Requirements Elicitation for a Long Term Digital Preservation System: A Case Study from the Financial Sector

Claudia-Melania Chituc
University of Porto, Faculty of Engineering
Informatics Engineering Department
Portugal
cmchituc@fe.up.pt

Petra Ristau
JRC Capital Management Consultancy & Research GmbH
Germany
+49-30-847 88 22-0
pristau@jrconline.com

ABSTRACT
Companies face challenges towards designing and implementing a preservation system to store the increasing amounts of digital data they produce and collect. The financial sector, in particular the investment business, is characterized by constantly increasing volumes of high frequency market and transaction data which need to be kept for long periods of time (e.g., due to regulatory compliance). Designing and developing a system ensuring long term preservation of digital data for this sector is a complex and difficult process. The work presented in this article has two main objectives: (1) to exhibit preservation challenges for the financial sector/ investment business, and (2) to present and discuss preliminary results of the requirements elicitation process, with focus on the financial sector - work pursued towards the design and development of a preservation system, within the scope of the on-going R&D FP7 project ENSURE – *Enabling kNowledge Sustainability Usability and Recovery for Economic value* (http://ensure-fp7.eu). Requirements, use cases and scenarios identified for the financial sector are presented and discussed. The agenda for future research work is also presented.

Categories and Subject Descriptors
H.3.4 [Information Storage and Retrieval]: Systems and Software

General Terms
Documentation, Design.

Keywords
Requirements elicitation, financial sector, digital preservation system.

1. INTRODUCTION
The financial domain, in particular the investment business, is characterized by the increasingly incoming streams of high frequency market data. The digital data produced and collected by financial institutions (e.g., market data, transactions data) needs to be preserved for long term (e.g., for regulatory compliance, research purposes). While in the last decade a particular focus of R&D in the financial business was on performance improvements of the IT-infrastructure in an attempt to keep pace with the constantly increasing volumes of data, today the need of financial institutes for support in compliance to regulations and legal standards takes an increasingly important role. In the light of the financial crisis, it can be well expected that the relevance of this issue will rise further, since various expansions of regulations are being discussed, like, for example, full transparency of transactions [1][2].

The design and development of a system ensuring long-term preservation of digital data is a complex and difficult process. Existing approaches focus mainly on preserving homogeneous data (e.g., cultural heritage data). In the case of the financial sector, the task to elaborate a preservation system is even more challenging since it is required to ensure long term usability of heterogeneous data, integrity and authenticity of significant intellectual property or highly personal data, while conforming to regulatory, contractual and legal requirements. Such challenges are addressed by the on-going R&D FP7 project ENSURE – *Enabling kNowledge Sustainability Usability and Recovery for Economic value* (http://ensure-fp7.eu) targeting the financial sector, health care, and clinical trials domains.

Once the necessity to elaborate a system ensuring long term preservation of digital data is expressed, the process of requirements elicitation starts. During the requirements collection and assessment phase, the properties (e.g., functional and non-functional properties) that the system should have to meet the stakeholder's needs are identified.

The work presented in this article has two main objectives: (1) to exhibit preservation challenges for the financial sector/ investment business, and (2) to present and discuss preliminary

results of the requirements elicitation process pursued towards the design and development of a preservation system, within the scope of the on-going R&D FP7 project ENSURE – *Enabling kNowledge Sustainability Usability and Recovery for Economic value* (http://ensure-fp7.eu). Relevant functional and non-functional requirements, use cases and scenarios identified for the financial sector are presented and discussed.

The rest of this article is organized as follows. The next section briefly introduces challenges in requirements elicitation in the context of long term digital preservation (LTDP). Section three describes the elicitation approach taken in the ENSURE FP7 project. Section four then portrays main characteristics of the financial sector and challenges towards long term digital preservation identified in this project. Relevant requirements, use cases and scenarios identified for the financial sector towards the design and development of the ENSURE Preservation System are then presented and discussed. The conclusions of the work pursued and directions of future research work are presented in section six.

