Cloud technology options towards Free Flow of Data

Version: 19 June 2017

Authors:

Erkuden Rios, Fundación Tecnalia Research & Innovation, MUSA project and DPSP Cluster coordinator.
Bernd Prünster, Bojan Suzic, Graz University of Technology, SUNFISH project.
Elsa Prieto and Nicolás Notario, Atos, WITDOM project.
George Suciu, BEIA Consult International, SWITCH project.
Jose Francisco Ruiz, Atos, Coco Cloud and TREDISEC project.
Leire Orue-Echevarria, Fundación Tecnalia Research & Innovation, OPERANDO project.
Massimiliano Rak, University of Campania "Luigi Vanvitelli"/CeRICT, SPECS project.
Nicola Franchetto, ICT Legal Consulting, CloudWatch2 project.
Paolo Balboni, ICT Legal Consulting, CloudWatch2 project.
Plixavra Vogiatzoglou, KU Leuven Centre for IT & IP Law, CLARUS project.
Rafael Mulero, Fundació Clínic per a la Recerca Biomèdica, CLARUS project.
Sabrina De Capitani di Vimercati and Pierangela Samarati, Università degli Studi di Milano, ESCUDO-CLOUD project.
Simone Braun, CAS Software AG, PaaSword project.
Stephanie Parker, Trust-IT Services, CLARUS project.
Stephan Krenn, AIT Austrian Institute of Technology GmbH, CREDENTIAL project.
Thomas Carnehult, SICS Swedish ICT, PaaSword project.
Thomas Länger, UNIL Université de Lausanne, PRISMACLOUD project.
Thomas Lorünser, AIT Austrian Institute of Technology GmbH, PRISMACLOUD project.
Thierry Chevallier, AKKA, CLARUS project.

Abstract:

This whitepaper collects the technology solutions that the projects in the Data Protection, Security and Privacy Cluster propose to address the challenges raised by the working areas of the Free Flow of Data initiative. The document describes the technologies, methodologies, models, and tools researched and developed by the clustered projects mapped to the ten areas of work of the Free Flow of Data initiative. The aim is to facilitate the identification of the state-of-the-art of technology options towards solving the data security and privacy challenges posed by the Free Flow of Data initiative in Europe. The document gives reference to the Cluster, the individual projects and the technologies produced by them.

Keywords: Free Flow of Data, Digital Single Market, DSM, Free movement of data, Ownership, Cloud computing, data protection, security, privacy, DPSP cluster.

Acknowledgment:

The authors would like to thank the European Commission (DG CONNECT) for its support to the clustering of projects working on Cloud Computing research, and in particular for encouraging the work of the DPSP Cluster. We send a special thanks to Francisco Medeiros who is the Commission official facilitating the work of the Cluster since its launch in April 2015.

Disclaimer:

The contents of this document reflect only the authors' view and the European Commission is not responsible for any use that may be made of the information it contains.

Table of Contents

1. Introduction

This Whitepaper describes solutions, methodologies and technologies, as well as technical and legal considerations for addressing the issues related to the Free Flow of Data initiative #14 of the EU Digital Single Market. The work is the result of the collaborative effort by the Cluster of EU-funded research projects working on the areas of data protection, security and privacy in the Cloud, the DPSP Cluster launched in April 2015 by the DG Connect Software & Services, Cloud Computing (DG-CNECT) of the European Commission.

The solutions and considerations collected herein are the result of the synergy effects of all the clustered projects working together. Currently 28 projects participate in the DPSP Cluster with a total EU funding of approximately €86M, corresponding to 19 projects funded in H2020 (€64M), 6 projects in FP7 (€17M), 2 projects in CIP (€5M). For more information on the cluster and the projects within please visit the cluster's website[1].

In the following, in order to better understand the context of the whitepaper, we first provide in Section 2 a short description of the Free Flow of Data (FFD) Initiative of the Digital Single Market strategy, and the main challenges addressed by it. In Section 3 we describe the methodology followed for collecting and describing the identified project contributions to address some of the aspects of the FFD initiative. Section 4 provides the collection of technical and methodological solutions and approaches from the projects that contribute to solve some of the aspects raised by the FFD initiative.
In Section 5 we provide a summary of the available outcomes from the projects ordered by FFD areas of work, as well as some clarifications that can help towards addressing FFD.
Finally, the Section 6 concludes the whitepaper.

2. The Free Flow of Data Initiative

The Digital Single Market (DSM)[2] is the Pillar I of the Europe 2020 Strategy[3]. The DSM strategy *aims to open up digital opportunities for people and business and enhance Europe's position as a world leader in the digital economy*[4].

As part of the DSM strategy, the Free Flow of Data, Initiative #14 of the DSM, was described as: The *'Free flow of data' initiative tackles restrictions on the free movement of data for reasons other than the protection of personal data within the EU and unjustified restrictions on the location of data for storage or processing purposes. It will address the emerging issues of ownership, interoperability, usability and access to data in situations such as business-to-business, business to consumer, machine generated and machine-to-machine data. It will encourage access to public data to help drive innovation. The Commission will launch a European Cloud initiative including cloud services certification, contracts, switching of cloud services providers and a research open science cloud.*

[1] https://eucloudclusters.wordpress.com/data-protection-security-and-privacy-in-the-cloud/
[2] http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52015DC0192&from=EN
[3] http://ec.europa.eu/priorities/digital-single-market_en
[4] https://ec.europa.eu/digital-agenda/en/digital-single-market

The main topics addressed by the Initiative #14 have been underlined in the text above. These topics are the main areas of work that will be used in this whitepaper to relate the projects' outcomes to the FFD initiative.

On October 2016, the issued the *Inception Impact Assessment of the European free flow of data initiative within the Digital Single Market*[5] proposing the roadmap for a Legislative proposal.
On January 2017, the European Commission adopted the *Communication on Building a European Data Economy*[6], accompanied by a *Staff Working Document on the free flow of data and emerging issues of the European data economy*[7], where it:

> *•looks at the rules and regulations impeding the free flow of data and present options to remove unjustified or disproportionate data location restrictions, and*
>
> *•outlines legal issues regarding access to and transfer of data, data portability and liability of non-personal, machine-generated digital data*[8].

In April 2017 finalised the public consultation by the Commission on Building the European Data Economy[9] started in January 2017 and the summary report on the results can be found here[10].
Note that *this consultation does not cover any issues related to personal data protection. These are extensively regulated elsewhere, namely in the new EU data protection rules*[11]*, as well as through the review of the ePrivacy Directive*[12].

3. Methodology

Based on the Cloud challenges described in the Cluster's Whitepaper *Challenges for trustworthy (multi-)Cloud-based services in the Digital Single Market*[13], we identified a grouping of research gaps on data protection, security and privacy in Cloud that served for creating the Working Groups of the Cluster and distributing the areas of work as follows:

WG1: Advanced security and data protection mechanisms:

- Full control of data flow (including cross-border).
- Efficient (searchable) encryption and key management.
- Secure and privacy-preserving multi-tenancy.

[5] http://ec.europa.eu/smart-regulation/roadmaps/docs/2016_cnect_001_free_flow_data_en.pdf
[6] https://ec.europa.eu/digital-single-market/en/news/communication-building-european-data-economy
[7] https://ec.europa.eu/digital-single-market/en/news/staff-working-document-free-flow-data-and-emerging-issues-european-data-economy
[8] https://ec.europa.eu/digital-single-market/en/building-european-data-economy
[9]https://ec.europa.eu/digital-single-market/en/news/public-consultation-building-european-data-economy
[10] https://ec.europa.eu/digital-single-market/en/news/summary-report-public-consultation-building-european-data-economy
[11] http://ec.europa.eu/justice/data-protection/reform/index_en.htm
[12] https://ec.europa.eu/digital-single-market/en/news/proposal-regulation-privacy-and-electronic-communications
[13] https://eucloudclusters.files.wordpress.com/2015/05/dpspcluster-whitepaper-v3-1.pdf

- Fully secure APIs.
- Security and privacy-by-design
- Security and privacy Requirements modelling
- Fine-grained policy definitions
- Risk assessment frameworks (scalability, multi-technology)
- Secure dynamic composition (brokering, CSP benchmarking)
- Continuous control and assurance

WG2: Trust & Interoperability:

- Data protection legal framework transparency
- Security & privacy aware cloud SLA management.
- Cloud security certification
- Interoperability mechanisms

As it can be seen many of these areas are related to the areas of work of the Free Flow of Data initiative. The specific challenges of the FFD were collected in a paper[14] for CLOUD FORWARD 2016 conference.

As part of the work of the Cluster to help the European research landscape advance towards future challenges, the Cluster decided to collaborate in describing in this whitepaper what technologies and methodologies related to the FFD are already discussed and developed within the clustered projects, i.e. explain to what extent the path towards making the FFD technologically possible was already initiated by the projects.

