motivates the control of this process through a service orchestrator.

The problem of preserving sensor data differs from the preservation of traditional documents, mainly because sensor data is dynamic and might have heterogeneous data representations. Moreover, the context information is much more complex, since sensor data depends on the sensor properties, calibration constants, etc.

The proposed solution, adopted in the scope of the SHAMAN project, is able to preserve sensor data. However, as future work, the dependencies of sensor data both in the infrastructure, but also on the acquisition process, require the problem to be addressed from the perspective of the overall business process, instead of a data-centric approach. This is the vision proposed by the TIMBUS[9] project, where digital preservation is seen as a business continuity issue, where business processes that ran in the past should be able to be reproduced in the future. In the case of sensor data, the recreation of the overall production environment (to simulate the sensor data acquisition) can be used to study the behavior of structures under a controlled (simulated real) environment.

ACKNOWLEDGMENT

This work was supported by FCT (INESC-ID multiannual funding) through the PIDDAC Program funds and by the projects SHAMAN and TIMBUS, partially funded by the EU under the FP7 contracts 216736 and 269940.

REFERENCES

[1] M. Wieland and R. Mueller. Dam safety, emergency action plans and water alarm systems. International Water Power & Dam Construction January 2009

[2] IEEE Standard Computer Dictionary: A Compilation of IEEE Standard Computer Glossaries, 1990.

[3] P. Lord and A. Macdonald. E-Science curation report – Data curation for e-Science in the UK: an audit to establish requirements for future curation and provision, 2003.

[4] Consultative Committee on Space Data Systems, ISO 14721:2003 - Reference model for an open archival information system (OAIS), 2003.

[5] ASCE – American Society of Civil Engineers. Guidelines for instrumentation and measurements for monitoring dam performance. 2000. ISBN 0-7844-0531-X.

[6] ICOLD - International Commission on Large Dams. Guidelines for automated dam monitoring systems. 1999.

[7] RSB. Dam safety regulation, DL n.344/2007, October 15th. DR, Lisbon, 2007 (in Portuguese).

[8] M. Franklin, A. Halevy and D. Maier. From databases to dataspaces: a new abstraction for information management. SIGMOD Record, 34(4):27-33, 2005.

[9] B. Ludascher, I. Altintas, C. Berkley, D.Higgins, E. Jaeger, M. Jones, E. Lee, J. Tao, Y. Zhao. Scientific Workflow Management and the KEPLER System. March, 2005.

[10] I. Altintas. Lifecycle of Scientific Workflows and Their Provenance: A Usage Perspective. IEEE Congress on Services. 2008.

[11] J. Barateiro, G. Antunes, M. Cabral, J. Borbinha, and R. Rodrigues. Using a GRID for digital preservation. International Conference on Asian Pacific Digital Libraries, Bali, Indonesia, December 2008.

[12] J. Borbinha. Authority control in the world of metadata. Cataloging Classification Quarterly, 38 Issue: 3/4, 2004.

[13] A. Chervenak, I. Foster, C. Kesselman, C. Salisbury, and Tuecke. The data grid: Towards an architecture for the distributed management and analysis of large scientific datasets. Journal of Network and Computer Applications, 23:187-200, 2000.

[14] J. Barateiro, G. Antunes, F. Freitas and J. Borbinha. Designing Digital Preservation Solutions: A Risk Management-Based Approach. International Journal of Digital Curation 5(1):5-17, 2010.

[15] H. Brooks, A. Kranstedt, G. Jaschke and M. Hemmje. Modeling context for digital preservation. Smart Information Knowledge Management, 197-226, 2010.

[16] M. Westhead, T. Wen and R. Carroll. Describing data on the GRID. GRID Computing, 134-140, 2003.

[17] M. Jones, C. Berkley, J. Bojilova and M. Schildhauer. Managing Scientific Metadata. IEEE Internet Computing 5: 59-68, 2001.

[18] C, Becker, H. Kulovits, M. Guttenbrunner, S. Strodl, A. Rauber, and H. Hofman. Systematic planning for digital preservation: Evaluating potential strategies and building preservation plans. International Journal on Digital Libraries (IJDL), 10(4):133-157, December 2009.

---

[9] *http://timbusproject.net*