2. LONG TERM DIGITAL PRESERVATION

2.1 Definition and Preconditions

Among the definitions found in literature, we exemplarily cite the following three. Digital preservation concerns the processes and activities related to preserving digital data over long periods of time, ensuring its accessibility and usability to future generations [3]. Digital preservation involves the retention of digital data/ object, and its meaning [4]. According to the Consultative Committee for Space Data Systems [5], long term preservation refers to the act of maintaining information, independently understandable by a specific community, supporting its authenticity over the long term.

The OAIS Reference Model (e.g., [5],[6]) presents a technical recommendation establishing a common framework of terms and concepts which make up an Open Archive Information System (OAIS). It comprises six functional entities (and related interfaces): Ingest, Archival Storage, Data Management, Administration, Preservation Planning, and Access.

Ensuring long term digital preservation (LTDP) is a complex process, and several challenges need to be addressed, such as: digital (technology) obsolescence, lack of standards and generally accepted methods for preserving digital information, deterioration (e.g., of digital data recording media). Although several approaches for digital preservation exist (e.g., emulation, information migration, encapsulation), as emphasized in [4], still, after ten years, there is a lack of proven preservation methods to ensure that the preserved digital data will (continue) to be readable after long periods of time, e.g., 20 years, 50 years, 100 years.

Although initially the focus has been on relatively homogeneous heritage data, currently organizations from the private sector (e.g., financial institutions, private clinics and hospitals) are increasingly concerned with preserving the growing amounts of digital data they produce and collect. These data tend to be heterogeneous, which represents an additional challenge for the preservation process itself, but also for the elicitation of requirements.

2.2 Challenges in Requirements Elicitation

Commonly used approaches for requirements elicitation in the context of LTDP are: the questionnaire (e.g., CASPAR FP6 project [7]), and internal surveys and interviews (e.g., KEEP FP7 project [8]).

Although these methods for requirements elicitation allowed the identification of requirements and scenarios, they have several limitations, such as: lack of standardized procedures for structuring the information received through interviews, difficulty to integrate different answers/ interpretations, different goals, different communication styles or terminology into a single requirement.

In the context of long term digital preservation, several challenges appear in the requirements elicitation phase, such as: difficulty to validate the collected requirements (e.g., due to the very long life-time of a preservation system and preservation period), hardware and software constraints during system's design and implementation phases, changing requirements over time. The functionalities of a preservation system may also change over time. It is also very challenging to integrate into a single requirement (or a set of requirements) needs expressed by stakeholders from different industry sectors for one preservation system.

3. REQUIREMENTS ELICITATION FOR THE ENSURE PRESERVATION SYSTEM

3.1 The ENSURE FP7 Project

The on-going ENSURE FP7 project aims at extending the state-of-the-art in digital preservation by building a self-configurating software stack addressing both the configuration and preservation lifecycle processes in order to create a financially viable solution for a set of predefined requirements [9]. It analyzes the tradeoff between the costs of preservation against the value of preserved data, addressing also quality issues. ENSURE draws on actual use cases from health care, clinical trials, and financial services.

ENSURE Reference Architecture for long term digital preservation is based around four main areas of innovation[1]: i) evaluating cost and value for different digital preservation solutions, ii) automation of the preservation lifecycle in a way which can integrate with organizations' existing workflow processes, iii) content-aware long term data protection to address privacy issues like new and changing regulations, and iv) obtaining a scalable solution by leveraging wider ICT innovations such as cloud technology.

3.2 Requirements Elicitation Approach

Considering the specificities of the LTDP domain, the objectives of the ENSURE FP7 project, and the characteristics of the three sectors on which this project is focusing on (financial, health-care and clinical trials), a use-case scenario approach has been

[1] http://ensure-fp7.eu

chosen for requirements elicitation, to reflect all the tasks the stakeholders will need to perform with the ENSURE preservation system. This was then combined with a traditional elicitation approach (e.g., where the stakeholders indicate what they want the system to do).

