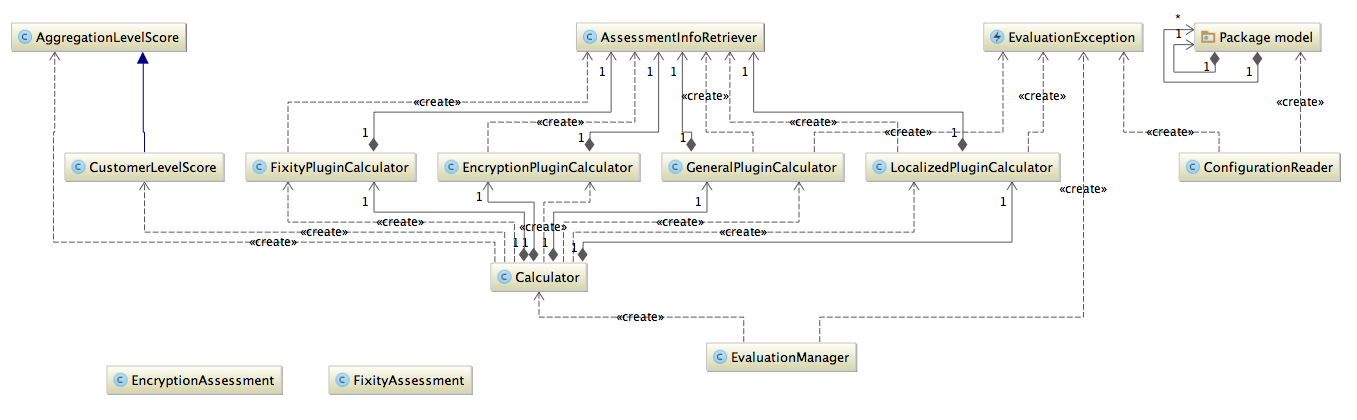
Quality model

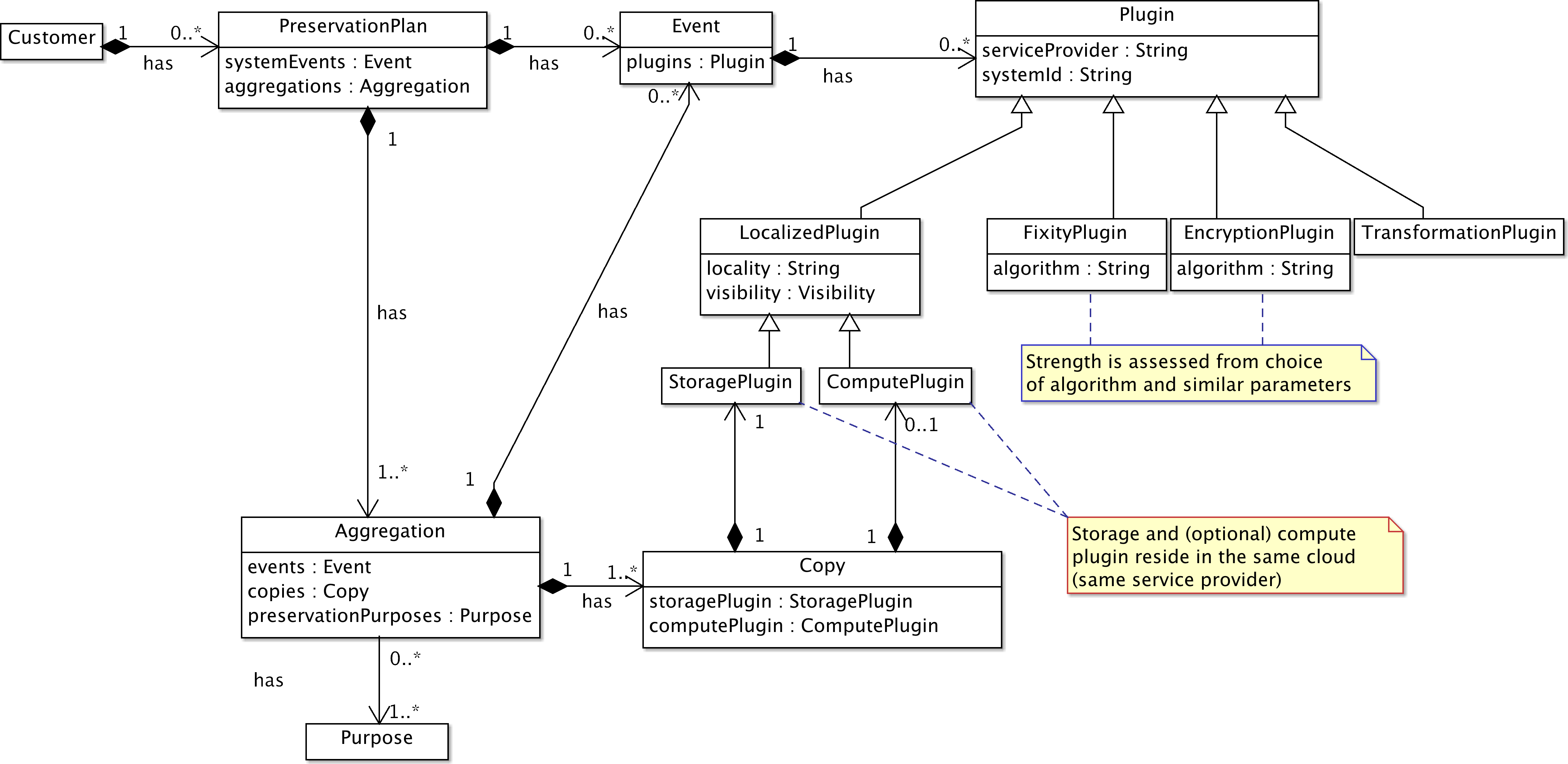
The Quality Engine receives preservation plan proposals from the Preservation Plan Optimizer (PPO). The Quality Engine then assesses the quality of the proposals in order to find the most suitable configurations from a quality perspective. The Quality Engine returns the estimated quality metrics with the preservation plan under consideration.

Internally, the Quality Engine is structured as per figure 4.11, which covers the online quality assessment classes. The model package, referred to in the internal structure model is detailed in figure 4.12.

Figure 4.12: Quality Engine Internal Structure

Upon receiving the preservation plan in the shape of a Global Preservation Plan (in xml format) from the PPO, this data is translated into an internal model used by the quality engine calculators, see figure 4.12.





4.8.1 Calculations

Figure 4.13: Quality Engine Preservation Plan Model

The data processing diagram in figure 4.13 shows how an estimation of quality is computed in detail by the Quality Engine; included in the diagram are the input and output parameters. An Aggregation corresponds to a class of digital objects, a collection, to be preserved. To handle a specific collection, a number of services are put to use. One task for the Quality Engine is to compute a quality score for each of these services (modeled as plugins).

One service (plugin) (in the figure numbered as 1) that is a subject for a quality assessment is the Ingest service which has the tasks of receiving control and creating an Archival Information Package suitable for preservation regarding content and structure. Another service that is useful in this context is the Fixity Service Plugin responsible for ensuring the integrity and authenticity of each digital object in the collection. A further example of a necessary preservation service is a Transformation service for migration from one format to another, since formats inevitably change in long-term preservation. The most fundamental and important service plugin that has to be assessed is the Storage service that takes care of the physical storage of objects. Another important service is the Access (Compute) service supporting the search and retrieval of digital objects from the Archive.

The basis for the quality score calculation of each service plugin (2) are service properties. These could be based on technical considerations such as ”Meantime between Failures”, ”Access Speed”, ”Loss Rate”, ”Storage Location” etc, and also on intersubjective considerations. We base the latter category of assess- ments on the Trustworthy Digital Preservation (TDR) standard, identified as measurable properties of a Service plugin. Examples of TDR metrics are organizational abilities such as ”Staff skill”, ”Staff turnover”, ”Use of community supported software”, ”Ability to handle anticipated volumes”, ”Mechanisms to detect bit corruption or loss”, ”Disaster recovery support”, ”Ability to keep track of changes”, ”Audit capacity”, ”Documentation of repository processes”, ”Written processing procedures”, ”Logs of content information types”, ”File format specifications”, ”Definitions of SIP completeness and correctness”, ”Listing all of the digital objects in the repository” and ”Metadata sufficient to validate the integrity” etc. Other formal and informal service quality frameworks can be utilised if they bring meaning to the customer.

The preservation purpose (purpose of use) (4) could be ”Business, Evidence, Research, Historical” and together with metrics originating from Service plugin properties or TDR and the importance of the metric could be boosted by a value from a list of weighting scores (5).

