REALIZATION OF A RESEARCH AND DEVELOPMENT PROJECT (PRE-COMMERCIAL PROCUREMENT) ON "CLOUD FOR EUROPE"

TECHNICAL SPECIFICATION: SECURE LEGISLATION-AWARE STORAGE SOLUTION

ANNEX IV (C)
TO THE CONTRACT NOTICE

TENDER NUMBER <5843932>
CUP <C58I13000210006>

CLOUD FOR EUROPE

FP7-610650

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 610650.

EXECUTIVE SUMMARY

Note: This is an informative summary of the document. The actual specification relevant for the bids is in the remainder of the document.

Data privacy, protection and exercising control over data in the cloud is among the major objectives taken into account during development and deployment of cloud storage services. Strong encryption, secure authentication and authorization represent technical means to exercising control over the data stored in the cloud and to avoid unauthorized persons access the stored data.

Archiving and storage systems are safe for long periods of time and for different types of data. Current cloud storage services are not up to date with regard to the legal requirements of citizens and public administrations.

This lot and the associated challenges propose an archiving platform and legislation-aware cloud storage solution that meets governments and citizens requirements on security of stored data.

This document characterizes the general and functional requirements for a secure legislation aware storage solution.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 SECURE LEGISLATION-AWARE STORAGE

According to Table 1 of *Annex IV(A): Cloud for Europe: Challenges and General Requirements* this lot covers the challenges as listed in Table 1 below. An overview of all defined challenges and a short informative description is given in Annex IV(A). A detailed description and specification of the lot is given in the remainder of this document.

Table 1 - Secure legislation-aware storage lot challenges

Phase	No.	Challenge Summary	Lot	Award Criteria
Design, Procurement	3	Assessing the legislation compliance of a cloud storage service to its contractual and functional description	2: SLAS	S1, S2
Transition, Delivery	5	Digital archiving and preservation of data	2: SLAS	S1, S2, S4
Operation	6	Legislation-aware data provision and access	2: SLAS	I2, I3
	11	Seamless change of service provider	1,2,3 (all)	Q1, I2, I3

1.1 INTRODUCTION

Note: This is an informative introduction to the topic. The actual specification relevant for the bids is in the remainder of the document.

Data protection and exercising control over data are major goals in deploying critical services to the cloud. Citizens' and public administrations' requirements are similar in a sense that data shall be protected against unwanted access by third parties. Public authorities' access policies to their data are determined by organisational structures or by legal constraints. It is widely recognized that the lack of full EU harmonization of data protection rules is perceived as a recurring legal barrier.

Strong and secure encryption is, apart from secure authentication, a technical mean to exercise such control. For data location outside a public administration's legislation, encryption under the administration control could be a mean to avoid seizing data by foreign authorities.

Archiving systems provide secure long-time storage for different kind of data, help to manage the lifecycle of information and enable organisations to cope with the exponential growth of information.

Storing data and documents in clouds is problematic due to the varying national (and sometimes regional) legal requirements on the protection of information. These legal requirements also depend on the type of content to be stored (personal data, sensitive data, health data, etc.).

The storage services currently offered in the cloud market are usually not explicitly targeting legal requirements of public administrations. That imposes additional difficulties to public procurers, since they have to assess case by case, whether a given storage service offering conforms to the applicable legal framework. Moreover, the consumer and the provider of a cloud storage service could be established in different countries.

To give an illustrative example: Let us take into account the National Agency for Cadastre from country A which needs to buy a cloud service in order to store a certain amount of cadastral data. Within this context, data has to be kept for defined (often long) timespans in a secure manner to fulfil legal or internal compliance regulations. Due to the sensitive nature of information, the data first needs to be encrypted, archived and finally stored in a secure storage system.

It is required for the Encryption, Archive and Storage Service to be compliant with the subset of laws of the legislation framework of country A that applies to cadastral data.

The National Agency for Cadastre needs to know whether the selected services satisfy the legal requirements or not and needs to easily select offerings that are compliant.

The set of Cloud for Europe challenges described in the present Annex ask for solutions to develop a cloud enabled encryption, archive and storage platform which complies with the requirements of EU public administrations, citizens and supports their collaboration.

1.2 BUSINESS CONTEXT

Storing data is a well-established use case for cloud services given the fact that storage services are included in the portfolio of most cloud providers. Providing security mechanisms for ensuring the integrity of information and providing services for the long-term legislation-aware accessibility of information is at the time being only addressed by a minority of service providers.

Currently encryption in the cloud is a niche where either the user uses the cloud as a storage and performs encryption and key management locally, or where the cloud provider claims to store all data encrypted, but the user does not have any control over the encryption keys.

Encryption solutions currently seen e.g. in SaaS are that the cloud provider owns the keys. A more secure multitenant SaaS service needs to use encryption keys owned by the tenant organization, not by the SaaS service.

Generic archiving services, targeting the specific technical and regulatory requirements in the public sector – providing a secure data space for eGovernment applications – are at the moment only provided within the environment of public service data centers for their own applications, or in the best cases at the national level.

The compliance of storage services to regulatory frameworks is rarely incorporated in contracts and has to be managed case by case. Commercial cloud service platforms are currently unable to manage legislative compliance in a structured, systematic and dynamic way.

Currently encryption, archive and storage services are basically separate services in most cloud systems.