The use-case approach has been successfully used during the last years for requirements elicitation and modeling, e.g., [10], [11], [12]. As emphasized in [13], the objective of the use-case approach is to describe all the tasks the users/ stakeholders will need to perform with the system. Each use case can consist of several scenarios capturing user requirements by determining a sequence of interactions between the stakeholder and the system.

Within the scope of this work, use case scenarios were regarded as assets that can be exploited as reference scenarios within the context of the ENSURE project and the ENSURE Preservation System [1][2]. Thus, the description of the ENSURE use case scenarios focused on the expectations the technological solution (e.g., the ENSURE preservation system) should address. Similar to [10], these descriptions should bring out the goals (and assumptions) the technological solutions should encompass.

Although the use case approach is successfully used for capturing requirements (e.g., [10]), several weaknesses are documented, e.g., [14], such as: the use cases are written from the system's (and not stakeholders') point of view.

To avoid such pitfalls, a well-documented template for requirements elicitation was provided to the stakeholders (e.g., targeting functional and non-functional requirements, use cases and scenarios). In addition, sector-specific discussion groups/ teleconferences were set where experts in the area of LTDP participated, as well as representatives of the Architecture team.

The use of this approach allowed the identification of: functional and non-functional requirements for the ENSURE preservation system, stakeholder's constraints for the technical solution, main classes of users, as well as the identification and documentation of use cases and scenarios.

4. THE FINANCIAL SECTOR

4.1 Main Characteristics and Challenges towards Long Term Preservation of Financial Digital Data

Before going into details about the data itself, we would like to address the sources and flows of the data as well as the stakeholders dealing with the preservation system. For the requirements elicitation process previously described, we restricted ourselves to the sector of investment banks and smaller financial investment institutes, whose core business is to offer advisory and asset management services to institutional as well as private investors.

So, one large source of data consists of the transaction documentation mainly in form of individual account statements from the custodian bank, broker or clearing house for each client. The constantly increasing amounts of documents go along with sharply increased requirements on information protection, safekeeping and risk management, and contain expanded overall record retention obligations issued by the regulating public authorities an both national and European level. This concerns, for example, any information received from or provided to the client that needs to be preserved for the whole duration of the contractual relationship that could actually overpass the minimum statutory preservation time of typically 5 years.

On the other hand, almost all investment institutes base nowadays their advisory services on decision support systems for forecasting and algorithmic trading. These systems automatically analyze incoming real time data and come to trading suggestions when to buy and sell a specific instrument. Before actually being applied to real time data, the underlying trading models are developed and tested on the basis of long histories of market data, often reaching back for more than 5 to 10 years.

While there are no regulatory or legal requirements on the retention of these large volumes of high frequency *financial market data,* they have to be stored and retained, due to their associated business value, i.e. their high recovery and restocking costs [1][2].

Although each institute and situation is unique, there are common organizational characteristics shared by all investment institutes.

Figure 1 portrays the flows of data between the different departments of a typical investment bank:

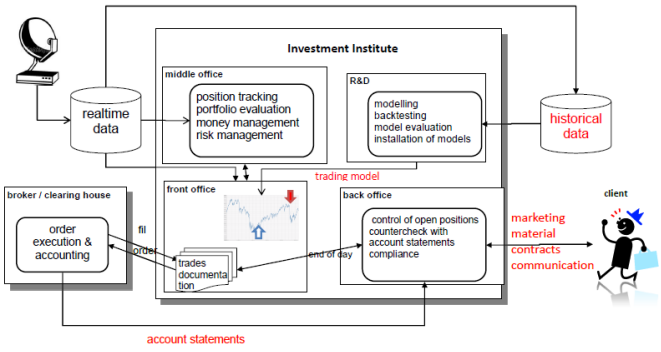

Figure 1. Flows of data between departments in an investment bank

Market Data is received from the real time feed and distributed to the historical data base from where it is accessed by the **financial engineers** located in the R&D department. Besides, it is of course immediately displayed on the charts on the traders' desktops in the front office as well as fed to the middle office, where it is used for tracking of open positions and risk management.