This means that specific properties can be treated as being more significant in the calculation of a Quality

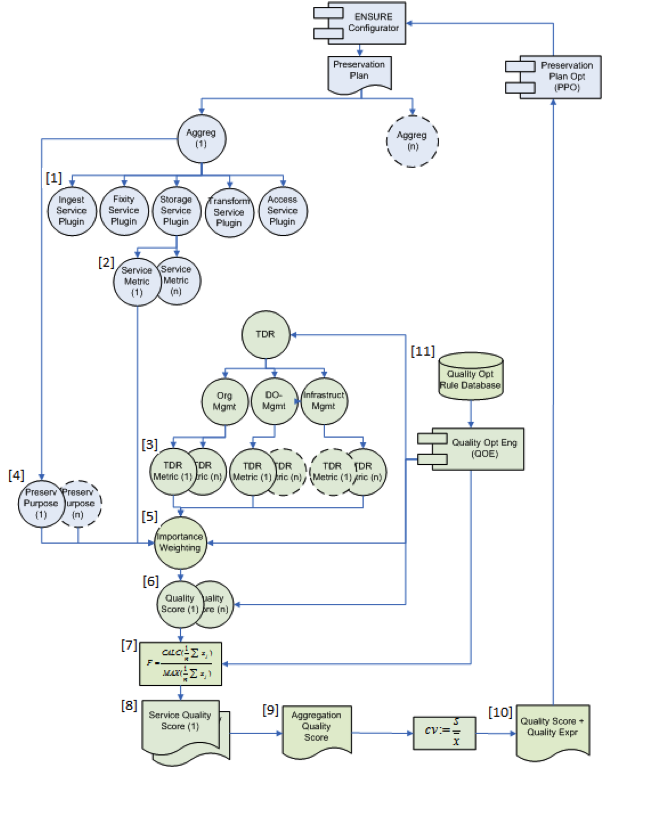


Figure 4.14: Quality Calculation Model

score (6) for each plugin. The calculation (7) for each service plugin;

CALC (􏰀 x) F = MAX(􏰀x)

where C ALC (􏰀 x) is the outcome of a measurement of a specific plugin. M AX (􏰀 x) corresponds to the maximum outcome that a plugin may have. By comparing these two we get the ratio (F), a value in the range [0, 1]. (8)

A calculated value for each plugin, calculated from (7), gives a value per aggregation. Item (9) Produces a calculated average over all aggregations for a given GPP. Item (10) c := s/x ̄ is the measure of variability calculated on the GPP level. It consists of the value of s that is the standard deviation of all aggregations divided by the average calculated from all aggregations. The quotient c gives a value with range [0, 1].

The Quality Engine keeps an internal database (11) containing opinions on services (plugins) which are used in the calculations. This database is maintained by the ENSURE service provider and managed through a separate web user interface. The database schema is depicted in figure 4.14.

Figure 4.15: Quality Engine Internal Database Schema

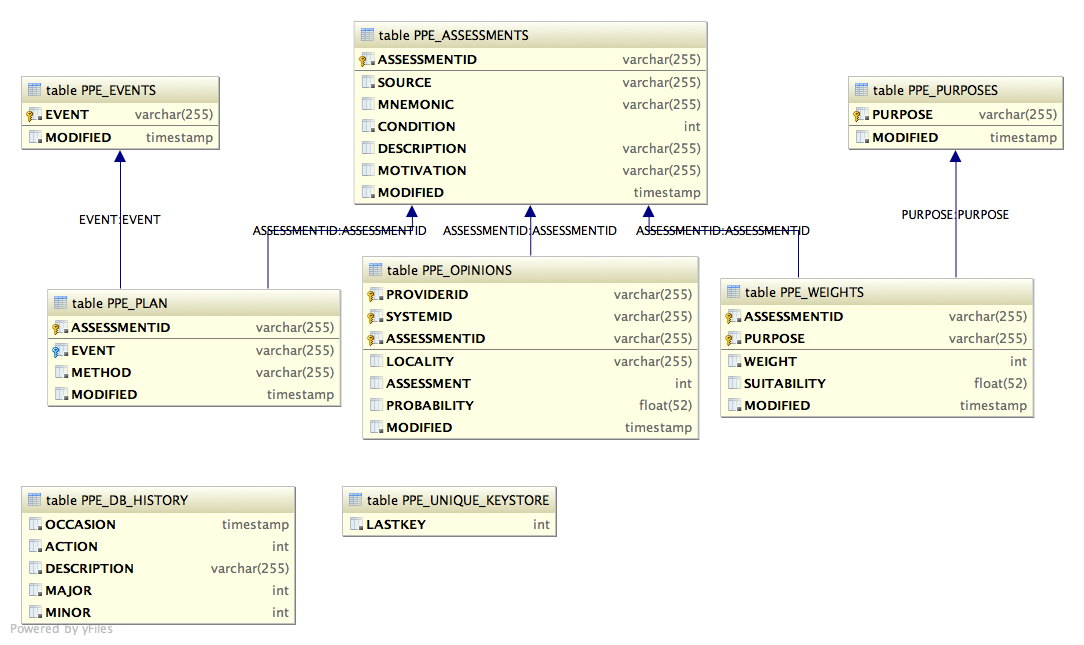
4.9 Graphic User Interface

The Configurator graphical user interface (GUI) was designed with the intention of hiding the complexity of designing a preservation solution, but still enabling the administrator of such a solution to make informed and effective choices.

The screens can be divided into input screens, where the administrator:

• Sets the parameters for the preservation solution both globally, and per aggregation: see Figure 4.15.

• Sets preferences for the various metrics that are produced by the preservation plan optimizer: see Figure 4.16.



Preservation Planner

3.1 Quality Model

This section describes how ENSURE models information systems quality, or more precisely *preservation systems quality* and how the Quality Engine (QE) uses this model in order to provide services to the ENSURE framework.

3.1.1 Overview

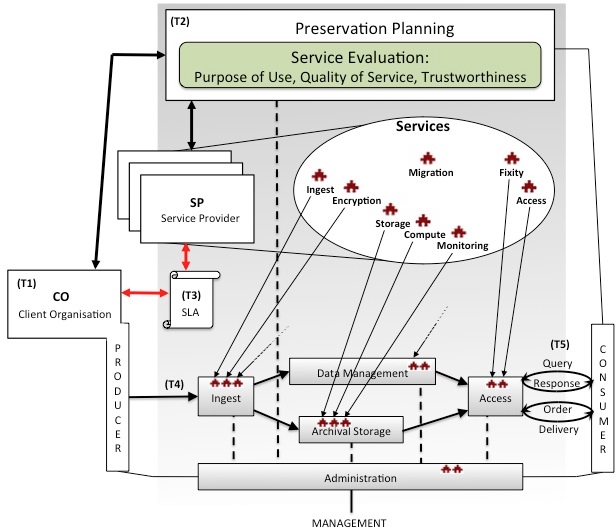


Figure 3.1: The Quality Engine Described In Terms of OAIS Functional Entities [9]

Many commercial-oriented organizations are facing serious problems of economically preserving a wide variety of data for future use, while ensuring protection of sensitive business data and privacy of individuals. There has been a rapid increase in suppliers that provide new service models such as cloud computing and it has become a realistic option for long-term digital preservation. With a variety of service providers – how do an organization choose to whom they will entrust their data for long-term preservation? It is not only a matter of sustainability neither of the service provider’s infrastructure nor of the issues surrounding the data itself, such as format, or even safety. Other issues have emerged as the requirements for client’s preservation of information change and the ENSURE project shows that including the purpose of preservation in the decision process makes sense.

The development of a framework for evaluating the quality of cloud computing services for long-term digital preservation is described in this section. Three main issues were identified during the development of a quality framework for commercial organizations:

1. What factors are important for describing the quality of a (preservation) service?

2. Is it possible to group or summarize system requirements based on the purposes of use of preserved information?

3. Is it possible to weigh the significance of quality factors to reflect different requirements for a preservation system relative to the purpose of use?

Three conceptual dimensions were identified as basis for the framework: *trustworthiness*, *quality of service*, and *purpose of use*. Engaging the financial and health care use case partners have been fundamental in capturing and understanding the different requirements related to quality aspects of preservation systems in these sectors. Figure 3.1 illustrates the role of the QE in the preservation system, figure 3.2 presents an overview of the quality assessment framework, and and figure 3.3 illustrates the flowchart of the quality assessment framework and its main components. The components are further explained in later sections.