To this aim, the clustered projects were asked to describe the technologies, methods, techniques, mechanisms, etc. from their work that were addressing free flow of data areas of work. For a better understanding of the solutions motivation, they were asked to do so by explaining first the case studies of the projects which are the context in which the solutions are being or were developed.

Note that some of the projects in the cluster contributing to this whitepaper are already finished though their open source solutions are still available in the projects' websites and/or in public repositories like github[15], bitbucket[16] or AppHub[17]. These code references, as well as the ones of on-going projects, can always be found in the corresponding project website.

[14] http://www.sciencedirect.com/science/article/pii/S187705091632107X
[15] https://github.com/
[16] https://bitbucket.org/product
[17] https://www.apphub.eu.com/bin/view/Main/

4. Contributions of projects towards Free Flow of Data needs

This section explains the on-going results and contributions of the clustered projects that address the different work areas of the Free Flow of Data initiative.

The section a) first describes the cases study areas of the contributing projects, so the research solutions are better contextualised. Then we explain the research topics and outcomes per project: in section b) we provide the collection of methodological results while in section c) we collect the technological results, i.e. methodology supporting tool oriented results. Each of the sections is concluded with a summary table.

a. Projects' Case study areas and solutions related to Free Flow of Data

In this section we provide a short description of clustered projects' case studies explaining the issues or challenges they bring related to the different areas of the Free Flow of Data Initiative. Some of these challenges are the current focus of the projects which are working on solutions addressing them. Such solutions are explained in next subsections.

Case study topics addressing some of the issues in FFD:

- Design of data formats in a secure way
- Risk assessment of data movement
- Design location-aware services
- Design of fully secure APIs
- Data flow monitoring
- Data protection
- Data anonymization
- User-centric consent management
- Privacy requirements formalization
- Security and data privacy in a holistic way
- Safeguard personal & business data in the cloud
- Protect the data persistency layer
- Facilitate context-aware access to encrypted and physically distributed data
- Data sharing agreements
- Data-centric security
- Data storage efficiency
- Multi-tenancy
- Access control
- Confidentiality of data
- Secure data migration to a cloud
- Trusted authentication

CLARUS

The objective of CLARUS, www.clarussecure.eu, is to enhance trust in cloud computing services by developing a secure framework for storing and processing of data outsourced to the cloud. This model change will give control back to data owners and increase transparency about data management, privacy and security. It thus improves levels of acceptance of cloud technology and creates new business opportunities.

CLARUS service proposition is a proxy that will be installed in the trusted domain of the end-user to provide a transparent solution to preserve the confidentiality of personal data and guarantee data protection before data are outsourced to the cloud for storage and processing. The proxy relies on the assumption that the Cloud Service Provider is "honest but curious", as such it will perform honestly the operations on the data as requested by the user, but it might also attempt to learn from the data. To address the need for privacy while leveraging the computational and storage capabilities of public Cloud Service Providers (CSPs), CLARUS proposes a set of privacy-preserving techniques leading to the concept of security as a service, implemented by the CLARUS proxy, which holds the keys and manages the knowledge to restore outsourced and secured data. The security-enabling techniques are a set of cryptographic primitives useful in the cloud context: searchable encryption, access control, homomorphic encryption, and secure multiparty computation. In the context of privacy-preserving techniques, a set of non-cryptographic techniques for the cloud has been defined: statistical disclosure control, data coarsening, data splitting.

Case Study 1: Geospatial Data Demonstration

With a view to the first case study, the CLARUS solution is demonstrated on sets of geospatial data, which refer to environmental and geographical information. Datasets in the environmental domain possess interesting characteristics like the enormous size of the available data, the different degrees of access rights and the availability of metadata, which must be considered while applying the CLARUS solution. Environmental information is also highly relevant for the Free Flow of Data within the Digital Single Market, as highlighted in the EC Staff Working Document on Building a European Data Economy through free flow of data and cloud computing services[18].

In general, the nature of the data included in geospatial information varies, as they may be non-personal or personal, confidential or public, requiring thus diverse levels of access. While the term personal data refers to information relating to an identified or an identifiable person, non-personal data is the exact opposite, in the sense that no person may be identified in relation to them. An example of personal data in the context of geospatial data may refer to people who have accessed, read or downloaded geospatial data or who have used a particular service in relation to environmental information. This way, actors operating in the geospatial scenario might own and manage information, which either may be available in the public domain or may be confidential and therefore must be protected. Security tools are needed for protected data used in commercial settings by private companies, analysts or other institutions in the public sector.

[18] https://ec.europa.eu/digital-single-market/en/news/staff-working-document-free-flow-data-and-emerging-issues-european-data-economy.

The CLARUS geospatial demonstration case focuses on specific scenarios in the field of environmental data management, where cloud technologies are used and where CLARUS could bring solutions to important security expectations, such as storage of geo-referenced data; geo-publication of groundwater borehole data; geo-processing of mineral concentration data and geo-collaboration on gas supply network data.

A key advantage of this approach is being able to extend these scenarios to more generic use cases in the field of Geographic Information Systems (GIS), namely storing geospatial data; searching and retrieving geospatial data and performing computations on geospatial data and updated geospatial data.

Another important aspect of the geospatial use case is the requirement of making services interoperable, and thus for making them compliant with standards. Standards in the geospatial domain are defined by the Open Geospatial Consortium (OGC). Among these standards, the OGC web services standards are of utmost importance for implementing the scenarios described above, namely the Web Map Service (WMS) for serving maps on the web from several geo-referenced data sources, the Web Feature Service (WFS) for exchanging geographical features across the web, and the Web Processing Service (WPS) for invoking transformation services on the Internet.

While cloud architectures provide actors in the geospatial domain with a high-quality, robust and cost-effective service, some geospatial data is confidential and their usage in the cloud raises security issues. Thus, some European public institutions and Data Providers are still reluctant to “move to the cloud”, due to the perceived threats on data security, user control on their data, and data location.

Securing the publication and the processing of their data is a key challenge for geospatial data providers, who often want to limit access to some of their spatial datasets and data services, due to public security concerns or to commercial concerns. This is notably the case for European geo-survey organisations whose mission includes the management of confidential environmental data, beside the legal obligations to share public data to a large audience.

The geospatial use case is of great interest for potential CLARUS adopters, as it shows how the solution can adapt to a highly-standardised landscape (cf. OGC standards), via its plug-in mechanism for protocol support.

In addition, the geospatial use case applies to data held by public authorities and thus it shows how CLARUS end-users can monitor, audit and retain control of their data without impairing the functionality and cost-saving benefits of cloud services.

One of the key features of CLARUS is to support multi-usage scenarios for outsourcing data to the cloud by applying different security techniques. The geospatial use case demonstrates this feature through a variety of scenarios, showing the broad range of technical solutions available, for example:

- Hiding precise location of objects to non-authorised parties thanks to anonymisation/coarsening techniques.
- Protecting geographical features thanks to distributed data splitting among different CSPs.
- Protecting the result of a geo-statistical computation thanks to encryption.

The CLARUS security framework for outsourcing data to the cloud is in line with the security expectations of actors in the geospatial domain. Adding CLARUS to a spatial data cloud infrastructure will mitigate the security threats and strengthen the trust from cloud users, i.e. data providers and data consumers. CLARUS helps geospatial data providers gain confidence in the cloud, providing them with control of their data in the context of honest but curious cloud service providers (CSP).

Among the numerous use cases for datasets and services in the geospatial domain, geo-publication and geo-processing in the cloud are probably the most common scenarios where CLARUS will provide a solution to important confidentiality requirements. The CLARUS solution will address the concern of security in geospatial data sharing, particularly in the event of a regional or national disaster, one of the major reasons cited by organisations for failing to share data e.g. in the case of emergency response.

In addition, as location data may provide for the identification of individuals, including their habits and routines, CLARUS could in the near-future be an answer to the problem of privacy in the use of location based services (i.e. location privacy issues). Other possible applications of CLARUS in the geospatial domain could be satellite imagery (protecting sensitive data in very high-resolution products) and health geo-statistics (privacy-preserving health statistics related to environmental factors).

Legal analysis of the geo-publication use case

In applying the CLARUS solution to geospatial data there are various legal aspects that need taking into account. As mentioned above, geospatial information is mainly non-personal data but they may include personal data, as for instance the personal log-in details to a geo-data related service. Furthermore, different sets of non-personal data fall under different legal obligations of publication or protection. This way, data held by the public sector that are critical for public safety or security or data with a strong business potential and personal data, will need to remain confidential while other environmental information may need to be made available according to national or EU laws.