Client data, i.e. account statements, are received from the custodian bank, broker or clearing house. It is the **back office staff** that manages this type of data together with any other client related data and is hence responsible for the preservation of this type of data.

Trading Models, finally, are developed by the R&D department and installed at the trader's desks. So, the financial engineers are

the responsible actors for the retention of all models and market data.

A big challenge in the financial sector is the heterogeneity of the data that needs to be stored in the preservation system. For example, the *client information* consists of heterogeneous data, such as: contractual paper work, marketing information, digital information (e.g., e-mail correspondence, web content, electronic account statements, accompanying notes, like telephone protocols, front-office documentation about trading decisions and order execution).

Market data, on the other hand, may exist in heterogeneous formats as well. Technological developments on both data feed side but also concerning charting and modeling software typically lead to occasional switches between data and software providers. So the data format may vary between different proprietary formats from the various vendors as well as source and software independent formats, typically in comma separated ASCII files (*csv*), distributed by third party vendors of historical data.

Finally, the trading models themselves shall be preserved and build a third category of data. Although it is not a requirement yet, the preservation of software applications should also be considered since trading models critically depend on software versioning, e.g., in order to avoid eventual deviations in model interpretation.

Next are briefly described **data formats** for each category:

- *Client Documents* need to be kept in paper form (e.g., signed originals) and digitalized form (e.g., client contract, risk disclosure information and limited power of attorney will build the foundation of a client's record). Only standard formats will be used to preserve the data in the preservation system, e.g., *.pdf*, *.jpg*. In addition, some current MS Office document formats will have to be supported by the preservation system, such as protocols of telephone conversations.
- *Financial Market Data*. With the constant introduction of new financial trading instruments, but also observable in the well-established high volume instruments (e.g., Forex, Stock Index Futures like Dax and Nasdaq), the financial markets are characterized by extraordinary increasing data loads during the past years. Market data has to be distinguished by its sources (e.g., Thomas Reuters, Bloomberg, Morningstar), which use different data formats. It refers to numerical price data, reported from trading venues, such as stock exchange, but also from non-regulated interbank markets (i.e., foreign exchange rates). The price data is attached to a ticker symbol and additional data about the trade like stamp and volume information. The data is then stored in a database, typically one that is integrated with a charting environment software used by the trader to graphically display the data. Although the databases have their own proprietary format, almost all of them offer the possibility of exporting data to *csv* format.

  A sample of financial market data in such a general format is illustrated in *Figure 2*. The column descriptors are contained in the first line (or in a separate file), and data is given in a ten minute compression format, i.e., all price data arriving within a 10 minute time window is reduced to 4 values only: the first price (Open), the highest (H); lowest (L) and last one (Close), the number of up-ticks (U, prices higher than the previous one), and down-ticks (D, prices less or equal to the previous one).

```
"Date","Time","O","H","L","C","U","D"
08/05/2009,1530,95.28,95.33,95.26,95.32,1528,1526
08/05/2009,1540,95.32,95.33,95.21,95.28,2520,2574
08/05/2009,1550,95.28,95.31,95.21,95.22,2634,2666
08/05/2009,1600,95.22,95.29,95.13,95.28,2936,2842
08/05/2009,1610,95.29,95.29,94.96,94.97,3930,3882
08/05/2009,1620,94.97,95.07,94.86,94.94,3710,3612
08/05/2009,1630,94.94,95.00,94.92,94.96,2946,2768
08/05/2009,1640,94.94,94.98,94.89,94.96,2526,2452
08/05/2009,1650,94.96,94.98,94.86,94.92,2618,2624
```

Figure 2. Sample of Financial Market Data
(Source: [1-2])

- *Trading Models* shall be preserved as code written within a proprietary programming language (e.g., depending on the modeling environment used), which typically contain some commands close to natural languages, allowing also trader with no or restricted programming skills to formulate their own trading rules. *Figure 3* presents a simple example of a trading model code snippet.

```
if Close crosses above Average(Close,10) then Alert;
If Marketposition=0 and Barssinceexit(1)> Barsflat  and
                  if Close < Close[1]
then Sell Short 10 contracts next bar at Market;
```

Figure 3. Example of Trading Model Data (Source: [1-2])

Challenges for the long term digital preservation of the financial data concern not only the heterogeneity of the data that needs to be preserved, but also the retention of the preserved information, e.g., the retention of client information, the retention of proprietary applications due to business purposes, and the retention of very large amounts of market data stored over time, while meeting regulatory directives and business goals.