A number of changes have been implemented during year three. This includes the following:

* Identification of quality key factors and improved algorithm for calculation of the importance of each quality metric based on Analytic Hierarchy Process (AHP) [35].
* A new Graphical User Interface (GUI) that supports interpreting the results of quality calculation based on quality key factors.
* A new GUI that supports entering of quality metrics and reconfiguration of its significance for the calculation of quality assessments.
* Development of a new syntax for expressing risks based on the GPP quality measurement results.
* Theuse(utility)ofthequalitykeyfactorsanditspresentationhasbeenvalidatedinENSUREuse-case  organizations.  3.1.2 Design

3.1.2.1 Solution Overview

Digital preservation can be understood as a series of managed activities necessary to ensure continued access to digital information over time and as long as necessary. This requires planning, resource allocation, and application of methods and techniques for preservation to ensure that digital information remains accessible and usable and retains its value.

In order to standardize practice of digital preservation and contribute to recommendations for constructing preservation system, the reference model for an Open Archival Information System (OAIS) [9] have been developed. OAIS describe the technical aspects of a digital object’s life cycle. Figure 3.2 illustrates the preservation system’s overall functional entities: *ingest*, *storage*, *data management*, *management*, *access* and *preservation planning*.

The advent of cloud computing offers new opportunities in the preservation-planning phase to select and configure (mix-and-match) a preservation solution consisting of services from different service providers (SP) in accordance with the client organisation needs. Among the choices of interest to the discussion is the possibility of choosing type of storage and compute services, choosing migration services, choosing fixity services to verify unintentional changes or bit rot, and choosing adequate encryption services to protect against non-authorised access and corruption of content.

The configuration of a client-tailored preservation solution is made in the preservation-planning phase by means of a decision support system. This section describes a framework to enable a decision based on quality of the preservation service.

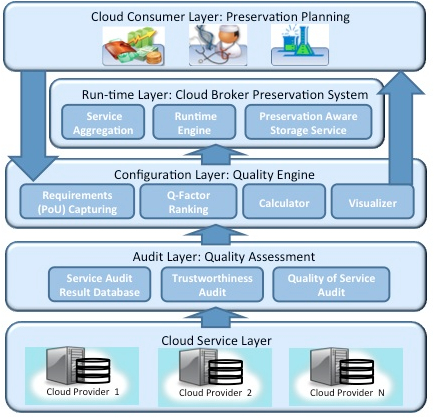


Figure 3.2: A Conceptual Overview of The Quality Assessment Framework 3.1.2.2 Case Study Findings: Identification of Preservation Systems Quality Requirements

A series of interviews were conducted with use-case owners to capture their requirements for a digital preservation service. The three organizations in the study represented three different commercial sectors: health-care (HC), medical clinical trials (CT) and finance (F). They are briefly described next.

HC was responsible for keeping medical records for a considerably long period of time. Part of the medical records need to be stored for seventy years and medical images need to be retained for a minimum of forty years. A large amount of the data that the organization needs to preserve is represented by textual data, images etc. They came in a range of formats, many of which were non-standard and proprietary. CT conducted clinical trials on newly developed drugs. Every trial collects a vast amount of personal data on each participant. These data has to be handled in such a way that it would not be at risk of being compromised, manipulated or mixed up. It was vital that the preserved data should be easily searched and accessed. These data were important for future research and traceable in case of litigation. They had to be stored for fifty years and allow secure sharing over time while protecting privacy rights. F was a capital management company that used high-volumes of data for predictive decision making models for the financial sector. Unlike HC and CT, F did not have to comply with regulatory requirements for preserving data. However, the data had business value as they are used to verify the predictive models in use against historical data. Preserving these data in a safe and affordable way was an important business issue. While the characteristics and the needs of the three organizations varied, they had one thing in common; while the value of their data holdings and their use of the information would change over time, the preservation objects had to be correctly managed at each point in time.

Finding 1 - Concerns related to the preserved data

One key challenge mentioned by the informants from all three organizations was that the preservation solution must handle the data in various formats. They used many non-standard formats, which were often proprietary, to represent a variety of preservation documents such as lab test results or X-rays. These could come as text, images, or raw binary data. Each of these types raised specific requirements. For example, in the HC case, preserving X-ray images was vital. The digital object must contain important patient informa- tion. It faced the choice of retaining its original format, which would result in large files on disk, or the high cost of compressing the data to another format which brought in the risk of loosing important data.

This issue also has different meanings in relation to the purpose of use and where a greater risk of loss of data is acceptable, if it according to current regulations has expired. For F, the challenge was that the data they used and needed to preserve originated from several different sources that used proprietary formats that are only supported by specific software for visualization and analysis. Since a large proportion of the data was preserved to meet regulatory requirements of the length of time it has to be preserved, the organizations also had to deal with the issue of what to do after the regulated time expired. Here responses varied; some saw the opportunity to erase some of the data or to compress it, as it was not subjected to requirements for fast access or high availability. They all however were unanimous on one count; the organization would define what change to the data is acceptable.

Finding 2 - Concerns about the service provider

The informants stressed that it was vital that they received quality service from the service provider. When we asked them to elaborate further on what they considered to be important aspects of quality, they mentioned several: availability, reliability, integrity, maintainability, privacy, performance and safety. Inter- estingly though, they had a variety of ways of evaluating service providers in terms of quality. They were largely based on reputation, financial position / stability and experience, low staff turnover and competence. A very important issue for respondents was the security on the external service provider was engaged to be responsible for their data and periodic status reports on the costs and if something happens or is done with the data.

Finding 3 - Purpose of preserving

When asked to give a reason for their organization’s rationale for keeping information over the long- term, the informants ranked legal reasons as most important, followed by economic, historic / memory and research purposes. An informant from CT gave a concrete example: in the case that some new issue surfaces around a drug that had undergone clinical trial at some point in time, they must be able to contact all the patients enrolled in the trials long after the trials were over. Another example was of probative value as external audits are performed regularly. Interestingly, the informants had difficulties imagining any alternative use of their data in the future other than for research and traceability for safety reasons, e.g. monitoring of a drug.

Finding 4 - Competency needed for preservation planning

Almost to a person, the informants emphasized that their main skill was not IT, and therefore they found it difficult to rank the benefits and risks of different services such as cloud-based solutions. They mentioned several risks such as unexpected changes in the data, security, privacy, confidentiality, and limited availabil- ity of connectivity problems. At the same time, they also mentioned expected benefits such as cost savings, error reduction with improved consistency and completeness, and availability (which they defined as access to data regardless of time and place.)

These organizations should also consider the extent to which safety and quality mechanisms are required in relation to cost, and the sensitivity and importance for their business if data be compromised or lost.

3.1.2.3 A Quality Assessment Framework (Preserv-QUAL)

Our proposed framework builds on existing frameworks augmented by the results of our discussions with the three client organizations described in section 3.1.2. This section presents a conceptualization of our framework in the form of the central parts of our framework; *trustworthiness*, *quality of service*, and *purpose of use*.

Trustworthiness: Trustworthiness (T) is a characteristic of a service provider and is defined as the extent to which a client organization perceives it safe to entrust its preserved objects to a service provider. The dig- ital curation and preservation communities conceptualize trustworthiness as a repository’s ability to meet certain criteria. If a repository adheres to specific criteria regarding organizational infrastructure, digital object management and technical infrastructure, including security it can be considered trustworthy. Trust- worthiness does not express opinions on the quality of the data to be preserved; it focuses on a repository’s ability to preserve digital materials. The factors that are part of T are:

* Authenticity (Integrity): capabilities that protect against data corruption or data loss, support of mechanisms to assert the digital record is a true copy of the original; track of actions and support of unbroken chain of custody mechanisms (auditability).
* Viability: organizational strength; skills/competence, financial; the risk of provider goes bankrupt.
* Security (Privacy, Confidentiality): This includes the assessment of hardware and software infras- tructure, and operational procedures to minimize risks. Includes mechanisms to avoid unintentional data breaches or corruptions (e.g., through loss or intrusions). Support of security audits, firewall, usage of encryption, and antivirus technology.  Quality of Service: Quality of Service (QoS) is a characteristic of a service provider and is defined as the level expected by a client organization of a service provider. QoS depends on the Client Organization’s needs. The factors that are part of QoS are:
  + Accessibility (Availability, Reliability): power failure risks, denial of service, speed of access, data transfer capacity, connection error rate, mean time between failure, scalability etc.
  + Portability (Flexibility): ability to move data from one SP to another; data lock-in risks; use of stan- dardized APIs. The infrastructure components of the repository should be able to export its holdings to a future custodian possibly another provider.  Purpose of Use: Purpose of Use (PoU) is the reason behind preserving the preservation objects. It is also synonymous to “purpose of preserving”. The use of information in an organization is subject to context, which determines the usefulness of the information in that context. As the context of information use in an organization may evolve over time, PoU cannot be fully determined at the time of ingestion. At any time, information may be put to concurrent use in different contexts, but it is generally expected that the requirements on multiple uses will diminish over time - as may the business value of the information. The boundary between the (objective) affordance offered by the preserved information and the (subjective) utility is defined by the purpose of use (context). The categories of PoU are:
    - Business: Timely access to information in the normal operation of business, where we stress the importance of having access to the right information at the right time. (Information must be timely, accessible, complete, accurate, relevant, and protected with authorized access)
    - Evidence: The information must meet legal requirements and must be acceptable as evidence in a court case, claims that differ between jurisdictions with free assessment of evidence and jurisdictions with admissibility considerations. An example could be the ability to provide provenance information

for preserved objects, support transfer of data if necessary and handled by a trusted organization.