Indeed, access to information held by the public is dictated by national freedom of information (FOI) laws, as at an EU level this is dominated by the principle of subsidiarity. According to this principle, there are areas which do not fall within the EU's exclusive competence but rather remain within the competences of the Member State due to their national character, as it is agreed in the Treaties signed for the birth and function of the European Union[19]. In the said areas, the Union acts only if and in so far as the objectives of the proposed action cannot be sufficiently achieved by the Member States, but can rather be better achieved at Union level. Even though, on a national level FOI laws may stipulate different conditions for providing access to information, there are three European Directives regarding access related to environmental and spatial data which have significance in relation to the geospatial data being used during the CLARUS project. The ACCESS Directive regulates public access to environmental information[20], the INSPIRE Directive establishes a legal basis for the creation of the

[19] Treaty on European Union, article 5(3), http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A12012M%2FTXT , Treaty on the Functioning of the European Union, http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A12012E%2FTXT

[20] Directive 2003/4/EC of the European Parliament and of the Council of 28 January 2003 on public access to environmental information and repealing Council Directive 90/313/EEC, L 41/26.

Infrastructure for Spatial Information in the European Community[21] and the PSI Re-use Directive refers to the re-use of public sector information[22].

According to the ACCESS Directive, public authorities are required to make environmental information available to the public either through express request or proactively of their own initiative. As such, it ensures that citizens are able to access environmental data in order to participate and assess the governmental decision-making process. This Directive defines environmental information broadly, as information on the state of the elements of the environment, on factors such as energy, on measures such as policies affecting or likely to affect the above, on reports on the implementation of environmental legislation, on economic analyses within this context and on the state of human safety and health. The framework includes the way relevant information should be disseminated, for example through policies, plans and programmes relating to the environment, data or summaries of data derived from the monitoring of activities affecting, or likely to affect, the environment or environmental impact studies and risk assessments concerning the environmental elements. In addition, it provides for grounds not to make this information available, in situations where there is a legal obligation to maintain the confidentiality of the data, as, for instance, under the data protection regime. More specifically, these content related exceptions can only be invoked if the disclosure of the information would "adversely affect" the interests that are protected and they must be interpreted in a restrictive way in a balancing of the respective interests, *in casu* the right to the protection of personal data.

The INSPIRE Directive focuses on the exchange of spatial data between public authorities regarding the performance of public tasks related to the environment and the facilitation of public access to this information to the point necessary. 'Spatial data', as defined in this Directive, is a narrower term relating to data with a direct or indirect reference to a specific location or geographical area, while 'spatial data set' means an identifiable collection of spatial data. As such, there is a small overlap with the above-mentioned ACCESS Directive. The latter prevails over the INSPIRE Directive in case of conflict though. However, the INSPIRE Directive goes further in creating detailed rules on the availability of high quality metadata for all data sets and services. In fact, 'metadata' within the framework of this Directive, refers to information on the conformity of spatial data sets with the implementing rules, to the conditions applying access to and use of spatial data sets and services, to the quality and validity of spatial data sets, to the public authorities responsible for the establishment, management, maintenance and distribution of spatial data sets and services and to the limitations on the public access. Limitations are defined depending on the service information is used for. In this way, public access to data sets provided for discovery may be limited only for severe reasons while public access to data sets provided for other services can be limited for additional reasons that are the same as the ones provided for by the ACCESS Directive.

Finally, the PSI Re-use Directive provides the minimum rules for public authorities to make their data available for non-commercial reuse of existing and public-sector information that is generally available.

[21] Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), L 108/1.

[22] Directive 2003/98/EC of the European Parliament and of the Council on the re-use of public sector information, L 345/90, 17 November 2003 as amended by Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 on the re-use of public sector information, L 175/1.

The rationale behind this Directive is that the public sector collects, produces, reproduces and disseminates a wide range of information in many areas of activity, such as social, economic, geographical, weather, tourist, business, patent and educational information. Making public all generally available documents held by the public sector — concerning not only the political process but also the legal and administrative process — is a fundamental instrument for extending the right to knowledge. However, safeguards must be implemented to protect confidential information, as it is the case with the aforementioned legal instruments. Under this legal framework, the dissemination of these sets of data must not interfere with national security and third parties' intellectual property and data protection rights.

These directives aim at promoting the accessibility of publicly held information to the public and thus stimulating the EU information services market, taking into account the data protection safeguards when this information includes personal data. To that end, the European Commission adopted the European 'Free Flow of Data' initiative regarding non-personal data, as one of the actions within the Digital Single Market strategy[23]. Non-personal data are data that do not relate to an identified or identifiable natural person, such as anonymized data. At the moment, there is no comprehensive legal framework regulating non-personal data amongst Member States, while on the contrary there is a plethora of national laws imposing technical and legal barriers to their free movement across the EU. In particular, the main problem identified is data localisation restrictions, i.e. rules or practices that specify a particular, often geographically defined, area where specific data needs to be collected, processed or stored, while issues like data ownership, data portability and access to and transfer of data are similarly troubling.

As it is pointed out in the EC Communication and Staff Working Document on Building a European data economy, data localisation restrictions facilitate scrutiny and access by competent authorities as well as security of the data but they also become financially and practically cumbersome for businesses[24]. In the context of cloud computing, data localisation restrictions hamper the very nature of cloud computing, while ensuring data portability guarantees an enhanced use of cloud computing services. At the same time, as vast amounts of data are generated by machines or processes based on emerging technologies, such as the Internet of Things, access to those data and possibility of transferring them should be provided for in order to extract maximum value out of them. Limitations to protect confidentiality, personal data, intellectual property and so on should also be imposed as a counterbalance however.

In order to tackle these issues, the European Commission is taking actions towards the abolishment of unnecessary national data localisation restrictions and is engaging in dialogues with the stakeholders to explore manifold solution. This initiative is also complemented by the European Cloud Initiative in

[23] Free Flow of Data Inception Impact Assessment (IIA), November 2016, available at http://ec.europa.eu/smart-regulation/roadmaps/docs/2016_cnect_001_free_flow_data_en.pdf

[24] EC Communication, "Building a European Data Economy", COM(2017) 9, 10.01.2017, available at https://ec.europa.eu/digital-single-market/en/news/communication-building-european-data-economy and EC Staff Working Document on the free flow of data and emerging issues of the European data economy Accompanying the document Communication Building a European data economy, 10.1.2017 SWD(2017) 2 final, available at https://ec.europa.eu/digital-single-market/en/news/staff-working-document-free-flow-data-and-emerging-issues-european-data-economy

enhancing the digital economy and the free movement of data[25]. The CLARUS solution is set to benefit from these initiatives as the barriers on cloud computing will be mitigated, as well as promote them, as its technology can contribute to the different degrees of access to data, the secure transfer of data and data portability.

Case Study 2: The eHealth Demonstration Case

eHealth is a key vertical for the European Digital Single Market. However, concerns about data privacy and security abound, also in the light of high numbers of data breaches at healthcare facilities in both the U.S. and Europe. It is also important to note that healthcare is a highly regulated vertical, making compliance a key driver for securing sensitive data.

In CLARUS, the eHealth use case concerns a distributed e-health scenario that requires immediate access to medical data outsourced to cloud providers. The main actor in this use case is the hospital responsible for treating the Electronic Medical Records (EMRs) of the patients, which contain information that is highly identifying or confidential. A series of functionalities are needed, like creating, managing and updating medical histories, including results of clinical visits, searching for specific patients/histories, as well as shared and cooperative access to these data based on the defined access policies.

In this scenario, the CLARUS solution will need searchable encryption methods used to ensure robust protection and data retrieval capabilities on outsourced health records, but also anonymisation techniques will be employed to securely outsource medical datasets that are still useful for research (e.g., data analysts outside of CLARUS). The use case focuses on passive Electronic Health Records (EHR), which refer to information from patients that have no contact with the hospital, for any reason, for a period of 5 years or more.

This information, which is stored on premise at the hospital, represents a large amount of data that the hospital needs but does not frequently use, making it a storage consumption problem. In addition, this information can be processed for research purposes, requiring several computation resources that can impact negatively on the performance of the hospital information system performance.

Outsourcing this information to the cloud is an opportunity to solve these problems. The space and computation resources used for this information will be available to handle active EHR data (information about patients who are in contact with the hospital for healthcare purposes).

According to several data privacy laws, clinical data is characterised as sensitive, confidential and private. While there are many cloud solutions available on the market, they are perceived as “honest but curious”, and therefore not considered an option for handling this kind of data. The key challenge for the e-Health case is making it possible to store, retrieve and compute sensitive information in a secure way, applying different security mechanisms to avoid that data privacy can be compromised if the cloud is accessed by unwanted users.