The system ensuring the preservation of data for the financial sector needs also to allow conformance to regulatory, contractual and legal requirements, and management of long term authenticity and integrity of intellectual property and personal data.

With these considerations, an extensive work has been pursued to identify functional and non-functional requirements, use cases and scenarios for the financial sector for a digital preservation system.

4.2 Main Findings: Use Cases, Scenarios, Functional and Non-Functional Requirements

Main stakeholders for the ENSURE Preservation System, for the financial sector, include [1][2]:

- *Back-office Staff*, the primary user of the ENSURE Preservation System, is responsible for all administrative tasks related to the organization's clients.
- *Model Developer/ Financial Engineer* is the employee of the organization responsible for the implementation, development and testing of the trading models.
- *Auditor* is an actor external to the organization, who periodically checks the compliance to regulatory standards and completeness of documentation.
- *System Administrator* is the employee responsible for the technical management of the organization's system, e.g., installing and technically managing the ENSURE Preservation System in a specific context.

The UML use case representation of the ENSURE Preservation System functionality for the financial domain is illustrated in *Figure 4*. The interactions between the actors and the ENSURE Preservation System are indicated.

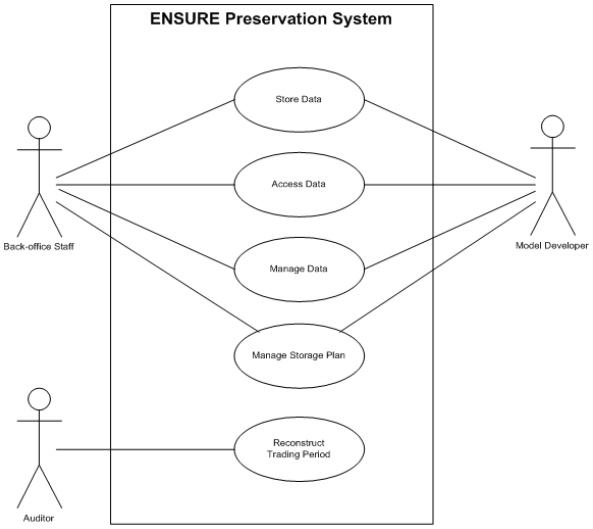

Figure 4. Main actors for the financial domain and the ENSURE Preservation System (UML) (Source: [1][2])

A brief description of the use cases illustrated in *Figure 4* is presented in *Table 1*.

Table 1. Overview of the Use Cases

Use Case	Description
Store Data	Data that is not accessed regularly is being preserved due to regulatory requirements or because it represents a business value.
Access Data	The data being preserved is accessed by the Back-office Staff or Financial Engineer.
Manage Data	Accompanying descriptions or metadata related to existing stored data of any type may have changed and needs to be updated, e.g., a client contract may have terminated or a trading model may no longer be in use. It is important to update these changes; especially in the case of client documents the safekeeping period starts with the termination date.
Manage Storage Plan	A storage plan describing how and where to store which type of data is set up once and used for each storage/ access operation. With respect to the available storage hierarchy and considering the related costs and risks of each device, the system determines the optimal storage location for each type of data and makes a suggestion to the user. The user may agree to the suggested plan or alter it according to additional constraints. The plan remains then fixed until it is updated on user request, e.g., when new storage devices are available or when related costs have changed.
Audit documents	A sample of client documents is requested (e.g. client XY for the period 01/2011 – 12/2011) by the auditor. It is then checked for correctness and in particular for completeness.