Currently no cloud based encryption, archive and storage service exists that complies with the requirements of public organisations regarding confidentiality, IPR, information integrity, efficient archiving and secure storage. This service might introduce new technical challenges.

From a legal point of view this service seems highly challenging:

- Establish trust by contractual means;
- The service processes data of different classifications: highly sensitive information, citizen-related information, or even confidential data. The data needs to be protected and only particular individuals or groups of individuals have access to specific files;

- Tackle with cross-border law enforcements;
- All stored data may fall under regulations concerning archiving and disclosure. Even the access to backups needs to be restricted to granted individuals;
- Retain data for long-term periods in a secure manner.

1.3 LOT-SPECIFIC REQUIREMENTS

Note: Annex IV(a) contains a general requirement section which applies to all three lots. Bidders are expected to also respond to the common general requirements.

The following is a list of definitions required in the following listing of the general requirements (these definitions below are in addition to the ones in the overall tender Glossary, as it is definitions specific to this lot):

- *"Data objects"* are stored data that can be associated with a "data subject"[1], a data owner and metadata and are analogous to files within a file system.
- *"Container objects"* are the fundamental grouping of stored data and are analogous to directories within a file system;
- *"Domain objects"* represent the concept of administrative ownership of stored data within a storage system;
- The "*data consumer*" is an organization or person that is accessing stored data in order to process it;
- The "*data producer*" is an organization or person that is publishing/saving data in the data consumer storage space;
- The "*data controller*" is an organization or person that has ownership of the stored data. It shall mean the natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data; where the purposes and means of processing are determined by national or Community laws or regulations, the controller or the specific criteria for his nomination may be designated by national or Community law[2];
- The "*legal rules*" are constraints that are applicable to the stored data;

[1] *"Data subject": an identified or identifiable natural person; an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity; (See Directive 95/46/EC article 2, lett. a).*

[2] Directive 95/46/EC article 2, lett. d)

- A cloud service is "*legislation-aware*" in regard to a specific legal rule if the process which is realising the service is executed in a legislation aware environment where it is constrained by legal requirements associated with that particular legal rule. The legislation-awareness affects the service behaviour.
- "Data subject": an identified or identifiable natural person; an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity (See Directive 95/46/EC article 2, lett. a).

1.3.1 ***S2LR1*** *– DIGITAL ARCHIVING AND PRESERVATION*

Table 2 - Digital archiving and preservation challenge

ID	**S2LR1**
Actors	Cloud consumer, cloud provider, data consumer
Description	According to SNIA[3], a digital archive is a storage repository or service used to secure, retain, and protect digital information and data for periods of time less than that of long-term data retention, while the preservation addresses the archiving of data in the long-term. According to the definition in ISO 14721 standard, Long Term is intended to be long enough to be concerned with the impacts of changing technologies, including support for new media and data formats, or with a changing user community. Long Term may extend indefinitely. Digital preservation in particular raises two key issues: the fragility of digital media (its 'shelf life' compared with, say, non-acidic paper is extremely short) and the rate at which computer hardware and software become obsolete. [1]. We are looking for a solution for long term preservation of data that could be implemented in new cloud storage services or as an improvement of existing cloud storage services. The solution will have to address the following: • the data object has to be easily discovered, identified and associated with other data objects; • the preservation of data must be guaranteed in case the data formats and associated metadata or the hardware and software platform has become obsolete; • the integrity of data objects must be guaranteed in time and upon data

[3] http://www.snia.org/education/dictionary/d

	replication and migration; • data formats should be monitored so that the cloud consumer is aware of risks of becoming obsolete and has time to mitigate the issue; • saving space and reducing costs is a key issue in archiving, thus specialized data compression solution should be provided; • strong long-term encryption is a must; • the system must support Litigation Hold and Permanent Deletion; • the system should allow data consumers to authenticate using pan-European eID solution such as STORK2.0[4] and FutureID[5]; • compliance with different European and other applicable regulations: MoReq 2, SOX, etc.) within the field of archiving. The solution comprehends a Cloud Management Interface that should expose the cloud storage functionality to the cloud consumer. The interface should support functions such as: • hierarchizing and managing of the data, container and domain objects; • compatibility with the major hypervisors and cloud orchestrators; • allow metadata to be associated with domains, containers and the objects they contain; • allow clients to discover the capabilities available in the cloud storage offering; • managing the delegation to external user authorization systems (LDAP, Active Directory, STORK2.0, etc.); • multi-layer logging and audit trailing.
Constraints	Though data compression and de-duplication are usually implemented and used at hardware level, the solution should come up with recommendations enabling the cloud provider to be able to identify the hardware that brings most value for the cloud archiving and preservation service.