The case can be used as an example of how to outsource highly sensitive and confidential data to cloud applying security techniques like searchable encryption and anonymisation (k-anonymity and t-

[25] More information on the site of the European Commission available at https://ec.europa.eu/digital-single-market/en/cloud

closeness). Searchable encryption allows encrypting data before moving to cloud and then performing the search process directly over the encrypted data without prior decryption, which guarantees that data is shown in clear only to allowed users. Anonymisation techniques (k-anonymity and t-closeness) allows to mask data in a way that data can be processed and computed while privacy is not compromised if the cloud is accessed by unwanted users.

Adopting the CLARUS solution could be an opportunity for the e-Health sector to start using cloud platforms, improving, among others, data sharing between different healthcare entities and the quality of research studies related to different healthcare areas.

Legal analysis of the eHealth case study:

The eHealth use case, as described above, includes medical data and in this sense, personal data and more specifically special categories of personal data, also known as sensitive data. This term refers to data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for uniquely identifying a natural person, data concerning health or data concerning a natural person’s sex life or sexual orientation. The data protection regime, as it was regulated by Directive 95/46/EC, soon to be replaced by the General Data Protection Regulation 20016/679/EU, has introduced a wide range of rules that will be applicable in this use case[26].

It is important to emphasise, however, that although the aim of the GDPR is to harmonise the legal framework, the laws of the Member States are allowed to diverge from the Regulation, when explicitly foreseen. For example, regarding the processing of sensitive data the Regulation provides a margin of manoeuvre for Member States’ to restrict or specify its rules and thus Member States are allowed to specify or introduce further conditions for the processing depending, inter alia, on the nature of the data concerned.

Concerning the different messaging and format standards used by different medical institutions, making it thus difficult to exchange information in a common way between hospitals, the GDPR also addresses this issue in Article 20, which establishes the new right to data portability, under certain conditions. In particular, where controllers process personal data through automated means, data subjects have the right to receive the personal data concerning them from the controllers in a structured, commonly used, machine-readable and interoperable format, whenever data subjects provided the personal data and the processing of this personal data is based on their consent, the processing is carried out by automatic means or the processing is necessary for the performance of a contract.

[26] 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 (Data Protection Directive), O.J. L 281, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML , Regulation 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), http://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1495790670928&uri=CELEX:32016R0679

As far as the processing of sensitive data for research purposes is concerned, the GDPR aims to promote innovation and encourage research. Thus, it defines the term "research" broadly (recital 159) by stipulating that research "include(s) for example technological development and demonstration, fundamental research, applied research and privately funded research(..)".

More specifically, regarding the primary use of research data relating to health, meaning when personal data is originally collected for research purposes, the legal grounds for processing the data will be with the consent of the patient. Nevertheless, consent is not always a prerequisite for processing health data for research purposes. For instance, the Belgian Data Protection Act determines that health data may also be processed if necessary for substantial reasons of public interest or when necessary for population screening.

Regarding the secondary use of research data, meaning the further processing of data for historical, statistical or scientific purposes, the GDPR addresses the issue of compatible use more extensively compared to the current legal framework set by the Directive 95/46. It explicitly mentions that "further processing for scientific, historical and research purposes shall not be considered incompatible with the initial purposes and foresees specific conditions regarding compatibility." Therefore, at the European level, the mechanism for further processing of data for research purposes can be summarised as follows. When the purposes of the research can be fulfilled by further processing data which do not permit or do not any longer permit the identification of data subjects, the research should be fulfilled in this manner: Pseudonymisation can be included as a technical measure, as long as it allows the purpose of the research to be met. But if the latter is not met, then other appropriate safeguards (incumbent to the Member States to define) should be put in place to protect the rights and freedoms of the data subjects.

As mentioned above, under the national regime for research, a Member State law may also foresee derogations to the right of the data subjects to access data processed on them, to request rectification, to restrict processing, and to object, unless the research is of significant public interest. Also, reflecting the difficulties of pure anonymisation, the GDPR encourages the pseudonymisation technique, to which it refers in numerous provisions.

Finally, regarding the access of public data by law enforcement agencies and the respective authorisation procedures, they can vary significantly across jurisdictions with differing oversight mechanisms since this is an area largely regulated by national legislation. Thus, it should be reiterated that such agencies will generally be afforded express powers by Statute to operate and gain access under certain circumstances, and particular controls. For many countries, this involves the exercising of some form of warrant dependent on, inter alia, the type of information to be accessed and the urgency of the matter (i.e., a matter of national security). Furthermore, article 48 of the GDPR includes a provision concerning the recognition and enforcement of 'any judgment of a court or tribunal and any decision of an administrative authority of a third country requiring a controller or processor to transfer or disclose personal data'. Therefore, such judgments or decisions may only be recognised or enforceable in any manner, if based on an international agreement, such as a mutual legal assistance treaty, in force between the requesting third country and the Union or a Member State.

As a general conclusion, the GDPR aims to strengthen data subject's rights, but also to promote innovation by encouraging research initiatives. To this end, it provides a broad definition for research, as well as numerous exceptions from the purpose limitation principle and other data subjects' rights in this context. However, in order to benefit from the above-mentioned exceptions, both researchers and member states must have in place adequate safeguards for the effective protection of personal data. With particular reference to research, the key changes introduced by the GDPR can be summarised as follows.

(a) Increased responsibilities for research organisations: Under the GDPR the principles of privacy by design and by default, in respect of the main principle of data minimisation will be the standard approach for data collection and use. Controllers and processors will now have more enhanced accountability obligations to maintain extensive records on data processing activities. In addition, organisations will have to undertake privacy impact assessments, to notify risky data breaches to the DPAs and to affected data subjects, in cases of high risks and damage caused by a breach, as well as to appoint data protection officers, when the organisation is involved in regular and systematic monitoring or processing of sensitive personal data on a large scale.

(b) Profiling: The GDPR explicitly prohibits the use of an individual's sensitive personal data for profiling purposes, unless (a) that individual has given his/her explicit consent (except where a law provides that such prohibition cannot be lifted by the individual's consent); or(b) such profiling is necessary for reasons of public interest.

(c) Consent: Consent must be specific and evidenced by clear affirmative action and explicit consent is required from individuals to process special categories of data (i.e. health related data). All information notices including privacy policies and research consent forms must be written in plain and intelligible language, while consent must be as easy to withdraw as it is to give. It is noteworthy to mention at this point that data for research purposes can also be processed by relying on the "legitimate interests of the data controller", thus without the need to obtain consent from the data subject, under the condition that this does not override the rights of individuals. At this point, it should be noted that consent is a matter usually addressed also on a national level by ethics committees, which provide for additional standards that need taking into account.

(d) National regimes for scientific, statistical and historical research: A considerable margin of manoeuvre is provided to the Member States to derogate from the obligations of the GDPR regarding research purposes, under the condition that they provide adequate safeguards. This possibility provides a dispensation from data subject rights to access, rectification of inaccurate data, restriction of processing and to object, including processing for research purposes. However, it should be highlighted that in that case, research must be done in line with recognised ethical research standards and by implementing appropriate technical and organisational safeguards, such as data minimisation and pseudonymisation.

(f) Penalties and fines: The significant penalties for non-compliance with fines of up to 4% of worldwide turnover or €20 million, point out the significance of the obligation to comply with the rules as set out in the GDPR.

CloudWatch2

As Coordination and Support Action, CloudWatch2 did not address specific case studies in its research. As a partner of this project, ICT Legal Consulting mainly provided legal advice to Public Administrations and SMEs when dealing with cloud service providers.

Coco Cloud

Coco Cloud project aims at allowing the cloud users to securely and privately share their data in the cloud. This will increase the trust of users in the cloud services and thus increase their widespread adoption with consequent benefits for the users and in general for digital economy.

Case Study 1: Pilot on secure data access in the Italian Public Administration sector

A Municipality provides access to the civil data (i.e., the vital events of citizens and residents) it manages through a cloud-based infrastructure in order to facilitate the ubiquitous interactions with other Public Administrations (PAs). In particular, the Municipality office responsible for managing the civil data and willing to share them provides an online web-based service (available at the institutional web site) for the use by other PAs. The data access is regulated through a number of contractual clauses specified in a legal binding or data sharing agreement. The clauses concern both the rights and obligations of the parties involved in the agreement, the duration of the agreement, the responsibilities of every party, and the technical rules that regulate the access to the data and to the infrastructure hosting the data themselves.

The Municipality office defines the agreement in natural language. The convention specifically refers to the technical guidelines on secure data access, the document released by Agenzia per l’Italia Digitale; that is, both the agreement and technical rules compose the overall set of data access policies. The data access control and management is undertaken by the Municipality office by means of the definition of authorization profiles for specific users. In particular, there will be a person in charge of the management of the data at the Municipality office, and a person in charge of using the data at the PA that accesses the data.