Table 2 contains a brief description of the *Store Market Data* scenario for the financial sector.

Table 2. Store Market Data Scenario

(Source: adapted after [1] and [2])

Name of scenario	Store market data
Actors	Financial Engineer, System Administrator
Operational Description	Once a month market data for all markets received by the real time data feed will be stored in the ENSURE Preservation System.
Problems, Challenges	Data integrity. Data protection. Large amounts of input data in one chunk.
Expected Benefits	Prevent real time database from overload. Secure business value of data.
Risks	Losing or corrupting data. Degradation of system performance during ingest of data.

Figure 5 illustrates the UML sequence diagram for the Back-office Staff for a simple scenario on client data request. After

the *Back-office staff* submits the login and password, the ENSURE preservation system shall send a success/failure notification. A successful login allows the *Back-office staff* to submit requests (e.g., query on client data, query on client name). The ENSURE preservation system shall verify the access rights for the query received and return the data requested.

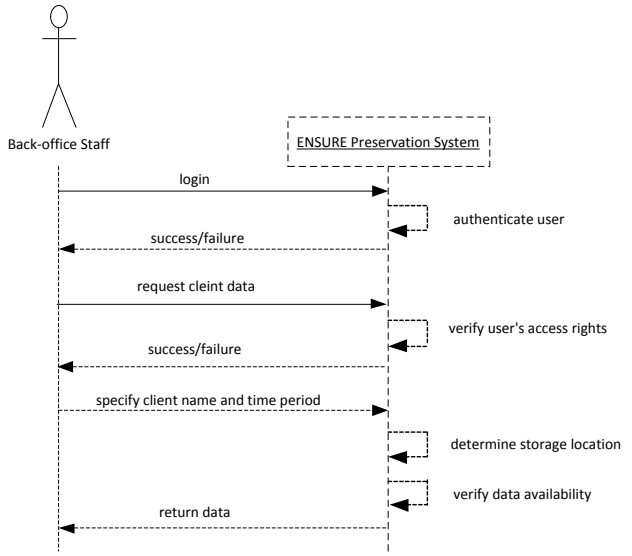

Figure 5. Client data request sequence diagram for the ENSURE Preservation System (UML)

The approach used also allowed the collection of functional and non-functional requirements. Five relevant *functional requirements* for the ENSURE preservation system identified for the financial sector are listed below [1][2]:

- *Authenticate User*. The ENSURE preservation system shall ensure that only authorized users have access to the system and are restricted to executing only those tasks necessary within their specific assigned role.
- *Encrypt Data*. The financial data is subject to security restrictions and shall be encrypted before being stored on an external storage device. The ENSURE preservation system shall detect automatically which data has to be encrypted and shall start the encryption process.
- *Notify Registration*. When a user is registered by the System Administrator, the ENSURE preservation system shall send an e-mail to the user with his/ her access details.
- *Keep Track of Minimum Regulatory Storage Times*. The ENSURE preservation system shall attach to each record of client documents the minimum storage duration. This requires also the possibility of entering the date when a client relationship has ended. The ENSURE preservation system shall automatically label the data with the earliest possible deletion date.
- *Delete Old Client Documents*. Client documents that have reached the minimum storage duration shall be deleted in order to relieve the system from unnecessary volume. The ENSURE preservation system shall detect such data automatically and ask for user confirmation before deleting the data.
- *Evaluate and optimize a preservation solution* considering cost, quality and economic performance in order to support differently experienced users.
- *Allow for the Reproduction of Results of Trading Models* (which is version sensitive).