1.3.2 ***S2LR2*** – *LEGISLATION-COMPLIANCE OF A CLOUD STORAGE SERVICE*

Table 3 - Legislation-compliance of a cloud storage service - challenge

ID	**S2LR2**
Full name	**Assessing the compliance of the contractual and functional description of a cloud storage service to a legislation framework**

[4] https://www.eid-stork2.eu/
[5] http://www.futureid.eu/

Actors	Cloud provider, cloud consumer
Description	Public administrations need to assess the compliance of storage cloud services to applicable laws and regulations. The assessment can be performed during the procurement process or during the operation of the service. Cloud Service Providers will benefit of a solution to easy assess their storage cloud services with respect to legislation frameworks of different countries and / or specific domain of applications (health records, justice…) regarding, by example but not exclusively data privacy, data protection, data security, data location. We are looking for solutions to low cost assessment of the adequacy and compliance of a given cloud storage service to a given target legislation framework and, hopefully, the effectiveness of implementation. The implementation of the solutions can be a composition of any kind of tools, software, process definitions or human contributions that realize an effective and efficient solution to the challenge. The conceptual input to the solution is a representation of a specific customer need, including an appropriate representation of the set of laws and eventually internal regulations regarding data management (protection, privacy, security…). The conceptual input to the solution includes an appropriate description of the type and characteristics of data / documents to be stored and processed. In some situations, the type of data is implicitly or explicitly described in the set of laws and internal regulations mentioned above. The conceptual input to the solution includes an appropriate description of the cloud storage service and of the contract ruling the relationship between the provider and the consumer. The conceptual output of the solution is a list of assertions about the compliance of the audited service to the input set of laws and regulations with an associated level of confidence.
Constraints	The auditing solution is not a *per se* "certification" system. Certification needs a certification authority. But the solution may be used by cloud consumers to self-assess the conformance to legal rules.

1.3.3 ***S2LR3*** *– LEGISLATION-AWARE DATA PROVISION AND ACCESS*

Table 4 - Legislation-aware data provision and access - challenge

ID	**S2LR3**
Full name	**Legislation aware management of authorization services for the provisioning and access to personal data in the cloud.**
Actors	Cloud provider, cloud consumer
Description	Public administration (in the role of data controllers) use storage cloud services provided by cloud providers (in the role of data processors) on behalf of citizens and business entities (data subjects). Public administrations need to manage the rights of the data subjects in a heterogeneous legal environment (European, national, local laws; sectorial or administrative rules). To give some examples, such legal constraints can be: ▪ data access restrictions limiting access to the data controller, data subject and their representatives; ▪ data access permissions to government entities for special purposes (like access to health data for urgent first aid purposes) Public administrations will benefit of a solution for the assisted, semi-automatic or automatic, in any case low cost, management of dynamic data authorization. The solution must be deployable as a cloud service. Cloud Service Providers will benefit of the solution so that they can easily provide storage services that assure a proper and lawful access to data. The implementation of the solutions can be a composition of any kind of tools, software, process definitions and human contributions that realize an effective and efficient solution to the challenge. The conceptual input to the solution is an appropriate representation of a set of laws or internal regulations regarding data management (protection, privacy, security…) that is formally equivalent to a description of rules and constraints to data access. The solution is able to enforce the defined policy in an authorization system. The input of the authorization system includes a proper electronic identification of a citizen or of a legal entity asking access to a specific data resource (a document, a folder, a record, a database).

	The conceptual input of the authorization system includes an identification of a specific data resource. The conceptual output of the authorization system is an authorization value.

Constraints	

1.4 USE CASE SCENARIOS

Note: This informative section provides an incomplete list of some exemplary use case scenarios for the *Secure legislation-aware storage solution*. The aimed solutions by the bidders should be uniquely valid and should not specifically target one of the use case scenarios.

1.4.1 ***S2UC1*** *– NATIONAL CADASTRE AGENCY*

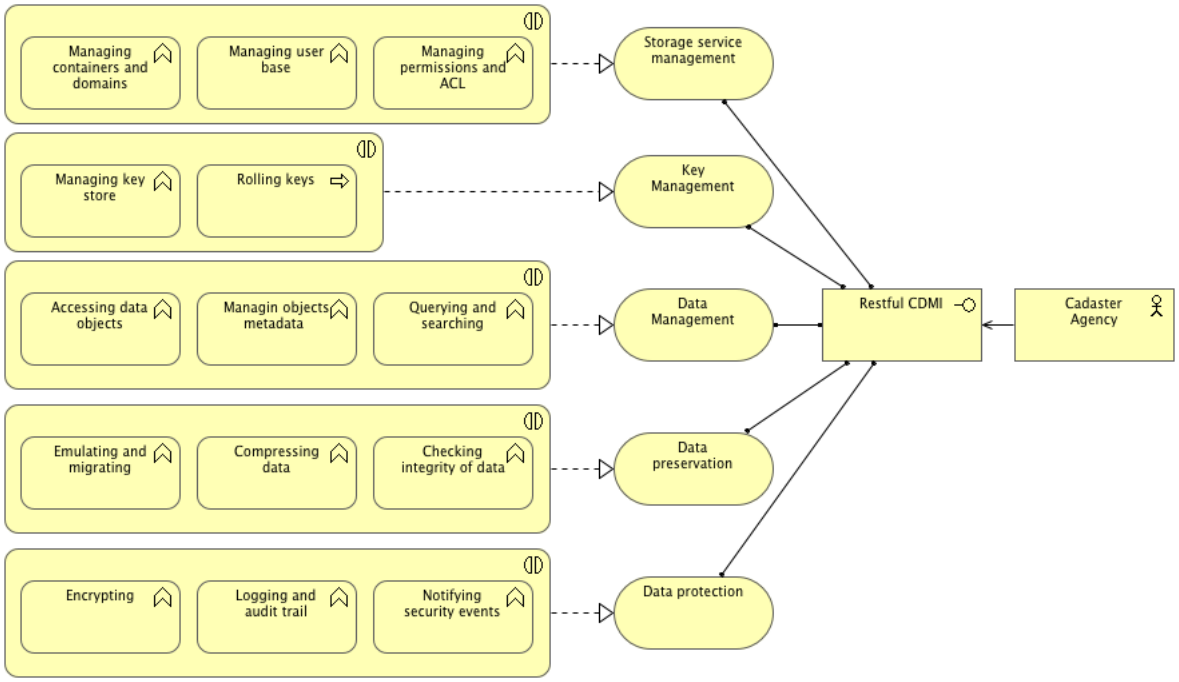
Table 5 - National Cadastre Agency use case scenario