According to the national regulation, for which data referring to vital events of citizens cannot be used for purposes different from the institutional functions, the deployment of the pilot will be based on simulated data that have the same structure as the real data. The pilot will be then run at Agenzia per l’Italia Digitale premises within a cloud-based infrastructure hosting the CoCo Cloud system. Agenzia per l’Italia Digitale will also investigate the possibility to engage an Italian municipality during the life of the project in order to collect further requirements and use real civil data and real conventions.

The pilot will show how agreements on secure data access, currently foreseen by the national legislation, can be monitored and enforced in an automate way within a cloud-based environment. Potentially, this can lead PAs to increase their level of trust in cloud solutions, thus paving the way to a larger adoption of cloud computing. This will also bring inherent advantages for the Public Administration sector in terms of reduction of maintenance costs and more efficient services to be provided to end-users and to other PAs. Finally, more effective collaborations between PAs and a higher number of ICT-based data exchange transactions can be envisaged.

Case Study 2: Mobile – Bring Your Own Device

The aim of the pilot is to demonstrate how Coco Cloud can significantly change and simplify the application of corporate security policies on confidential data. The pilot will be ran with a selection of the SAP mobile workforce, for example including Customer Development units. It will focus on secure storage, and it will deal with sensitive data to be consumed on mobile devices, according to the applicable SAP confidential policies. In particular, it will focus on showcasing a number of Coco Cloud contributions, in order to demonstrate:

- the transformation of natural-language policy(es), like the SAP's confidential document policy, into machine-readable DSA;
- the confidentiality-preserving functionalities of Coco Cloud-enabled cloud storage services, where confidential data will be stored and distributed, according to their previously mentioned DSAs;
- the consumption of confidential document on a mobile Coco Cloud client platform, that would allow users to cope with SAP's confidential policies, by enforcing DSA conditions without requiring any active user involvement in the process.

The application of the Coco Cloud approach represents a clear improvement with respect to the current situation, as an automatic enforcement of usage control directives would support employees in coping with corporate policies, thus preventing also unintentional violations. The pilot will demonstrate different results of the Coco Cloud project, from DSA and Cloud elements to the mobile enforcement infrastructure, and will involve mobile business units the validity of the approach.

The pilot will involve a number of Coco Cloud contributions, thus demonstrating:

- the translation from a natural language policy to a DSA;
- the distribution of DSA-regulated documents from a cloud service to a Coco Cloud mobile application;
- the mobile enforcement infrastructure. The pilot will be ran together with SAP mobile business units in evaluating the validity of the approach, and it will also make use of a mobile SAP SDK, the SAP's Sybase Unwired Platform (SUP) for the development

Case Study 3: eHealth Pilot

The system will enable a straightforward connection with the Hospital Cloud infrastructure by offering itself as a new service of medical imaging diagnosis and follow-up. This solution will consist of four kinds of components:

- System administration. A database located at the hospital to handle user management, alerts and information flow.
- Middleware. To ensure the integrity and compliance of all the medical imaging and reports information with medical information standards (HL7, DICOM).

- The Cloud connector. Acting as a bridge between the hospital systems and the Cloud where its services are deployed.
- Mobile devices, such as smartphones and tablets with specific software will also access the system. Depending also on the approach envisioned from the exploitation point of view (BYOD against providing the pair device+application), this software will be published on Apple Stores/Google Play.

Aligned with the market strategy of Health Market on Atos Research & Innovation, the monitoring module will be developed alongside with its Cloud infrastructure and services that will provide the building blocks of the global solution of Atos for Interoperability. This will enable Atos Cloud offerings capabilities and strengthen the presence of the group into the Health Interoperability market.

Indeed, by delivering a concrete solution for a specific Pilot Case of Quirón hospitals, new markets are opened to the ARI Health Market: by using Health Standards for communication and interoperability, the addition of new sensors and connectors with other Health systems is easier to build and maintain, which will enable Atos to provide products adapted to the real needs of Quirón hospitals without losing scalability of the system.

CREDENTIAL

CREDENTIAL (“Secure Cloud Identity Wallet”) is a EU-funded research project developing, testing and showcasing innovative cloud-based services for storing, managing, and sharing digital identity information and other highly critical personal data with a demonstrably higher level of security than other current solutions.

The main idea and ambition of CREDENTIAL is to enable end-to-end security and improved privacy in cloud identity management services for managing secure access control. This is achieved by advancing novel cryptographic technologies and improving strong authentication mechanisms.

The solutions and technologies developed in CREDENTIAL foster free flow of data in two way. On one hand it can be used as fully cloudified privacy friendly IAM solution which can be used to authenticate towards different service provider without leaking any information about one’s attributes to the identity provider. This IdP supports federated or multi-cloud applications. On the other hand is can be directly used as a central data sharing portal which preserves the privacy of the data by encryption. Users can dynamically and selectively control which data they want to share with which other users. To the extent possible, also metadata privacy problems (e.g., who accesses which files owned by which other user at which time) will be addressed within CREDENTIAL.

Case Study 1: e-Government

The eGovernment pilot considers citizens who want to remotely pay taxes or request financial support from their local tax office. For instance, the pilot considers a citizen of country A living abroad in another country B, who needs to pay local taxes in country B. Now, he can use his electronic identity card of country A to securely and strongly authenticate himself to the tax portal of country B, potentially using STORK and eIDAS to perform this cross-border authentication.

The CREDENTIAL platform is now used to host authentic personal data that goes beyond the data that is stored on the national eID card. For instance, such data might include pay slips or certificates of registration. The user can now grant the tax authority of country B access to this data. As granting access rights can also be done for documents that will be added to the wallet in the future, the user can easily file certain required documents later without having to contact the tax authority again, but by simply uploading the data to the CREDENTIAL wallet.

FFD issues being addressed: The main issues being addressed are related to strong (potentially cross-border) authentication while giving the citizen full control over which data to share or to keep private. This is done in a way that still guarantees the authenticity of the shared data towards the receiver. All developed solutions will put a special focus on usability and user information to increase the trust into, and adoption-rate of the results. Furthermore, interoperability with existing authentication methods (STORK, etc.) will be guaranteed.

Solutions:

- Privacy-preserving authentication
- Attribute-based credentials on encrypted attributes
- Secure management and backup of sensitive key material
- Integration of advanced cryptographic primitives like proxy re-encryption and malleable signatures
- Development of dedicated software components such as mobile apps

Case Study 2: e-Health

The eHealth pilot is concerned with a data sharing platform between patients, doctors, and further parties, in particular in the context of Type 2 Diabetes. Namely, the developed components will allow patients to record their health data (blood sugar level, weight, blood pressure, etc.) using external mobile devices. The data measured on these devices will be collected by a CREDENTIAL eHealth mobile app, which remotely stores this data in the CREDENTIAL wallet. The user can then define who is allowed to access which parts of this medical data, to share specific parts of the measurements, e.g., with the family doctor, diabetologist, nutritionist, or personal trainer. Based on the data they see, they can then provide recommendations back to the user.

Because of the confidentiality of medical data, it is of prime importance that only legitimate users are able to access a user's data. Furthermore, because of the potential consequences of wrong recommendations, the authenticity and integrity needs to be guaranteed.

FFD issues being addressed: CS 2 puts the user back into control over his own data by giving him full control over which data he wants to share with whom. All developed solutions will put a special focus on usability and user information to increase the trust into, and adoption-rate of the results.

Solutions: Same as CS1.

Case Study 3: e-Business

Besides a classical single sign-on (SSO) functionality, the eBusiness pilot showcases how easy the privacy offered by existing solutions can be enhanced through the integration of modular libraries implementing CREDENTIAL’s technologies. Encrypted mails are a requirement for many companies to protect their data and inventions, but they also represent a significant challenge when employees go on vacation. Currently, employees have to expose their private key material so that a substitute can still read and answer incoming mail. In contrast, with proxy re-encryption, an employee generates a re-encryption key for a substitute before leaving, with which the mail server is able to translate incoming mail during the absence.

FFD issues being addressed: This case study is related to (temporarily) granting access rights to potentially sensitive information to other users, and to delegate access rights to other users. All developed solutions will put a special focus on usability and user information to increase the trust into, and adoption-rate of the results.

Solutions: Same as CS1.

ESCUDO-CLOUD

The ESCUDO-CLOUD project aims at empowering data owners as first class citizens of the cloud. ESCUDO-CLOUD provides effective and deployable solutions allowing data owners to maintain control over their data when relying on Cloud Service Providers (CSPs) for data storage, processing, and management, without sacrificing on functionality.