(Information must be authentic, valid, transparent and portable)

• Historic: In a historical setting, information is often used for documenting “events” or “conducts”, e.g. documenting conduct and exertion of power in a government agency and also allows using for research purposes. This relates to keeping provenance information for the preserved object. (Infor- mation should be authentic, valid, available, handled by a viable provider)

3.1.2.4 The QE Framework in use: flowchart process overview for service (plugin) quality evaluation

The QE workflow in Figure 3.2 shows how the quality of the estimation process is divided into two major phases. One phase before the QE database that collects quality measures of the various services (plugins) considered as candidate services which may be components in a preservation solution. Each type of service has to be measured in different ways. The following describes the major process activities:

• Evaluate quality factors relevancy for purpose of use: this is an important activity assigned to estimate each quality factor’s relevancy for different purposes of use (business, evidence, historic). The estimation has to be made by experts from the preservation area using the method Analytical Hierarchy Process (AHP) [35] that supports pairwise comparisons of each quality factor in relation to each other for different purpose of use. The result of the AHP analysis is weighted values for each quality factor (authenticity, viability, security, accessibility, and portability) to reflect each quality factor’s relevance to different purposes of use. The result is saved in the QE database and later used as one basis for the calculation of quality executed by the calculate quality score activity as well as an explanation of the quality measurement exposed by the graphical chart presentation shown in figure 3.2.

• Evaluate storage and compute service: A cornerstone of a preservation solution is the storage and compute services (S/C). This activity relates to the measurement of the S/C in two different process activities. The first process relates to the measurement of quality factors related to the trustworthiness dimension of the QE-framework. One instrument suitable for this measurement is the Trustworthy Digital Repository Checklist (TDR). The checklist supports the measurement of service mechanisms that span from organizational staff competence and financial strength to technical infrastructure mech- anisms. The TDR-metrics are grouped according to defined quality factors of authenticity, viability, or security. Another type of S/C measurement is related to the Quality of Service dimension with re- lated quality factors of accessibility and portability. Metrics related to these factors are derived from cloud service benchmark tools and service specifications. Results can also be retrieved from public benchmark services. The result of the trustworthiness and quality of service measurement are saved in the QE database to be used as the basis for calculating the S/C plugin quality score.

• Fixity and encryption assessment: Fixity check is the process of verifying that a digital object has not been altered or damaged and is an important component of a digital preservation workflow. A fixity check is done by calculating and comparing the checksums or digests. A choice exists between different types of algorithms where a *checksum* is the simplest and least secure method, and the use of *cryptographic hash functions* is safer but more costly to compute. The difference lies in the complexity of the algorithm and a high cost of computation is in fact a necessary characteristic of a cryptographic hash function. Among the most common algorithms, we find MD5 (Message Digest 5) and various variants of SHA (Secure Hash Algorithm) [2]. Other things to consider in a preservation system are at what occasion fixity checks will be executed; upon receipt to the preservation system (ingest), upon dissemination from the preservation system (access), and upon regular intermediate intervals to detect “bit rot”.

Encryption is the process of encoding data so only authorized stakeholders are able to read it’s content. There are different classes of encryption methods, such as symmetric key cryptography and public key cryptography, each with their own merits.

Information about the choice of algorithms for fixity and encryption, together with some basic perfor- mance information, is part of the input to the QE. The QE also keeps records of individual merits of different algorithms and this information is used in providing a quality assessment.

• Assess transformation service: One of the main objectives of digital preservation is to prevent loss of access to information due to format obsolescence. Because of the complexity of many file formats, it may be difficult to determine if a file after migration from one format to another has retained all of its properties. Migration is a continual process, which must be repeated every time a medium reaches obsolescence. In work package 42 (WP42) we developed functionality that automatically validates information packages before, after and transitionally during a migration. It is also possible to use tools as Planets test bed for test of migration tools and file format suitability as input to QE.

• Calculate quality score: A quality score for each digital object (aggregation) and associated preser- vation services (plugins) is calculated by rules based on the results of various plugin measurements. Input of this process is a preservation plan that contains the specification of aggregation, events and plugins. Some preferences are also specified and part of the assessment. Output is quality score and risk expressions calculated and generated from the collected service (plugin) measurements and preservation plan specification.

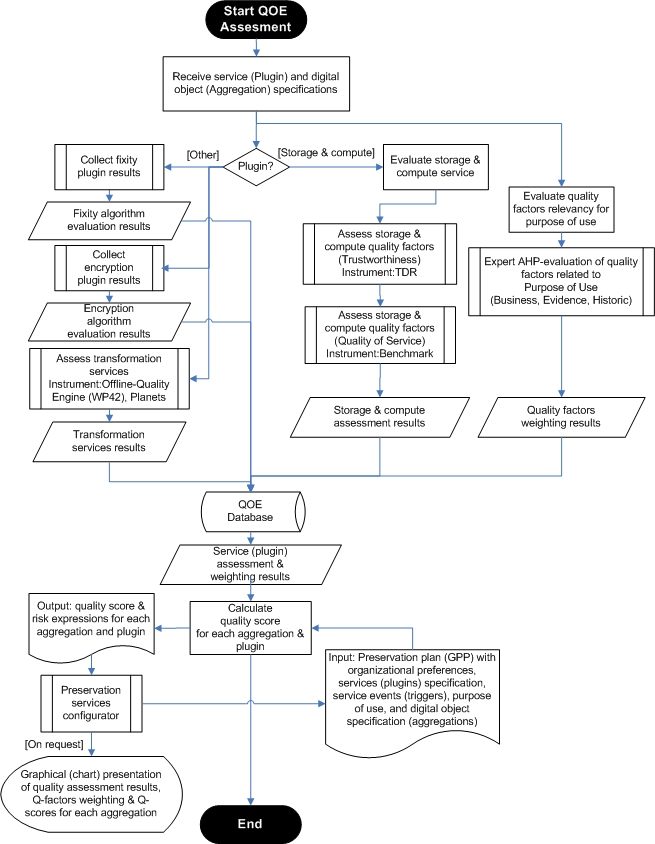


Figure 3.3: Process Flowchart for Preparation and Evaluation

3.1.2.5 Calculation of quality of fixity- and encryption plugins

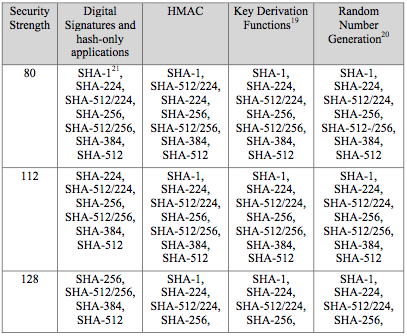
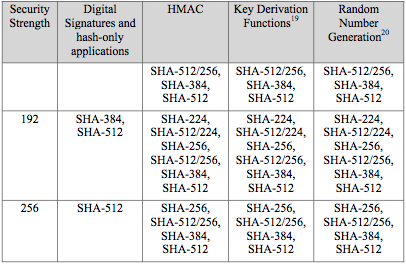
The calculation of quality for fixity and encryption plugins are treated separately from other plugin types since their qualities are primarily quantitative in nature. It is problematic to find a single authoritative source of strengths and weaknesses for all fixity and encryption algorithms, they are often analyzed separately, but we will refer to NIST (the US National Institute of Standards and Technology) for the sake of this discussion.

As presented in [2] and for the set of cryptographic hash algorithms governed by NIST, i.e. the SHA set of algorithms, the individual hash functions are given a strength – in this case as *bits of security* – as presented in figure 3.4 and figure 3.5.