ID	**S2UC1**
Actors	Cloud consumer (agency), data consumer (agency's users), cloud provider
Description	As a service, the cadastre holds information about the value and ownership of the land as a basis for taxation. The cadastre has to retain this information for decades, if not centuries to come. This is expensive and challenging even if the archive is kept in electronic format as in time the files might get corrupted or accidently deleted. Encryption is also an issue in the data preservation. To cope with this, the National Cadastre Agency decides to use a system based on the solutions to the challenges of Secure Legislation-Aware Storage cloud service. The National Cadastre Agency could have selected the storage service using the solution to the challenge "legislation compliance of a cloud storage service". The agency will use a **RESTful cloud management interface** to interact with the

	cloud service. Workflow and the main services the agency will be using: • the agency will start with defining **container objects**, each container corresponding to an administrative region; • the agency will also define a separate **domain** where it will store cadastre information which is not actively used and which is to be managed differently than the active information; • the agency will use an existing base of users or data consumers that need access to the archived data and associate these users with particular containers. The authentication of users is done using an **external authentication scheme** such as STORK2 or FutureID. • Thus, users from region A have access only to the container associated with region A. There are users that have global access and that can read, update or delete the archived objects. The access policy can be defined and enforced using the solution to challenge "Legislation-aware Data Provision and Access " • using the **metadata** service, the legitimate agency users will iteratively describe and store the archived files to appropriate containers. The metadata is also used by the users to **query and find** archived files; • the agency will add a set of **encryption** keys in the key store and schedule a **key rolling** every N months; • the agency can run jobs such as the **compression of** data objects and containers to save storage or **checking the integrity** of the data stored; • the agency is **notified** when a data object is deleted or found to be corrupted; • the agency can run **audit trails**. The agency can thus monitor who accessed a specific data object, what operations were executed, on what, by who and when; • the agency can **recreate the full functionality and exact look and feel** of the data objects that have been retained for very long periods of times, even if the formats of the data objects become obsolete; the agency can take **snapshots** of the data and can revert the data to a specific snapshot.
Constraints	

The following figure gives an overview of this exemplary use case.

Figure 1 - Services and underlying functionality needed by the Cadastre Agency



1.4.2 ***S2UC2*** – *PERSONAL DIGITAL ASSET*

Table 6 - Personal Digital Asset use case scenario

ID	**S2UC2**
Actors	Cloud consumer (agency), data consumer (agency's users), cloud provider
Description	Consider the case of a public school using a "digital box" service for each of its students (data subjects) where it put **data and documents** regarding the student's activity, such as grades, evaluations, education materials, etc. The school could have selected the storage service using the solution to the challenge "legislation compliance of a cloud storage service". The student can then authenticate and access data on his/her personal digital asset. The school also grants access to student's parents or tutors. The authentication of the students and her/his parents is done using an **external authentication scheme** such as STORK2 or FutureID. Following the authentication of the user, **another system** uses some access control rules to decide whether access requests from the authenticated user shall be granted or rejected. The cloud service is **legislation-aware** of the fact that when the student reaches

	the age of 18 she/he can decide for her/himself to whom he/she may want to grant access. Thus the service will automatically revoke the permissions of the parents or tutors and the student will have to grant them access explicitly, as their representative or delegates. Let's take in consideration the situation in which the parents of the student divorce, and for some severe reason, a judgment of a court prohibits a parent the access to the personal data of his son. The court will communicate this decision to the system to update the access control rules with this exception. Let's take in consideration the situation in which the students change school during the school year. The right to update the documentation in the digital box will switch from the old school to the receiving school. The personal digital asset use case can be generalized. There are many scenarios in which data access rights by subjects, their representatives and government entities must be accurately managed.
Constraints	

The following figures illustrate this exemplary use case.