Case Study 1: OpenStack Framework

The scenario of this use case relates to a Cloud-storage platform that supports server-side encryption with flexible key-management solutions. This use-case is particularly applicable for the development of internal Cloud solutions as well as for CSPs building private or public Cloud solutions using open source frameworks such as OpenStack. In particular, it focuses on data-at-rest encryption and key management solutions to be used with OpenStack Swift, an object-storage system that runs on commodity hardware and provides failure resilience, scalability, and high throughput in software. Encryption occurs on the server side under the governance of the storage provider; encryption inside the storage platform is an important feature for large-scale and enterprise-class storage systems. Coupled with a suitable key-management solution that is able to control and securely erase cryptographic keys, the encryption technology also supports the controlled destruction of data, called *secure data deletion*. Data-at-rest encryption and secure deletion are important requirements for enterprise-level Cloud storage services.

The goal of this use case consists of adding cryptographic protection technology inside a private Cloud platform, in particular, to storage systems. Clients of Cloud services and operators of the CSPs benefit from data encryption in the storage systems, so as to make the system resistant to attacks that target lower layers of the infrastructure.

FFD issues being addressed: The main expected results consist of technologies to protect the confidentiality and the authenticity of the stored data, and to protect data that is shared and concurrently accessed by multiple clients from being altered or modified.

Solutions:

- Integrated at-rest encryption with OpenStack Swift.
- Key-management solutions within OpenStack.

Case Study 2: Secure Enterprise Data Management in the Cloud

The scenario of this use case relates to the outsourcing of supply chain interactions in the aerospace engine maintenance industry. So called, maintenance, repair and overhaul (MRO) providers offer their services to several airlines leveraging cost savings by streamlining the process. In general, two main business-optimizing services have to be guaranteed in the aero engine overhaul supply chain: the Collaborative demand Forecasting (CF) and the Collaborative Planning and Scheduling (CPS) of the overhaul activities.

The first one allows MRO service providers to obtain demand forecasts from all customers based on on-condition engine status observations, reducing so overall costs due to a more accurate capacity planning; while the collaborative planning and scheduling guarantees better supply chain performance, since an ideal receipt point for each engine can be computed. Traditionally, each party on each stage of the supply chain has its rather isolated forecasting processes that are mainly based on data of historical demand that arose from their direct customers. The problem with these orders from the next stage is that they are again results of an isolated forecast and in general do not match the actual sales on the buyer's stage. Instead, they tend to have a larger variance. This effect of demand distortion results in amplified forecasting numbers and dramatic swings in demand increasing with every step on the supply chain further away from the end customer. This phenomenon is known as the bullwhip effect.

However, this information does exist, and CF is an attempt to bring them together to create a single, more accurate forecast that is accepted by all collaborating members of the supply chain. In a collaborative forecasting process ideally all supply chain members add their particular expertise and information to find the best possible forecast. The information about end customer demand is shared with the upstream supplier, so demand distortion can be reduced drastically. This again will drastically reduce the bullwhip effect.

A central issue for maintenance and support service providers concerns the management of the growing amount of information generated by the development of highly complex aircraft systems and by stakeholders' requirements in terms of dependability increase and LSC decrease. To face these problems, maintenance and support actors are depending more and more on ICT solutions. These are one of the main elements not only to improve the effectiveness and efficiency of the maintenance process for complex systems with a long lifecycle, but also to reduce the associated risks and to contribute to a more efficient business process. The benefits linked to the use of ICT systems in this business segment are:

- more controlled content sharing;
- information exchange and knowledge management;
- coordination of maintenance process with other processes;
- connection to strategic business objectives and external stakeholder requirements.

FFD issues being addressed: The main expected results consist of developing new supply chain cooperation systems, based on encrypted database technology. This technology is based on the search and aggregation of encrypted data and can be applied in planning the MRO service to different customers without knowing the actual status of the aero fleet nor the capacity usage of the service provider, nor the inventory status.

Solutions:

- Encrypted database for ERP and other applications with support for search.
- Support for fine-grained data access through encryption.

Case Study 3: Federated Secure Cloud Storage

This use case considers the application of data protection in Multi-Cloud environments including federated secure Cloud storage. It will offer data protection as a service via a Cloud service store that enables customers to protect their data stored on multiple Cloud platforms. As such this use case is particularly applicable for Managed Security Service Providers and for Cloud user organisations who want to control the protection of data at rest across multiple Cloud environments and apply uniform data protection and data access policies for heterogeneous data stored in a multiplicity of Cloud providers (including internal, private and public Clouds). The data is protected by leveraging a Cloud-based data protection service for encrypting independently of the Cloud provider hosting it, and ensuring that CSPs have no access to the encryption keys or its protection and access control policies. The encrypted data can be stored on multiple Cloud storage services like block storage, object stores and Big Data clusters. Access control and key management are offered as tightly coupled services that manage the protection of the data via an integrated policy framework. This tight integration will ensure that the decryption of the protected data is only possible in the client's environment following a policy-based approval procedure and the resulting release of the encryption key. The encryption and decryption process will be transparent to applications and end-users while the data-at-rest will always stay in encrypted state on the multiple Cloud platforms, in compliance with specific data security standards and regulations.

FFD issues being addressed: The main expected results consist of ensuring the confidentiality, integrity and availability of the customer's data through a client-side encryption approach. In addition to this core responsibility, the key challenges addressed in this use case are to offer the key management feature and the policy-based access control feature as a service through a Cloud service store. The instance of a key management service and the access control service have to be tightly coupled for each customer, which will allow them to specify key release rules that are applicable only under specific conditions.

Solutions:

- Roadmap for Cloud-of-Clouds support.
- Data-Protection-as-a-Service (DPaaS) for automating compliance.

Case Study 4: Elastic Cloud Service Provide

This use case considers the use of an elastic CSP. As such this use-case is of particular relevance for Cloud service brokers or intermediaries offering a secure Cloud data storage capability to their customers while possibly leveraging other Cloud providers for storing this data and ensuring that data are protected from such other Cloud providers and other users. In this use case scenario a data owner has access to his/her files stored in the Cloud using some middleware available on the client (web portal or agent). The middleware will enable the communication between the end user and the CSP. Furthermore, by using an elastic Cloud, it will be possible to adapt the capacity of the Cloud to the requirements of the user. Additionally, by using the ESCUDO-CLOUD middleware with an elastic Cloud, the encryption provided by ESCUDO-CLOUD will be present in the third party Cloud too (so, ESCUDO-CLOUD will be present for all data transfers). With the ESCUDO-CLOUD middleware, the users would have secure access to their data hosted by the Cloud providers such that the data are not compromised. They will also be able to manage their data from a web browser or an agent (installed in their devices). Finally, it will be possible to synchronize stored files from third-parties like Google Drive or Dropbox. So, by using ESCUDO-CLOUD the users will be able to leverage the Cloud services that they would typically request from a CSP but with a higher level of security, access control, and assurance.

FFD issues being addressed: The main expected results consist of an elastic, cloud-based secure storage and data synchronization service by leveraging 3rd party cloud infrastructures and data storage services while the intermediary assures and governs the data protection independently of the cloud hosts.

Solutions:

- Developed architecture for hybrid cloud security architecture.

MUSA

The MUSA project is working on the creation and operation of secure multi-cloud applications. Multi-cloud applications, with components deployed in or using heterogeneous clouds, ask for the **free movement of data**, because no restrictions should be applied to where the application components are deployed, unless for technical reasons (e.g. latency issues).

Case Study 1: Flight scheduling application

The application consists in a commercial product that Lufthansa Systems company develops for easing the nightmare of flight scheduling task. It is a multi-cloud based application where different components are deployed in heterogeneous clouds mainly due to performance and security reasons.

FFD issues being addressed: Some of the major issues MUSA is working to solve are the need of protecting the access to the different application components, ensuring the location of the components and the data processed by them. The project also works towards aligning the definition of the overall application security requirements with the offered composed Service Level Agreement (SLA) and with the measures (mechanisms) actually implemented in the application to ensure such requirements are fulfilled, i.e. the guarantees in the SLA are not violated. This involves the alignment with and monitoring of the cloud providers' SLAs.

Solutions:

- Integrated framework for the DevOps and agile based engineering and operation of the multi-cloud application taking into account the security requirements of the application and the security offerings of the cloud providers to use.

The parts of the framework which also work as independent tools are:

- Web Modeller of multi-cloud applications (based on CAMEL language) that allows the specification of the deployment and security requirements.
- DevOps oriented Risk Analysis tool.
- Cloud Service Providers (CSP) selection Decision Support Tool, including security features in selection process.
- Security and privacy-aware SLA Generator.
- Multi-cloud Deployer.
- MUSA Security Assurance Platform for continuous runtime monitoring of composite SLA fulfilment and security features enforcement.
- Security monitoring agents for multi-event monitoring in distributed applications.
- Security enforcement mechanisms (such as access control and scalability) actionable at runtime operation.