Figure 3.4: List of strengths of (some) cryptographic hash algorithms, adapted from [2]

Figure 3.5: List of strengths of (some) cryptographic hash algorithms, adapted from [2] (continued)

Some algorithms are already deemed non-applicable; “*SHA-1 has been demonstrated to provide less than 80 bits of security for digital signatures; at the publication of this Recommendation, the security strength against collisions remains the subject of speculation. The use of SHA-1 is not recommended for the generation of digital signatures in new systems; new systems should use one of the larger hash functions.*

*For the present time, SHA-1 is included here for digital signatures to reflect its widespread use in existing systems, for which the reduced security strength may not be of great concern when only 80 bits of security are required.*” [2]

The level of strength of the different hash functions are then compared to a general assessment of perfor- mance of contemporary and future computing machinery, and from this discussion follows time intervals delineating the applicability of the different hash algorithms, see figure 3.6.

Figure 3.6: Delineating applicability of different cryptographic hash algorithms, adapted from [2]

NIST provides this description of the the table in figure 3.6 [2]: “*If the security life of information extends beyond one time period specified in the table into the next time period (the later time period), the algorithms and key sizes specified for the later time period shall be used for applying cryptographic protection (e.g., encryption). The following examples are provided to clarify the use of the table:*

1. *If information is cryptographically protected (e.g., digitally signed) in 2012, and the maximum- expected security life of that data is only one year, any of the approved digital-signature algorithms or key sizes that provide at least 80 bits of security strength may be used. However, if only 80 bits of protection is used, there is some risk that needs to be accepted. Note that a digital signature that provides 80 bits of security could be processed (i.e., verified) after 2013 as indicated by the legacy use indication in the table.*
2. *If the information is to be digitally signed in 2012, and the expected security life of the data is six years, then an algorithm or key size that provides at least 112 bits of security strength is required.*

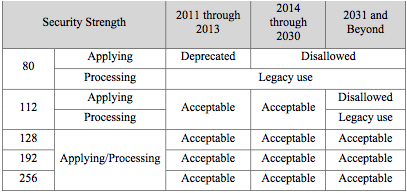
We have chosen to use the same reasoning around picking a cryptographic hash function for providing trust in the repository, in the meaning that the preserved information may be accepted as *authentic*. A fixity check calculation is very much the same as verifying a digital signature – with the key difference being that no verifiable encryption of the hash has been done.

The ramification is that, depending on retention period of the preserved information, the longer the in- formation is kept in the repository, the stronger the hash algorithm strength should be. Depending on the purpose of use of the preserved information, the trust expressed in the repository and in the preserved infor- mation may wary — if the purpose is *evidence*, we do require a suitable strength of the hash algorithm and will favour stronger algorithms before weaker which then affects the assessed quality of the fixity plugin in question.

The same reasoning applies to encryption.

3.1.2.6 AHP-based calculation of weighted Q-factors and quality score

Analytic Hierarchy Process (AHP) [35] is an efficient mathematical method for complex decision support. This is accomplished through a series of pairwise comparisons between different criteria. The outcome will be weighted values used in the calculation of quality. Figure 3.7 shows an example of the application of

D20.2.3 – February 19, 2014 44 of 78

ENSURE Consortium http://www.ensure-fp7.eu/ AHP to develop each quality factor’s importance for different purpose of use. The result is a weighted value

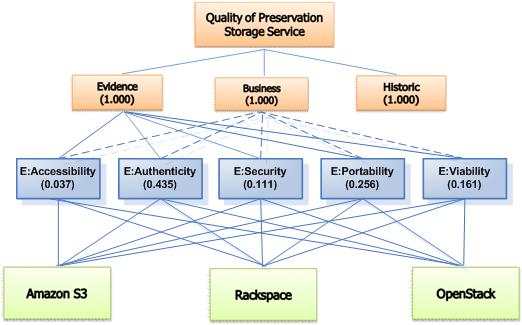
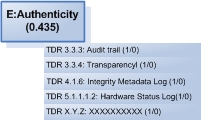
for each quality factor used in the calculation of the preservation plan quality score.

Figure 3.7: AHP evaluation of storage services based on the quality evaluation framework

Each quality factor has a number of quality metrics, which measure degree of quality factor fulfillment for each service. Figure 3.8 shows the quality factor authenticity and examples of metrics from TDR.

Figure 3.8: Arriving at the metrics

Using the scale 1-9, described in table 3.1, experts in the field can do a pairwise comparison of each quality factor relevancy in relation to each other for different purposes of use.

|  |  |  |
| --- | --- | --- |
| Intensity of importance | Definition | Explanation |
| 1 | Equal importance | Two elements contribute equally to the objec- tive |
| 3 | Moderate importance | Experience and judgment slightly favour one element over another |
| 5 | Strong importance | Experience and judgment strongly favour one element over another |
| 7 | Very strong importance | One element is favoured very strongly over another; it’s dominance is demonstrated in practice |
| 9 | Extreme importance | The evidence favouring one element over an- other is of the highest possible order of affir- mation |
| Intensities of 2, 4, 6, and 8 can be used to express intermediate values. Intensities 1.1, 1.2, 1.3, etc can be used for elements that are very close in importance | | |

Table 3.1: Fundamental scale for pairwise comparisons – adapted from [35]

For the sake of the following discussion; In the ENSURE project, we chose to use the following pairwise judgments for the *evidence* purpose, see table 3.2, but they are in no way magical and should be adjusted to match the domain in which the preservation service provider is active. As much as possible, we tried to adopt the expectations of the use case owners in this work – based on the interview and survey we did in Year 2 of the ENSURE project.

|  |  |  |  |
| --- | --- | --- | --- |
| Criteria A | Criteria B | More important criterion | Intensity of importance |
| Authenticity | Viability | A | 5 |
| Authenticity | Security | A | 3 |
| Authenticity | Accessibility | A | 7 |
| Authenticity | Portability | A | 2 |
| Viability | Security | A | 2 |
| Viability | Accessibility | A | 5 |
| Viability | Portability | B | 3 |
| Security | Accessibility | A | 4 |
| Security | Portability | B | 2 |
| Accessibility | Portability | B | 6 |

Table 3.2: Pairwise comparison of quality factors relevancy for *evidence* purpose of use

3.1.2.7 Pairwise Comparison of Quality Factors Relevancy for Purpose of Use

Based on the outcome of the AHP (above), for each purpose and depending on the number of Quality Factors under consideration (in this case 5), we establish a 5 × 5 matrix describing the correlation of individual quality factors, i.e. all the pairwise comparisons of quality factors for each purpose.

Let Ae denote a 5 × 5 matrix for the purpose *evidence*:

 a1,1

a1,2 a1,3 a1,4 a2,2 a2,3 a2,4 a3,2 a3,3 a3,4

a1,515372

a2,5  1/5 1 2 5 1/3

a2,1 Ae = a3,1

a3,5 = 1/3 1/2 1 4 1/2  

a4,1 a4,2 a4,3 a4,4 a4,5  1/7 1/5 1/4 1 1/5 a5,1 a5,2 a5,3 a5,4 a5,5 1/2 3 2 5 1

where, for instance, a1,2 corresponds to the pairwise comparison of *authenticity* against *viability* in light of the *evidence* purpose in table 3.2. The diagonal will be all ones (a1,1 = a2,2 = a3,3 = a4,4 = a5,5 = 1) and the lower left part of the matrix will be the inverse of the upper right part of the matrix.

In order to establish weighting values, i.e. quality factor relevancy for PoU, we square the correlation matrix Ae, which describe the *proportion* of variance in common between any two Q-factor. By squaring the correlation matrix Ae, we achieve comparable standard scores (weighting factors):

 5 189/10 87/4 61 257/30   x1,1 x1,2 x1,3 x1,4 x1,5 

 107/42 5 89/12 139/6 11/3   x2,1 x2,2 x2,3 x2,4 x2,5  Ae × Ae =  73/42 149/30 5 46/3 79/30  =  x3,1 x3,2 x3,3 x3,4 x3,5 

  239/420 103/56 121/70 5 737/840   x4,1 x4,2 x4,3 x4,4 x4,5 

163/42 21/2 51/4 73/2 5 x5,1 x5,2 We calculate a sum of the values in each row in the squared matrix