Figure 2 - Personal Digital Service use case

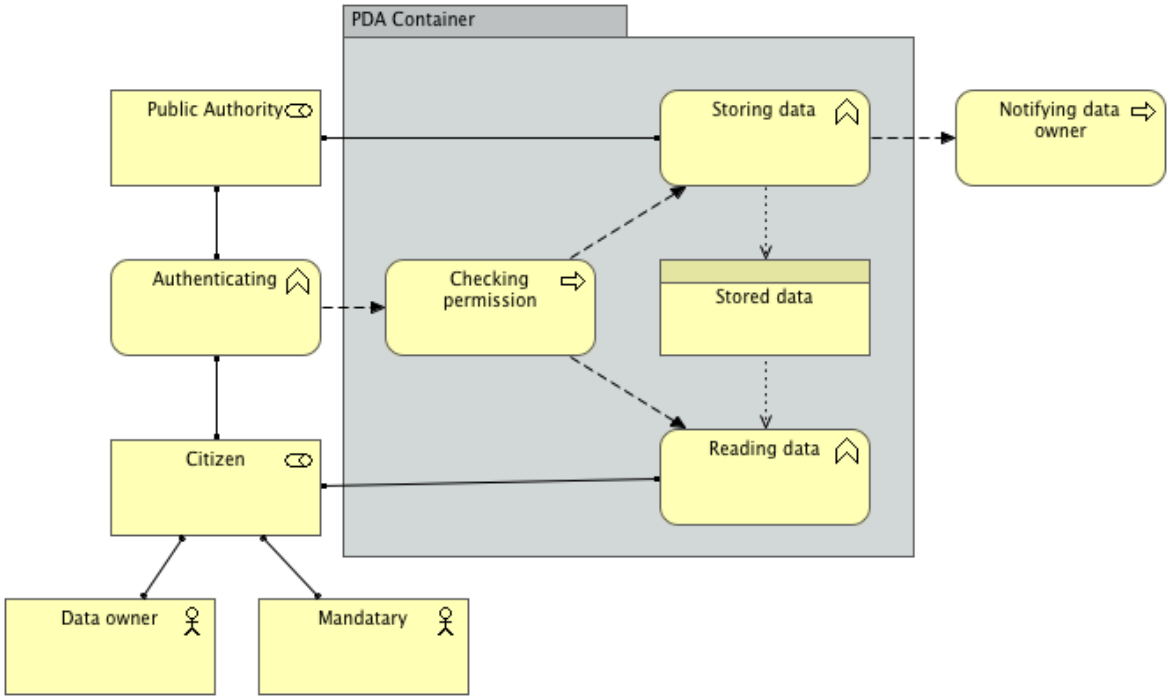
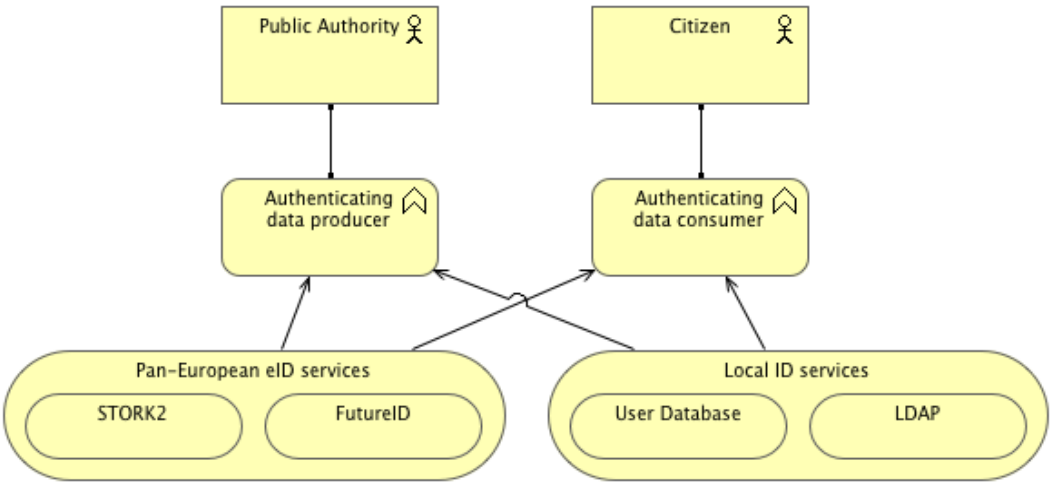


Figure 3 - Authentication schemes for data producer and data consumer



1.5 FUNCTIONAL REQUIREMENTS

Table 7 - Challenge "Legislation-compliance of a cloud storage service" – Specific Requirements

ID	Actors	Importance	Description	Award Criteria
S2FR1		High	A solution to assist the auditing of the cloud storage service in order to assess the adequacy and compliance with the target legislation framework MUST be provided.	C1
S2FR2		High	The description of the solution SHOULD include the list of tools, software, process definitions, data resources, knowledge resources, human contributions and their proper composition in order to solve the challenge	C1 C4
S2FR3		Medium	The solution preferably relies on existing and well established or standard cloud services and data/knowledge sources.	C1 C4
S2FR4		High	The solution MUST be deployable as a cloud service.	C1

S2FR5		Medium	The solution SHOULD be designed to scale up to about 100.000 customers (government entities) coming from the Member State, tens of cloud storage providers and hundreds of different cloud services	C3
S2FR6		Medium	The input to the solution SHOULD include the representation of the customer requirements regarding the cloud storage service, including at least: - an appropriate representation of a set of laws and eventually internal regulations regarding data management (protection, privacy, security, location…); - an appropriate representation of the type of data to be stored and processed, including as example, the presence of sensitive or personal data, the specific domain (justice, health …). In some situations, the type of data is implicitly or explicitly described in the set of laws and internal regulations mentioned above.	C1
S2FR7		High	The input to the solution SHOULD include the representation of the cloud storage service to be audited, comprehending at least: - the technical and functional specification of the service; - the clauses of the agreement ruling the relationship between the provider and the consumer.	C1
S2FR8		Medium	The solution SHOULD include: - a list and a description of the formats/syntax accepted to describe the input; - a description of capabilities and limits of the expressive capacity of the formats/syntax accepted as input.	C1
S2FR9		Medium	The formats / syntax of the input should be able to represent as many as possible laws and regulations regarding data management (protection, privacy, security, location…) in the Member countries of the European Union.	C3