Case Study 2: Smart mobility application for Tampere City (Finland).

This multi-cloud application developed by Tampere University of Technology provides Tampere citizens with a smart service for multi-modal and energy efficient commuting. This service exploits other existing services in the Transport Systems and Services (ITS) platform[27], which includes the public transport services APIs, traffic related APIs, etc.

FFD issues being addressed: The major issues being addressed are the protection of the citizens' personal data (e.g. the mobility footprint and their energy consumption profile) even in cases when the cloud provider does not provide strong security, and the creation and enforcement of the security and privacy-aware SLA that is the composition of the SLAs of the individual services being used. The current Finish law also asks for the Finish citizens' private data be located in Finland, therefore there is a location issue for one of the storage components in the application.

Solutions: Same as CS1.

[27] Intelligent Transport Systems and Services (ITS) Factory Wiki, http://wiki.itsfactory.fi/index.php/ITS_Factory_Developer_Wiki

OPERANDO

The OPERANDO project is working on the creation of a platform that will used by independent Privacy Service Providers (PSPs) to provide comprehensive user privacy enforcement. the project has two target groups, Business to Consumers (B2C) and Government to Consumers (G2C). The OPERANDO platform has as main aim to protect and safeguard the data privacy of a user of a digital B2C and G2C service, who also need to comply with strict legislation and regulations. In some B2C cases, however, in which Online Service Providers (OSPs) utilize users' private data coming from e.g. social networks login data, the user shall have also the opportunity to benefit from the exploitation and usage of his own data by means of a new concept named as "privacy for benefit".

Case Study 1: Food Coach by Ospedale San Raffaele (Italy)

Food Coach is an application that lets users on one hand, to take advantage of the dietary advice automatically provided by the Food Coach engine, and on the other hand, to provide people affected by pathologies, e.g., diabetes or obesity with a common infrastructure where patients' doctors can monitor the health status of the patients and interact with them, tuning their diets.

FFD issues being addressed: OPERANDO deals with safeguarding private data on the cloud, with a major focus on health-related data which are amongst the most sensitive ones. To achieve that, several anonymization techniques and algorithms are being used, researched and implemented. Furthermore, novel security methods for authentication are being extended.

Solutions:

- Integrated Platform to implement the "Privacy Authority" concept including components for Rights Management, authorization, anonymization, big data analytics, and so on;
- Monitoring of changes in the privacy settings of OSPs via a Privacy Watchdog;
- Provision of a policy computation engine, which acts as decision support engine for providing privacy aware services.
- Web Browser plug-ins for enforcing anti-trackers, manage identities, protect data leakage and scan reputation;
- Mobile Application supporting the services to be delivered by the privacy service provider, including an application permission scanner, a reputation scanner, and the corresponding identity manager;
- Several APIs, such as the Regulator API, which facilitates compliance with privacy laws as well as auditing and pro-active supervision of OSPs by privacy regulators;

Case Study 2: AmI (UK)

The AMI use case describes how local authorities in the UK can gather information about clients who require assistance with various day-to-day activities, and the volunteers who help with these situations. AMI using big data analytics in order to analyse needs against available resources producing reports that can be utilised in order to forecast service needs effectively. The AMI use case collects and stores data about both the clients who need help and the volunteers.

FFD issues being addressed: Same issues as above.

Solutions: Same as CS1. The anonymization techniques whatsoever change in this case, since not only patients' data need to remain anonymized but also the data from the volunteers.

Case Study 3: ASL Bergamo (Italy)

ASL Bergamo has started a new public service for helping people with a gambling addiction in to manage this mental disorder and to obtain information about this problem in the Bergamo Area.

FFD issues being addressed: Same issues as above.

Solutions: Same as CS1. The anonymization techniques whatsoever change in this case, since fake personal data will be used.

PaaSword

PaaSword offers a security by design framework enabling the cloud application developer to define security annotations, transparently integrated into IDEs. Those annotations are then transformed into context-aware security policies that enforce access control, cryptographic protection and physical distribution for securing sensitive data.

PaaSword provides encrypted storage and context-aware access control for Platform-as-a-Service. It thereby delivers key components to protect next generation cloud applications against internal and external adversaries.

In brief Paasword provides:

- PaaSword holistic framework
- Policy-based access control & context-aware security models
- Policy enforcement middleware
- Searchable encryption scheme for secure queries
- Dedicated IDE plug-in

Case Study 1: Protection of personal data in a multi-tenant CRM environment by CAS

The platform provides a core CRM application on top of which customer and sector-specific modules can be developed, deployed, and executed in parallel. A CRM system typically includes a large amount of sensitive data that need to be protected; in addition, this amount is further increased when customizing the system for specific domains like healthcare or education, naturally increasing the demand for security.

FFD issues being addressed: A multi-tenant cloud system like CAS OPEN needs to prevent not only external adversaries (for which various mechanisms are already in place) from gaining access to the sensitive data stored and exchanged, but also internal adversaries. Thus, by integrating the PaaSword framework in the CRM platform, CAS aims at ensuring that certain highly-sensitive data(-types) are inaccessible to internal adversaries.

Solution:

This case will showcase the following concrete usage scenarios:

1. Security support for part of the CRM/xRM solution.
2. Support for additional tenant isolation in the CRM/xRM solution through the provision of private storage.
3. Secure key management.
4. Management of user and role permissions with additional refinements based on contextual information (e.g. sensitive data accessible to a user only if the user is located within the campus of the educational institution adopting the CRM/xRM solution).

From an implementation and integration perspective, the following key steps will be performed and evaluated:

1. Integration of PaaSword framework into CAS Open Cloud Application Developer perspective
2. Implementation of a secure CAS Campus cloud solution on top of PaaSword-enhanced CAS Open Cloud-enabled application developer perspective
3. Installation, operation and maintenance of a PaaSword-secured cloud application vendor system engineer perspective and end user perspective

Case Study 2: Cloud application management platform by SixSq

SixSq provides cloud technologies that facilitate the effective use of cloud computing in the enterprise. SixSq's cloud management platform, SlipStream, delivers efficiency savings via automation and self-service IT. SixSq's turnkey cloud appliance, NuvlaBox, brings together all of these concepts, delivering an easy-to-use, affordable and unique private cloud solution.

SixSq provides a cloud application management platform, Nuvla (https://nuv.la), built with SlipStream, a multi-cloud application management platform (see http://sixsq.com/products/slipstream) that facilitates the effective use of cloud computing in the enterprise.

FFD issues being addressed: One major impediment that businesses and other institutes continually identify is the secure storage of data on the cloud. This becomes even more of a barrier as the legal requirements for data protection become more onerous. We would like to provide an option, using the PaaSword components, to allow sensitive data to be treated on the public cloud.

Solution:

By integrating the PaaSword components, the developers using Nuvla can quickly develop applications that securely handle sensitive data while avoiding the expense and delay of developing and validating similar functionality themselves.

In addition, SixSq can potentially benefit by integrating some of the ideas developed within the project within the SlipStream software itself. Incorporating context-aware authorization or protecting sensitive data within SlipStream's database are two examples where these ideas will positively impact SixSq's products.

Case Study 3: Secure Intergovernmental Document and Personal Data Exchange by Ubitech

UBITECH maintains a secure information exchange platform that can be used in order to facilitate legally binding document exchanges between various organisations. The product is called ubi:eXchange. UBITECH will augment the existing product by transforming it into a Qualified e-Delivery platform that can tackle the specificities of vertical domains. The financial domain is the first one that will be targeted since it offers concrete business opportunities related to the rapidly growing demand of trust services. Such demand extends horizontally to any domain; yet in the frame of ubi:eXchange, the focus is on the vertical needs of the banking sector.

FFD issues being addressed: The bank sector and financial institutions are primarily affected by the leveraged, in terms of strictness, regulatory framework at the European level that obliges them to satisfy harsh requirements that relate to security and privacy. Such requirements refer both to internal and external processes.

Solution:

The PaaSword-enabled ubi:eXchange platform will focus in particular on utilizing the following features:

- Security Model adaptation
- Dynamic Policy enforcement
- Data Fragmentation

This pilot will showcase the following concrete usage scenarios:

1. Create Complex Policies and Policy sets that will be transparently applied (Design time conflict resolution & Reasoning benefits)
2. Apply controlled-fragmentation schemes for the e-delivery process
3. Use multiple encryption algorithms for the TDE for tenant encryption

Case Study 4: Secure Sensors Data Fusion and Analytics by Siemens

The scope of SIEMENS' pilot is to offer a test-bed for the evaluation of PaaSword's implemented mechanisms under a data fusion and analytics setting. More specifically, SIEMENS will assume a smart transportation scenario in which an array of sensors constantly monitors and reports relevant transport data (e.g. cargo temperature, vibration, humidity, transport times and routes).