5

􏰁xi,j =Rei j=1

and calculate the sum of each of these row-sums

5

􏰁Rei =Se i=1

where we then arrive at a weighted value for all rows

1 · Re = We Se

which in the *evidence* purpose case corresponds to

 0.44 

 0.15  We =  0.11   0.04 

x5,3 x5,4 x5,5

0.26 This vector represents *evidence*-weighted values for each quality factor under consideration.

We use the Analytic Hierarchy Process (AHP) to calculate the weighted parameters for each purpose of use (evidence, historic, and business). These are then used to analyze the quality factors (Q-factors). The result is seen in table 3.3:

|  |  |  |  |
| --- | --- | --- | --- |
| (α) Authenticity | 0.44 (αe ) | 0.47 (αh ) | 0.27 (αb ) |
| (β) Viability | 0.15 (βe ) | 0.26 (βh ) | 0.08 (βb ) |
| (γ) Security | 0.11 (γe ) | 0.03 (γh ) | 0.27 (γb ) |
| (δ) Accessibility | 0.04 (δe ) | 0.13 (δh ) | 0.35 (δb ) |
| (ε) Portability | 0.26 (εe ) | 0.11 (εh ) | 0.03 (εb ) |
| Sum | 1.00 | 1.00 | 1.00 |
| Each column of weighted Q-factors for a given purpose shall sum to unity | | | |

Table 3.3: Weighted Q-factors

It is now possible to calculate a quality score for a plugin, in light of how well it contributes to preserving information for a specific purpose of use (of the preserved information). In order to calculate the quality score for a plugin used for preserving information that must retain *evidential* qualities, we use the following setup:

􏰀 αcalc 􏰀 βcalc QAe =αe·􏰀α +βe·􏰀β

max max

􏰀 γcalc 􏰀 δcalc +γe·􏰀γ +δe·􏰀δ

max max

􏰀 εcalc

+εe·􏰀ε max

where 􏰀 αcalc is the sum of fulfilled values for the quality factors of authenticity α and 􏰀 αmax is the maximally possible outcome for the quality factors of authenticity α, and likewise for the other quality factors in table 3.3.

Using the sets of constants in table 3.3; {αe,βe,γe,δe,εe} for purpose *evidence*, {αh,βh,γh,δh,εh} for purpose *historic*, and {αb, βb, γb, δb, εb} for purpose *business*, we can calculate the assessed quality of individual plugins for any of these purposes.

The values of the individual αcalc (authenticity), βcalc (viability), γcalc (security), and εcalc (portabil- ity), which are either true (1) or false (0), come from individual assessments based on TDR that has been partitioned according to the individual Q-factors listed in table 3.3.

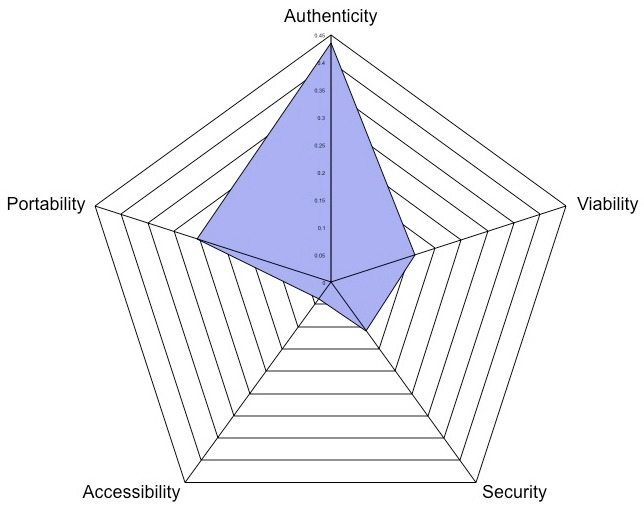
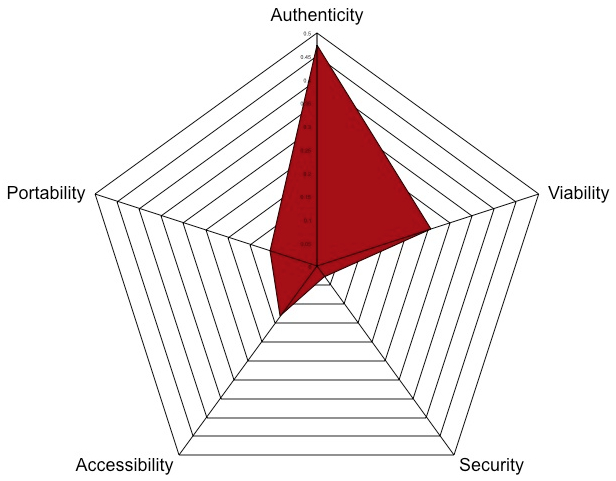
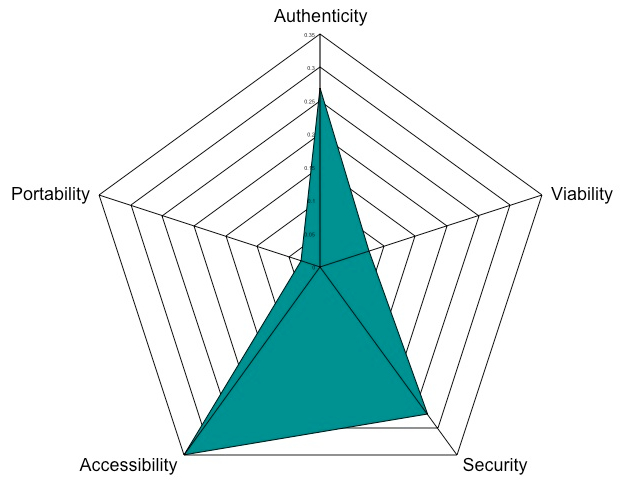
The values of the individual δcalc (accessibility), on the other hand, are natural number in the set {0, 1, 2, 3, 4, 5} — where the number 1 is minimum/worse, the number 5 is maximum/better, and where the number 0 indicates that no value exists for a given assessment. The individual assessments corresponds to object storage push/pull (mb/s), uplink/downlink throughput (mb/s) through measurements from members of Cloud Auditors, e.g. CloudHarmony.

Finally we arrive at an assessed quality score for individual aggregations by calculating the mean score of all plugins in the aggregation, combined with the variance of the individual plugin scores. These two values are provided to the Performance Optimization Engine for all aggregations in the global preservation plan.

Weight distribution for any given purpose, distributed in relation to the Q-factors are illustrated in the following polar diagrams — observe that the scales of the axes are a bit different among them, but they illustrate the differences in concern among the different purposes.

3.1.2.8 User-oriented GUI-presentation of QA-assessment results

This section shows a screenshot from the graphical presentation of quality calculation result. Figure 3.9 shows examples of the quality score of a service (plugin) – in this case Rackspace Cloud Services. The bar graph shows the degree of service fulfillment for each quality factor based on quality measurement of the plugin. The pie chart shows the distribution of the weighted values for each quality factor for the specified purpose of use. The weighting is based on expert judgment by means of the AHP-method.

Evidence Historic Business

Table 3.4: Differences in weight depending on purpose

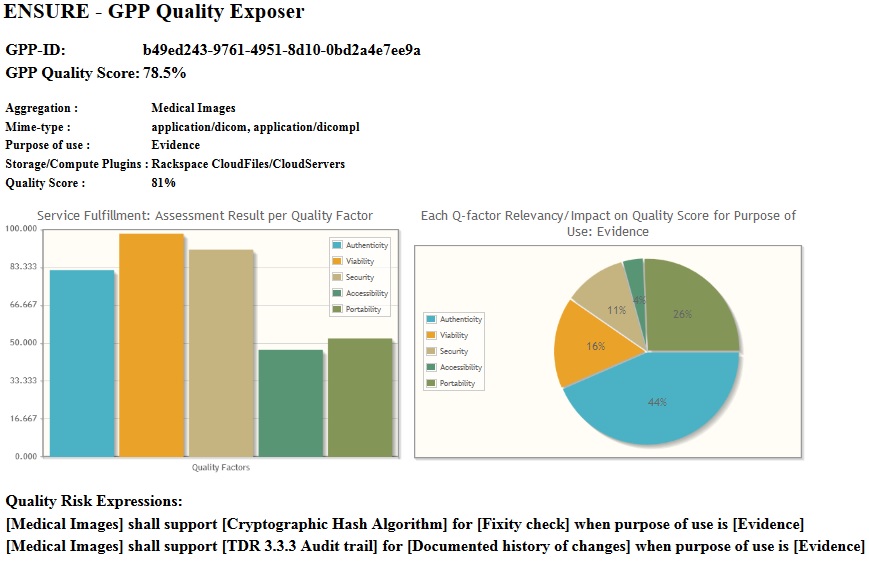


Figure 3.9: Graphical presentation of a quality assessment result



D20.2.3 – February 19, 2014 49 of 78

ENSURE Consortium http://www.ensure-fp7.eu/



Figure 3.9 shows syntax for expressing risks based on digital object, quality metrics for the purpose of use. The expressions that are exposed under the heading Quality Risk Expressions lists the specific metrics that do not achieve acceptable level. The GUI is designed to support preservation planners in the decision of the configuration of service-based preservation solution. The user interface exposes an explanation of quality scores’ meaning and exposes risks associated with the service configuration proposal (choice of plugins).