Examples of *non-functional requirements* for the ENSURE preservation system identified for the financial sector are [1][2]:

- *Availability*. The ENSURE preservation system shall be available every working day during normal working hours (e.g., 9h-18h). Since the storage of data will be done in weekly or monthly periods, and retrieval will be executed without time constraints, non-availability of the system for less than 24 hours is not critical.
- *Reliability*. Resilience and reliability of the system are critical factors with respect to compliance to regulations, e.g., loss of data is not acceptable.
- *Integrity*. Data integrity is very important due to compliance to regulatory prescriptions (e.g., changes or partial loss of client data is not acceptable), except for market data.
- *Confidentiality*, which concerns client documents and trading models. Stored trading models contain proprietary know-how and the essence and experience of many years of research. For this reason, the integrated trading models represent a high economic value for a financial institution that has to be protected against unauthorized use and reverse engineering attacks. Furthermore, in order to ensure the correct restoration of trading signals, the models have to be protected against external manipulation and corruption. The sensitive nature and strategic value of both such data cannot tolerate any unauthorized access or usage.

5. DISCUSSION

The data considered for long term preservation in the financial use case within the ENSURE project consists of client data (including transaction documentation as, for example, daily account statements) on the one hand, that has to be stored due to record retention obligations according to legal regulations like the EU's Markets in Financial Instruments Directive[2] (MiFiD), and market price data histories on the other hand needed by researchers and market analysts in order to build prognosis models to be used by the traders as decision support systems. As different as the goals for data preservation – legal reasons for client data and economic reasons due to high costs for

[2] http://europa.eu/legislation_summaries/internal_market/single_market_services/financial_services_general_framework/l24036e_en.htm

repurchasing, filtering and format conversions for market data – as different are the consequences of data loss and security violations on both data classes.

Since the legal consequences of losing client data can be severe, up to endangering the continuance of business execution, the preservation plan has to fulfill highest service level with respect to security standards including duplication of data and multiple fixity checks. For the same reason, cost considerations will only come in the second place for this type of data, after all data security requirements have been met. For market data, in contrast, storage costs represent the main decision criteria during preservation plan generation.

Another major criterion for the preservation strategy stems from the different data formats found in both data classes. While client data is typically stored in widespread document types like *pdf*-s, market data format is determined by the origin of the data (i.e., the data feed provider) and/or the proprietary data format of the data base, where the data is fed to. While the handling of such common data formats like *.pdf*-files does not represent a major challenge to a preservation system, proprietary data formats may call for two options to be followed: they may either be converted to more common or "standard" formats, independent from the original application software, or a virtualization path may be followed, preserving the whole application package, including data base, and making them accessible through the use of virtual machines.

A particular challenge consists of the fact that, as far as market data is concerned, its value decreases with its age. The target preservation plan shall therefore distinguish between relatively *recent data*, that has to be kept as the complete data set, so called *tick data*, consisting of every single price change, leading to hundreds of prices per minute, and *older data*, where compressed data format (so called OHLC[3] bar representation) would be acceptable, reducing the data to only four values per time period. The resulting information loss would be acceptable, as well as a downgrade of the service level.

Following the path indicated in [4] for the selection of preservation techniques, all of the above considerations lead to a twofold strategy. Although the complexity of the digital resource is low for both data classes, the data format is not known as far as market data is concerned. In this case, an emulation solution is recommended for market data preservation while encapsulation or migration would be the technique of choice regarding client data.

6. CONCLUSIONS AND FUTURE WORK

Requirements elicitation for a preservation system is a complex task, mainly due to the challenges associated to long term digital preservation, such as: difficulty to validate and test a preservation system due to its very long lifetime, the data and technology to be preserved are exposed to obsolescence.

So far, most approaches towards ensuring long term preservation of digital data focused on relatively homogeneous data (e.g., cultural heritage data). In this article, challenges for the financial sector towards the design and development of the ENSURE preservation system were identified and discussed, e.g., data heterogeneity, conformance to regulatory, contractual and legal requirements, integrity of preserved data.

The ENSURE preservation system aims at providing support to perform an analysis of cost, economic performance assessment and quality for a preservation solution, which represents a novelty compared to other initiatives (e.g., iRODS[4] policy-based data management system).

Results of the requirements elicitation process, with focus on the financial sector have been presented (e.g., functional and non-functional requirements, use cases and scenarios) which reflect the work pursued towards the design and development of a preservation system, within the scope of the on-going R&D FP7 project ENSURE: *Enabling kNowledge Sustainability Usability and Recovery for Economic value* (http://ensure-fp7.eu). The approach used in requirements elicitation for the ENSURE preservation system was also described.