S2FR10		High	The formats / syntax of the input should be able to represent as many as possible of the cloud storage services currently offered on the market. The solution SHOULD easily be adapted in order to audit emerging cloud storage services on the market	C3
S2FR11		High	The conceptual output of the solution is a list of logical assertions about the compliance of the audited service to the set of laws and internal regulations given in input. Each assertion MUST be associated with a level of confidence, expressing, in some way, the probability that the assertion is coherent with the evaluation performed by human experts. The solution SHOULD describe how the level of confidence is estimated. *Figure 4 sketches the conceptual form of the output of the solution.*	C1
S2FR12		Medium	The performance of the solution MUST be evaluable using the Precision indicator, i.e. the number of output true positive assertions divided by the total number of output assertions.	C1
S2FR13		Medium	The solution SHOULD hopefully have a "learning" behaviour, so that its precision monotonically increases in time with the usage.	C1,C3, C4
S2FR14		Low	The performance of the solution MUST be evaluable using the Time to Response indicator, i.e. the time elapsed between the submission of the last input to the delivery of the last corresponding output.	C1
S2FR15		Medium	The solution SHOULD comprehend methods and tools for the benchmarking (objective measurement) of the Precision index.	C1
S2FR16		Medium	The solution SHOULD support at least the English language and another official language of the European Union	C3
S2FR17		Medium	The solution SHOULD comprehend methods and tools for the localization to other official languages of the European Union.	C3

S2FR18		High	The solution SHOULD comprehend an estimation of its operational costs, including, as example: - an estimation of the cost of an industrial implementation of the solution - an estimation of the cost of usage (by example for transaction)	C4, I4
S2FR19		Medium	The solution MUST provide human-accessible and inter-system accessible managements interfaces.	Q8
S2FR20		Low	The solution SHOULD comprehend a risk impact assessment in order to determine the damage level for Confidentiality, Integrity and Availability breaches.	S1
S2FR21		Medium	The solution could collect and store persistent data. In this case the solution MUST be compliant with EU data protection standards and legislation regarding security and data protection.	Q4

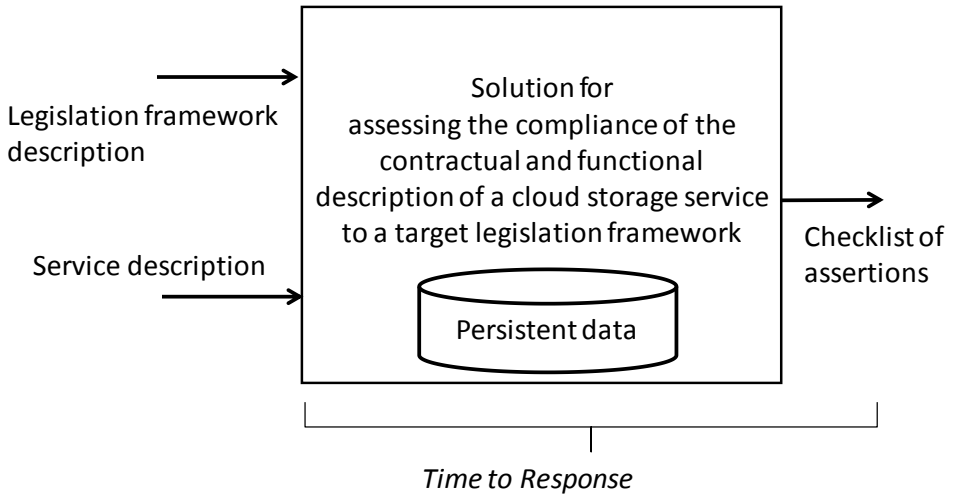


Figure 4 - Conceptual Input-output description of the solution

Figure 4 sketches a conceptual input-output model of the solution. The solution could comprehend a persistent storage of data collected from other sources or produced by the solution itself.

Audited Service: MegaStorage 1.2 by BestCloudprovider.uk VS ItalianLaw set		Confidence
The service is compliant with the Italian technical rules on digital preservation	V	95%
The service is ISO 27001 certified	V	100%
The service is compliant with article x of the Italian law xx/2005	X	98%
The service is compliant with security level three of the internal regulation for the archiving of administrative documents	?	77%

Figure 5 - Conceptual example of the output of the solution

Figure 5 sketches a conceptual example of the output of the solution in the form of a "checklist". The detailed definition of the format of the output should be part of the proposed solutions.