3.1.3 Validation

The technical validation of the Quality Engine consists of validating a central group property of the cal- culation. We show that the calculation is closed in the range 0 ≤ QE ≤ 1 for all quality assessments of individual plugins.

3.1.3.1 Closure

Since all individual quality assessments x of any plugin and x ∈ N0, we can validate closure for minimal and maximal cases.

Let all measurements of assessed quality be minimal, where each αi, βi, γi, δi, and εi are taken from a column in table 3.3. We then have:

00000

0 = αi · 􏰀 α + βi · 􏰀 β max

max

+ γi · 􏰀 γ + δi · 􏰀 δ + εi · 􏰀 ε max max max

Let all measurements of assessed quality be maximal, where each αi, βi, γi, δi, and εi are taken from a column in table 3.3. We then have:

max

since each column in table 3.3, corresponding to individual vectors We for the purposes we are studying, sums to 1 (an effect of how we arrived at We), we see that

αi +βi +γi +δi +εi =1

Thus the calculation is closed in 0 ≤ QE ≤ 1 for all measurements, which is the rationale for reporting quality calculations from the Quality Engine as a percentage of fulfillment towards a specified purpose of use of the preserved information.

3.1.3.2 Does it make sense?

Let us reason around different scenarios in order to ensure ourselves that the performance of the Quality Engine makes sense.

In table 3.5, we have chosen values for QE, where 0 ≤ QE ≤ 1, that corresponds to plugins having impairments of specific Q-factors in order to show how this affects the calculated quality for the individual plugins.

The resulting quality assessments for these plugins are shown in table 3.6.

If we read table 3.6 horizontally, we see that the quality assessment of the same plugin is influenced by the purpose of use of the information. If we read the table vertically, we see the quality assessments for a given purpose of use.

􏰀αmax 􏰀βmax 􏰀γmax 􏰀δmax 􏰀εmax

1 = αi · 􏰀 α + βi · 􏰀 β max

+ γi · 􏰀 γ + δi · 􏰀 δ + εi · 􏰀 ε max max max

D20.2.3 – February 19, 2014 50 of 78

ENSURE Consortium http://www.ensure-fp7.eu/



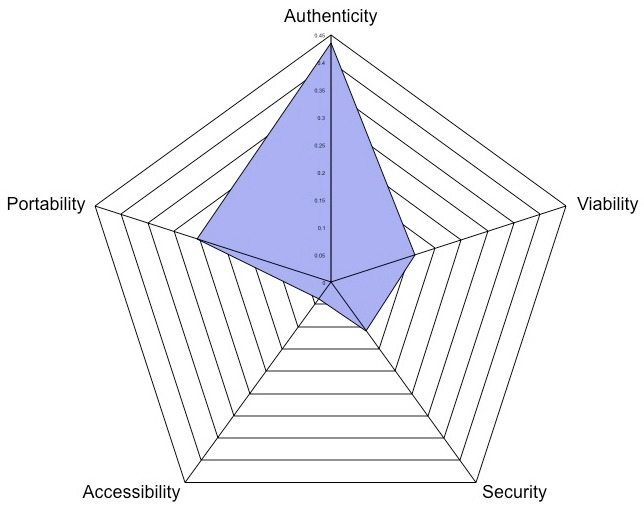
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Impaired authenticity | Impaired viability | Impaired security | Impaired accessibility | Impaired portability |
| 􏰀 αcalc 􏰀 αmax | 0.5 | 1 | 1 | 1 | 1 |
| 􏰀 βcalc 􏰀 βmax | 1 | 0.5 | 1 | 1 | 1 |
| 􏰀 γcalc 􏰀 γmax | 1 | 1 | 0.5 | 1 | 1 |
| 􏰀 δcalc 􏰀 δmax | 1 | 1 | 1 | 0.51 | 1 |
| 􏰀 εcalc 􏰀 εmax | 1 | 1 | 1 | 1 | 0.52 |

Table 3.5: Calculated quality values for plugins with specific impairments

Table 3.6: Assessed quality of impaired plugins weighted to purpose

Given the importance of the individual Q-factors for each purpose of use, the results depicted in table 3.6 is consistent with the intersubjective requirements stated in table 3.4, as outlined in table 3.7.

|  |  |  |  |
| --- | --- | --- | --- |
| *Plugin with ...* | Evidence | Historic | Business |
| Impaired authenticity | 78.25% | 76.23% | 86.58% |
| Impaired viability | 92.77% | 87.10% | 96.13% |
| Impaired security | 94.32% | 98.58% | 83.31% |
| Impaired accessibility | 98.19% | 93.58% | 83.04% |
| Impaired portability | 87.72% | 94.92% | 98.50% |



Evidence: The plugin with impaired authentic- ity scores the lowest among all (impaired) plugins from the *evidence* purpose point of view. This is an effect of scoring low on the most important Q-factor for this purpose of use.

The plugin with the impaired accessibility scored the highest, accordingly, since this Q-factor is least important from the *evidence* purpose point of view.

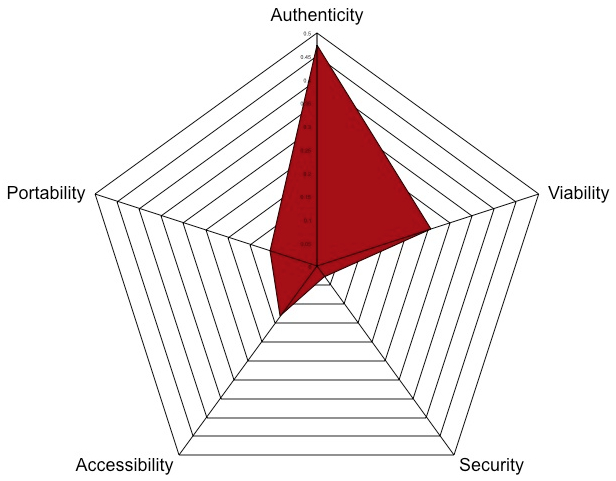
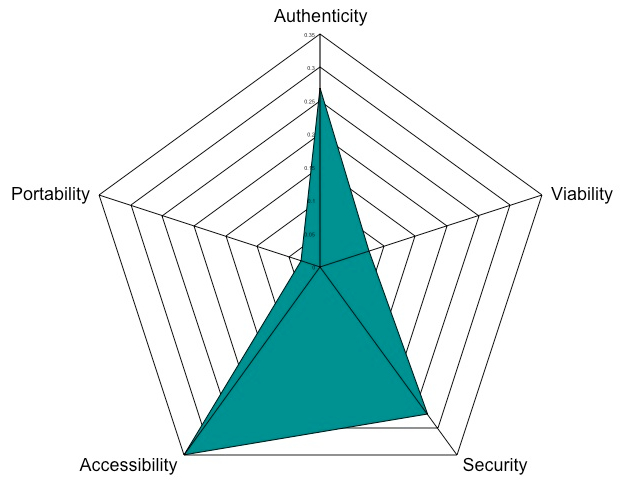
Historic: The plugin with impaired authenticity scores the lowest among all (impaired) plugins from the *historic* purpose point of view. This is an effect of scoring low on the most important Q-factor for this purpose of use.

The plugin with the impaired security scored the highest, accordingly, since this Q-factor is least im- portant from the *historic* purpose point of view.

Business: The plugin with impaired accessibility scores the lowest among all (impaired) plugins from the *business* purpose point of view. This is an effect of scoring low on the most important Q-factor for this purpose of use.

The plugin with the impaired portability scored the highest, accordingly, since this Q-factor is least im- portant from the *business* purpose point of view.

Table 3.7: Ranking of impaired plugins in light of purpose of use

This simple exercise, where we see relevant effects of viewing quality calculations of individual plugins in light of the purpose of use of the preserved information and where the intersubjective aspects of purpose of use could be derived from AHP in a logical manner, is sufficient for technical validation of the quality calculations in the Quality Engine.

3.1.4 Result

This work package lists these results:

* Developed a Quality Engine (QE) framework for estimating the quality of services in a cloud broker preservation LTPD system.
* Identified different quality factors (Q-factors) that describe the results of quality of service (QoS) measurements, which explains the result on an abstraction level appropriate for preservation planners.
* Identified purpose of use (PoU) as a dimension in the QE framework that group requirements based on client organizations’ (cloud consumer) preservation needs.
* Developed an algorithm based on the Analytic Hierarchy Process (AHP) that supports expert assess- ment of the Q-factors importance weighting related to PoU used in the calculation of a service quality score.
* Developed syntax for expressing risks associated with metrics that are not met as a (minimal) set of stated requirements for a specific PoU.
* A software component has been implemented that supports the execution of the QE framework for automatic calculation of service quality as part of the configuration engine.
* A validation of the QE framework algorithm for calculation of the service quality score has been conducted confirming its credibility. Experts in the field of digital preservation and cloud storage should do an estimation of Q-factors importance weighting in order to strengthen its credibility.
* Validationoftheuser’spartofthepreservationplanningprocessgraphicaluserinterface(GUI)should be done in larger scale in addition to the project’s business organizations to obtain results that increase the credibility of the QE framework.