The main result of the analysis presented in this article concerning the financial sector was that due to the heterogeneous character of the data to be stored a combined strategy will most probably be the most suitable one for the analyzed data. How exactly the strategy should look like and whether it is economically viable to use several storage devices in parallel will be the result of the next step in the ENSURE project. With the help of the ENSURE economic performance assessment engine (e.g., [15], [16]), several storage scenarios can be compared and their expected gains will be estimated and considered with respect to the given constraints for the data. Future work will also focus on the validation of the functional and non-functional requirements, use cases and scenarios identified, and their traceability in the implemented preservation system.

7. ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under the grant no. 270000. The authors acknowledge all consortium members who contributed to the ENSURE M4 and M12 Requirements Deliverables: D1.2.1a and D1.2.1b.

8. REFERENCES

[1] ENSURE: D1.2.1a, ENSURE Requirements Deliverable, M4, June 2011 (http://ensure-fp7.eu).

[2] ENSURE: D1.2.1b, ENSURE Requirements Deliverable, M12, January 2012 (http://ensure-fp7.eu).

[3] Borghoff, U.M.; Rodig, P.; J. Scheffczyk; Schmitz, L. Long-term Preservation of Digital Documents, Springer Berlin Heidelberg New York, 2003.

[4] Lee, K.-O; Stattery, O.; Lu, R.; Tang X.; McCrary, V. The State of the Art and Practice in Digital Preservation".

[3] Open-High-Low-Close.

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

Journal of Research of the National Institute of Standards and Technology, 107, 2002, pp. 93-106.

[5] CCSDS – The Consultative Committee for Space Data Systems, "Reference Model for an Open Archival Information System (OAIS)", Draft Recommended Standard, CCSDS 650.0-P-1.1, Pink Book, August 2009. (Accessed at: http://public.ccsds.org/sites/cwe/rids/Lists/CCSDS%206500P11/Attachments/650x0p11.pdf on February 20, 2012).

[6] Giaretta, D. Advanced Digital Preservation, Springer Verlag Berlin Heidelberg, 2011.

[7] CASPAR FP6 Project, D4101 User Requirements and Scenario Specification, 2006.

[8] KEEP FP7 Project, D2.2a Requirements and design documents for services and architecture of emulation framework, Part A, 2010.

[9] Edelstein, O., Factor, M., King, R., Risse, T., Salant, E. and Taylor, P., "Evolving Domains, problems and Solutions for Long Term Digital Preservation". Proceedings of iPRES 2011, November 1-4, 2011, Singapore, pp. 194-204.

[10] DiNetto E., Plebani P., "Describing case-studies: the S-Cube approach", 2010, S-Cube Networks of Excellence, http.//www.ss-cube-network.eu [Accessed at: http://www.s-cube-netowrk.eu/pdfs7EC-document-20100713.pdf on February 3, 2011.]

[11] Cockburn, A. "Writing effective use cases", Boston, MA: Addison-Wesley, 1999.

[12] Jackobson, M. "Software requirements & specifications: a lexicon of practice, principles and prejudices". New York, NY, USA: ACM Press: Addison-Wesley Publishing Co., 1995.

[13] Wiegers, K.E. "Software Requirements", Microsoft Press, 2nd Edition, 2003.

[14] Lilly, S. "Use case pitfalls: top 10 problems from real projects using use cases". TOOLS 30, Proceedings of Technology of Object-Oriented Languages and Systems, 1999, pp. 174-183.

[15] ENSURE: D20.1.1, ENSURE Activity II Scientific Report, February 2012.

[16] Chituc C.-M., Ristau P. "A service-oriented approach to assess the value of digital preservation". Submitted at the 2nd International Workshop on Performance Assessment and Auditing in Service Computing-PAASC2012, to be held at ICSOC 2012, November 12-16, 2012, Shanghai (*under review*).