Table 8 - Challenge "Legislation-aware data provision and access" – Specific Requirements

ID	Actors	Importance	Description	Award Criteria
S2FR22		High	A solution for the assisted, semi-automatic or automatic, in any case low cost, management of dynamic data authorization MUST be provided.	C1
S2FR23		High	The description of the solution SHOULD include the list of tools, software, process definitions, data resources, knowledge resources, human contributions and their proper composition in order to solve the challenge.	C1 C4
S2FR24		Medium	The solution preferably relies on existing and well established or standard cloud services and data/knowledge sources.	C1 C4
S2FR25		High	The solution MUST be deployable as a cloud service.	C1

S2FR26		Medium	The solution SHOULD be designed to scale up to support authorizations requests regarding: - millions of citizens and companies from different Member State; - hundreds of cloud storage services; -	C3
S2FR27		Medium	The input to the solution SHOULD include at least: - an appropriate representation of a set of laws and eventually internal regulations defining the data access policy to be enforced; - an appropriate representation of specific events that have an influence on the enforcement of the data access policy; - electronic identification of a citizens or a legal entity; - electronic identification of data objects.	C1
S2FR28		Medium	The solution SHOULD include: - a list and a description of the formats/syntax accepted to describe the input; - a description of capabilities and limits of the expressive capacity of the formats/syntax accepted as input.	C1, C4
S2FR29		Medium	The formats / syntax of the input should be able to represent as many as possible laws and regulations regarding data access the Member countries of the European Union.	C3
S2FR30		High	The conceptual output of the solution is the authorization to a given (legal entity) subject to access a given data object.	C1
S2FR31		Low	The performance of the solution MUST be evaluable using the Time to Response indicator, i.e. the time elapsed between the submission of an authorization request and the delivery of the corresponding output	C1
S2FR32		Medium	The solution SHOULD support at least the English language and another official language of the European Union	C3
S2FR33		Medium	The solution SHOULD comprehend methods and tools for the localization to other official languages of the European Union.	C3

S2FR34		High	The solution SHOULD comprehend an estimation of its operational costs, including, as example: - an estimation of the cost of an industrial implementation of the solution; - an estimation of the cost of usage (by example for transaction).	C4, I4
S2FR35		Medium	The solution MUST provide human-accessible and inter-system accessible managements interfaces.	Q8
S2FR36		High	The solution SHOULD comprehend a risk impact assessment in order to determine the damage level for Confidentiality, Integrity and Availability breaches.	S1
S2FR37		Medium	The solution should be applicable to as many as possible authorization systems already on the market (LDAP local authorization system, other standard or well known cloud Building Blocks)	C3

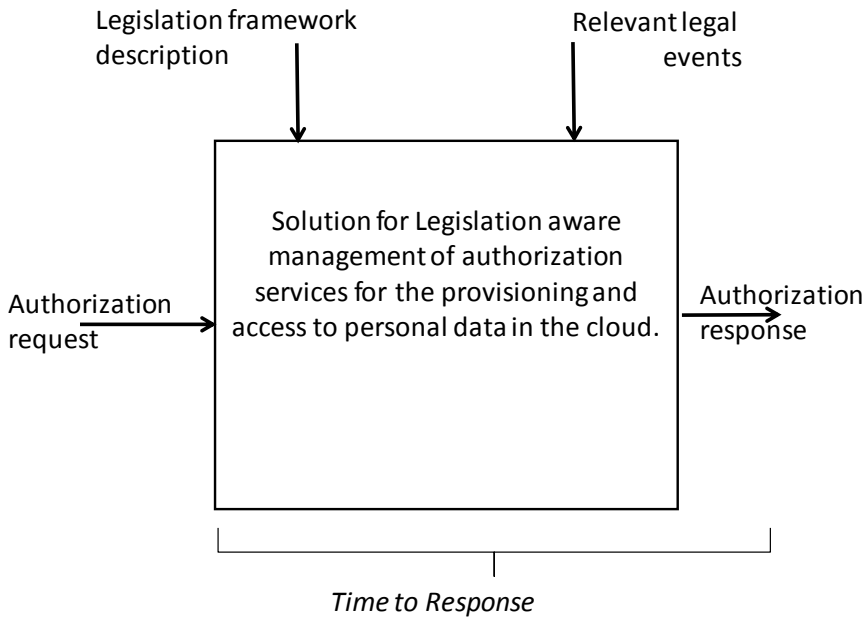


Figure 6 - Conceptual Input-output description of the solution

Table 9 - Challenge "Digital Archiving and Preservation" – Specific Requirements

S2FR38		High	A solution for long term preservation of data to be implemented in new cloud storage services or as an improvement of existing cloud storage services MUST be provided.	C1
S2FR39		Medium	The solution preferably relies on existing and well established or standard solutions, cloud services and data/knowledge sources.	C1, C4
S2FR40	Cloud Consumer, Cloud Broker, Cloud Provider	Medium	Data objects SHOULD be associated with metadata to help consumers to identify and process the data. Further data objects SHOULD be group able into containers. The containers again can be associated with each other.	C1
S2FR41	Cloud Consumer, Cloud Broker, Cloud Provider	High	The cloud consumer and the cloud controller SHOULD be able to use domain objects to separate the data owners and SHOULD be able to define data owners and associate the data owners with the domains.	C1
S2FR42	Cloud Consumer, Cloud Broker, Cloud Provider	High	The system MUST support CRUD (create, read, update, delete) operations and queries (including regular expressions) on data objects, containers and domains.	C1
S2FR43	Cloud Consumer, Cloud Broker, Cloud Provider	High	The system MUST provide a solution for data snapshots and data replication.	C1
S2FR44	Cloud Consumer, Cloud Broker, Cloud Provider	High	The cloud consumer MUST be able to define legal rules and constraints in a structured way and associate these rules to data objects, container objects and domain objects. The system MUST check against the legal rules upon accessing and processing of the data.	C1