Workpackage4.2

3.1 Overview

The purpose of this work package is to provide means to establish *opinions* on the quality of individual transformation plugins, to monitor the action of transformations on individual digital objects in the archive, to construct sample transformations to be used in preservation scenarios, and to monitor the behavior of these transformations over time.

During the assessment of a proposed preservation plan, the online quality engine depends on an auxiliary database of *opinions* on available plugins regarding quality features and has to be prepared ahead of production use. This work package provides a tool to assess some of the metrics associated with quality for a specific plugin.

Transformation tools are packaged as self-contained virtual appliances, but they may exhibit unwanted side effects. Unsuspected dependencies on previous transformations may trigger illicit behavior during subsequent transformations — prior transformations may have modified a digital object in such a way that it eventually falls outside the envelope of required behavior.

In order to detect deterioration of transformation tool chains, purpose-specific quality aspects of both AIPs as well as individual bit streams contained within these AIPs are measured and compared. This may be done based on relative comparison, such as comparing quality before and after individual transformations, or based on absolute quality measurements of samples as compared to demands stated in the purpose of the preservation. The former allows intervention while the latter may be used to detect existing problems in the collection.

Offline quality measurements are partly implemented by ENSURE-specific tools and partly by tools developed in earlier projects. The tools developed in ENSURE are based on the ENSURE quality model. We are not only interested in aspects of the contained bit stream itself (e.g. individual DICOM files in an AIP), but consider the AIP (or a part of an AIP) as a whole. In the case of DICOM files, the format of embedded images in the file as well as presence of individual key/value metadata-pairs in the file may influence the fitness for a specific purpose — and thus the quality of the digital object.

Depending on the purpose of preservation, we are looking at such diverse information as validating pack- age information components (fixity- or provenance-information), validating the representation information metadata, as well as looking into the individual bit streams (i.e. files).

Different approaches to defining data and information quality have been taken in the literature, many of which overlap or complement each other. We have chosen to express quality in terms of objective and inter- subjective parts. It is notoriously difficult to make use of subjective quality measurements in an automated preservation environment — traditionally being sampled by use of qualitative approaches and somehow fed back into the preservation system. By having the users of the preservation system document (parameterize) the purpose of the preservation *a ́ priori*, we study the possibilities of treating key subjective quality aspects as being technically objective aspects during the preservation process and what we can gain from this.

We also investigated specific transformations for DICOM Digital Pathology images for use in preservation scenarios. For DICOM images containing multi spectral image data, there is a need for specifying how to display the image data (either as a multi-spectral true color image, or as a pseudo color image) in a consistent and device independent manner. Additionally, for the application of molecular pathology, there is a need for describing such a transformation taking into account potential cross-talk between the multispectral channels, in order to arrive at the individual marker intensities (fluorescent or chromogenic).

Year 2 enhancements are:

• The engine responsible for unpacking and reading (Archival) Information Packages has been devel- oped further, adding to the robustness of the software. Also, the software will handle varied technologies for packaging information packages (such as TAR, ZIP, etc.) but also unpackaged information packages — applicable to running before the information packaging is done. These steps contributes to the usability of the software, even outside of the ENSURE project as the software is distributed as open source.

* The EQEL specification language has gotten more attention, but finalization is postponed till Year 3 until after having the online quality engine requirements finalized.
* Functionality was enhanced regarding the health case use case scenario, now also covering doctors’ encounters.
* A new multispectral presentation state is defined, with the following requirements:
  1. Ability to define how to display multi-spectral images as true color visible light images.
  2. Ability to define how to un-mix multispectral input channels for the purpose of deriving quan- titative representations of individual biomarker intensities, said markers can be fluorescent or chromogenic.
  3. Ability to define how to display (un-mixed) multi-spectral images (fluorescent, chromogenic) as pseudo color images. It should be possible to use the un-mixed output from 2) as input for this mode.
* Year 2 new technologies are:
* The online quality engine calculates quality metrics for proposed preservation plans. These plans specify a set of plugins for storage, computations, checks and various types of transformations. Re- garding transformation plugins, the online quality engine does calculations that partly are based on *opinions* that cover both subjective as well as objective qualities of the plugin. Functionality for pro- viding a basis for these *opinions* (on transformation plugins) has been added. This functionality is based on the analysis of the behavior of the transformation on an information package.
* As part of simulating aging, e.g. by introducing various types of errors into information packages (IPs), functionality was added that allows a sequence of; unpacking IP, modifying contents of IP, and repackaging IP. Since the offline quality tool already knows how to unpack IPs and trigger content- dependent functionality, adding repackaging functionality made sense to address the simulation of aging.
* New functionality was added for the financial use case.
* Implementation of DICOM multispectral presentation state  3.2 Design  Regarding the health care use case, we have specifically studied retrieval of original and ingestion of trans- formed DICOM objects. The objective of the transformation is to reduce the size of the stored object without losing relevant information from the diagnostic point of view. The transformation of objects is done off-line and is performed on a case level that may contain multiple images. Each image is represented as a DICOM object that has a DICOM structured report object associated with it. The DICOM structured report contains a description of the relevant Regions-Of-Interest (ROI). The ROIs are selected manually by a pathologist during diagnosis of an image and stored with metadata relevant for diagnosis e.g. whether it is a positive or negative sample (malignant or benign). The transformation will rely on the information stored in the DICOM structured report and use it to retrieve only the selected ROIs from an image. As a result a series of ROIs represented as a DICOM series will be stored in the preservation system.

In order to generate DICOM Pathology images, team members studied and implemented the DICOM supplement 145 (Whole Slide Microscopic Image IOD and SOP Classes) within Philips Image Management System. One of the key challenges for transformation of medical images that the team investigated was to define archiving strategies for transformation of DICOM supplement 145 images to maintain clinically valid information. As the result of that investigation, the concept of using DICOM structure report has been proposed as a container of clinically relevant information like for example information about manually selected regions used for reporting, or results of manual read or Image Analysis Applications.

Currently DICOM only allows for a limited use of ICC profiles for defining color transformations:

1. The ICC implementation in DICOM is limited to specifying an ICC input profile. Other profile types are not supported.
2. The ICC implementation in DICOM can only work on composite input images in the RGB color space — i.e. ICC input profiles can only be applied to composite 3 channel RGB images.  To provide the desired functionality one or multiple multi-spectral presentation states can be associated with each DICOM image. A flexible and straightforward way to specify the required multi-spectral transfor- mations in the presentation state is to make use of existing ICC functionality. The current ICC functionality provided in DICOM however is too limited to achieve this end. The ICC support will be extended on two fronts:

1.

2.

3.3

AllowtheuseofICCDeviceLinkprofiles.TheadvantageofusingDeviceLinkprofilesisthattheycan specify transformations from either *n* to *n* channels, or from *n* to 3 channels, and are not color space specific. This allows the chaining of multiple transformations in a transformation pipeline, simply by chaining ICC DeviceLink profiles.

Extend the ICC Input profile support already present in DICOM. The current limitations (the profile can only be applied to RGB color space composite images) have to be removed, to allow for mul- tiple input channels in an arbitrary color space. The ICC Input profile would always be the last (or only) component in the pipeline, since it transforms to Profile Connection Space (PCS), which can subsequently be transformed by the image consumer (viewing station, printer) to be displayed.

Design decisions

During Year 2 development, we saw the need to wait until the online quality engine functionality was added before finalizing the EQEL specification language. Additional requirements following from quality assessment of proposed preservation plans has to be stable before we dare encounter the unification of quality assessment requirements into a language. Therefore this functionality is postponed until Year 3.

3.4 Results

The main achievements during the second year cover:

* The handling of (Archival) Information Packages was generalized and strengthened.
* Support functionality for the online quality engine was added.
* Support functionality for the aging simulation functionality was added.
* A paper discussing quality assessment of medical information in automated preservation systems was published and presented at the eChallenges 2012 conference in Lisbon, Portugal [?].

• The DICOM multispectral presentation state proposal has been accepted by the WG26, and currently a multispectral presentation state addition to the DICOM standard is being written