S2FR45	Cloud Consumer, Cloud Broker, Cloud Provider	Medium	The cloud consumer SHOULD be able to define and associate access restrictions with the data objects, containers and domains.	C1
S2FR46	Cloud Consumer, Cloud Broker, Cloud Provider	High	The system MUST provide a plug-able concept for external authentication schemes (e.g. STORK2.0)	Q1
S2FR47	Cloud Consumer	Medium	Cloud consumers MUST be allowed to use different algorithms for data compression.	Q1
S2FR48	Cloud Consumer	Medium	All data MUST be encrypted and the cloud consumer has sole control over the encryption keys and key management. Further the system MUST provide a way for rolling the encryption keys. The encryption algorithms SHOULD be plugged in through a flexible system, so they can be extended easily.	C1
S2FR49	Cloud Consumer	Medium	The cloud consumer SHOULD be able to query data objects even if they are encrypted.	C1
S2FR50	Cloud Consumer	High	The cloud consumer MUST have access to all records regarding accessing of data objects and security and system events.	C1
S2FR51	Cloud Broker, Cloud Provider	Medium	The system MUST provide mechanisms for audit trailing.	Q7
S2FR52	Cloud Provider	High	The system MUST retain the data for long periods of time (long term retention) and MUST ensure data integrity and data recovery.[6] The solution should include a description of which types of changes (technological, at example) could affect the validity of the solution.	Q8

[6] Long Term is intended to be long enough to be concerned with the impacts of changing technologies, including support for new media and data formats, or with a changing user community. Long Term may extend indefinitely

S2FR53	Cloud Provider	High	Deleted data objects MUST be retained for a limited period of time in case of restoration	C1
S2FR54	Cloud Provider	High	The system MUST allow permanent deletion of data	C1
S2FR55	Cloud Provider	High	The system MUST not change or delete data that has been put on litigation hold.	C1
S2FR56	Cloud Provider	High	The system MUST automatically send notifications upon reaching legal rules, data deletion and data corruption or illegal access to data	C1
S2FR57	Cloud Provider	Medium	The solution should be compliant with most of the hypervisors on the market	Q1
S2FR58	Cloud provider	High	The solution must comprehend a Cloud Management Interface that should expose the cloud storage functionality to the cloud consumer. The interface should support functions such as: • hierarchizing and managing of the data, container and domain objects; • compatibility with the major hypervisors and cloud orchestrators; • allow metadata to be associated with domains, containers and the objects they contain; • allow clients to discover the capabilities available in the cloud storage offering; • managing the delegation to external user authorization systems (LDAP, Active Directory, STORK2.0, etc.); • multi-layer logging and audit trailing.	C1

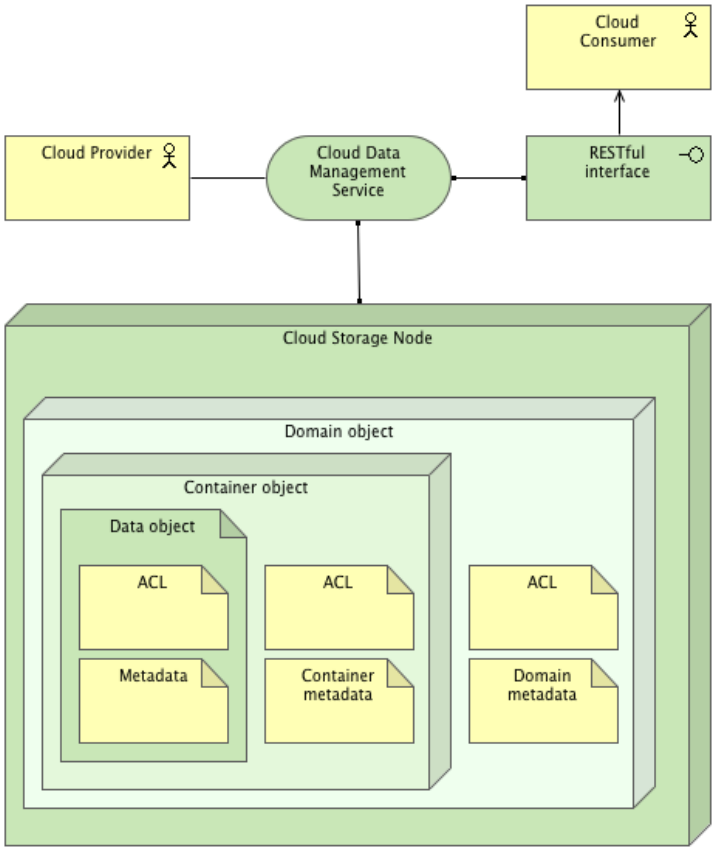


Figure 7 – Cloud Data management service

2 REFERENCES

[1] Granger, S. (n.d.). Digital Preservation & Emulation: from theory to practice.

[2] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data.

[3] COMMISSION DECISION of 5 February 2010 "**on standard contractual clauses for the transfer of personal data to processors established in third countries under Directive 95/46/EC of the European Parliament and of the Council**" (notified under document C(2010) 593) (Text with EEA relevance)(**2010/87/EU**)

[4] ISO 14721:2012 Space data and information transfer systems -- Open archival information system (OAIS) -- Reference model