FFD issues being addressed: The main features targeted by the SIEMENS pilot are:

- Secure storage and processing of sensitive business data used in monitoring and decision making.
- Support for non-standard NoSQL databases through the secure data storage and privacy-enhancing mechanism.
- Tenant isolation and secure tenant-specific applications.

Solution:

This Siemens pilot will showcase the following concrete usage scenarios:

1. Development of the trial through the inclusion of the PaaSword-based NoSQL database adapter
2. Definition and enforcing of the context-aware policies regarding data access of different actors to the available data
3. Integration of the key management mechanism into the trial architecture.

Case Study 5: Protection of Sensible Enterprise Information in Multi-tenant ERP Environments by SingularLogic

SingularLogic offers Enterprise Resource Planning (ERP) solutions which support single- and multi-tenant use cases, and which rely on IaaS deployment schemes. For the piloting phase of PaaSword, a multi-tenant ERP solution that is targeted to SMEs will be used.

FFD issues being addressed: SingularLogic expect this to improve their ERP solutions in terms of confidentiality and privacy, as for us, in multi-tenant ERPs the main security risk is the exposure of the data of one tenant to other tenants. In addition, they foresee to provide users with improvements in other aspects of security such as the protection of partially-compromised data as well as context-aware access control.

Solution:

The SILO use case is based on the following real world scenarios:

- Implementation of context-aware access control in the PaaSword-enabled ERP solution through the usage of PaaSword annotations.
- Encryption and secure storage of sensitive data of the PaaSword-enabled ERP solution through the use of PaaSword annotations.

PRISMACLOUD

The Horizon 2020 research project PRISMACLOUD (short for ‘privacy and security maintaining services in the cloud’) is dedicated to enabling secure and trustworthy cloud-based services by improving and adopting novel tools from cryptographic research. Several cryptographic primitives and protocols with TRL (technology readiness level) TRL3 or higher, are being advanced to TRL7 (‘system prototype demo in operational environment’), and are encapsulated in cryptographic tools. The tools are used to construct cloud services with unique end to end security and privacy features—for the benefit of cloud providers and end users. In one sentence, the main idea and ambition of PRISMACLOUD is to provide the tools for enabling end-to-end security and privacy protection for end users by the best technical means possible: by cryptography.

Achieving end-to-end security in data sharing systems, and building privacy features into them by design, can significantly increase the trust in the communication infrastructure for data sharing. Thus, our designs can contribute to the overall goals of the free flow of data initiative.

Case Study 1: Medical data sharing portal (e-Health)

The e-Health use case proposed by the PRISMACLOUD project aims at supporting secure and safe interaction between patients and healthcare providers, as well as between different hospital services and the clinicians. The main objective of this use case is to add several privacy and security features to the Trusted Healthcare Platform (THP) of the Fondatione Centro San Raffaele (FCSR) located in Milano, Italy.

In particular, three of the PRISMALOUD services will be used to extend the features of the THP platform: The *Selective Authentic Exchange* service uses malleable signatures and allows the system on behalf of the patient to redact health documents according to the recipient's need in order to maintain maximum patient privacy. The *Big Data Anonymization* service enables data sharing on open health platforms to the research community providing only anonymized datasets. The *Verifiable Statistics* service enables outsourcing computation (evaluating statistics) on signed data, providing the possibility to audit the correctness of the computed statistics. It is intended to support the privacy-friendly manipulation of user generated data via dedicated devices, e.g., like smart watches. In particular the following main features are provided by the medical data sharing portal:

- Medical data generated by physicians is signed and uploaded to a data sharing portal.
- The patient (i.e. end user) can select which data to share with other parties. He or she can e.g. generate a sick note for the employer, containing only a required minimal subset of the signed data.
- By use of redactable signatures, such selective disclosure reveals only relevant information, while the digital signature of the originator (i.e., physician) remains intact, i.e. the authenticity of the revealed medical data remains intact.
- Personal health data from wearable devices can also be uploaded, and statistics about the data can be shared with other parties in a privacy-friendly way, e.g. only revealing the overall jogging time during one week, without giving information on single tracks or times. The computation on the signed single data items is delegated to the cloud, and the party receiving (in this case) the sum, can cryptographically verify that it is authentic and was calculated correctly.
- Beside sharing on a personal level, the platform enables sharing bulk data for studies or other purposes. Therefore, an anonymization service for large data sets is integrated to protect individuals' privacy when participating in medical studies.

FFD issues being addressed: The issues being addressed by the PRISMACLOUD tools in this case study are *data ownership*, *access to data* and *public data access*. The core idea of the use case is that the authenticity of data is protected end-to-end. All data in the system is digitally signed, either by authorities or by the data subject itself, therefore ownership of data can always be proven to external parties and the whole database can be considered of high assurance when shared with stakeholders. Furthermore, by usage of redactable signatures selective sharing is supported for authentic data. Therefore, access to data can be restricted for different stakeholders but without destroying the authenticity of data. The user is in control of what parts of information are shared with whom, thus efficiently regulating access to his data. For the case of self generated data, statistics are processed in verifiable manner to convince others about the achievement of personal goals in a privacy preserving manner, e.g., imagine an insurance company giving discounts for active people running certain distances per week. In such a scenario you can prove the achievement of the goal without disclosing detailed activity data like location or time. In addition, when users decide to participate in clinical studies they can be assured that only anonymized records are shared by authorities, therefore protecting privacy when generating data for public access. In general, in PRISMACLOUD we believe, that building privacy features into systems by design could persuade even more people to share data and foster an even larger data economy.

Solutions:

PRISMACLOUD provides solutions on several levels or layers, which accompany the development of the demo applications for the case study.

Solutions on methodological level:

- Holistic security models for developed services and applications.
- A new development methodology (CryptSDL) which extends classical approaches to assist in the complexity of secure cryptographic service design.

Specific ready to use services:

- Selective Authentic Exchange service based on redactable signatures (SAEaaS).
- Verifiable Statistics service for authentic data aggregation (VSaaS).
- Big Data Anonymization service for large sets of data (BDAaaS).

Cryptographic research and development of tools:

- Flexible Authentication with Selective Disclosure tool (FLEXAUTH).
- Verifiable Data Processing tool (VERIDAP).
- Data Privacy tool (DATPRIV).

Case Study 2: Evidence sharing platform for law enforcement (Smart Cities)

The scenario is built upon existing systems developed by ETRA I+D, being a Parking Control System (ParkPlaz) and a CCTV System for traffic management purposes. ParkPlaz is a system for controlling any kind of parking facility. It receives input from different kinds of sensors, mainly presence sensors indicating the occupancy status of each of the parking lots, and license plate scanners installed in the entrances and exits of the parking facilities. These license plate detectors automatically provide pictures of the vehicles coming in and out of the parking, which are stored together with certain metadata at the parking facilities, as enforced by Spanish law 44/2006, respecting the existing legal constraints regarding the maximum storage period.

Due to the kind of information that is being collected (evidences of accesses to parkings in the first case, incident videos in the second case), law enforcement units do usually require the provision of pictures or videos to be used in legal processes. This usually involves a bureaucratic procedure that could be improved by using current cloud technologies. The purpose of this use case is to develop a cloud-based sharing system, where sensible data (pictures of license plates, video of incidents, and all the metadata linked to them) can be uploaded and shared in a secure and privacy friendly way, i.e. such that the cloud provider does not learn the data nor is able to tamper with the data, thus making the access of law enforcement units to this information easier and faster and more reliable and authentic. This system makes use of the novelties developed in PRISMACLOUD to ensure that this data can be securely stored in the cloud, also convincingly maintaining the chain of custody, and granting access only to legitimate actors.

FFD issues being addressed: The issues being addressed by the PRISMACLOUD tools in this case study are *access to data*, *usability*, *interoperability*, *switch of CSPs* and *cloud contracts*.

Basically the system is a cloud data sharing solution which builds on the concept of data fragmentation and distributed multi-cloud platforms. It is based on the Archistar system (http://archistar.at) which encodes data and disperses them over multiple clouds and trust zones in a way, that only a predefined subset of the data is required to reconstruct the data. However, single fragments reveal no information about the data, hence, the overall system is preserving the confidentiality and integrity of data while increasing the availability of the overall system.

Therefore, access to the data is protected, i.e. unauthorized access by single or small sets of colluding cloud storage providers is securely and provably prevented. However—because encoding is used and not encryption—no keys have to be managed, i.e. the system usability is improved compared to conventionally encrypted cloud storage services. Additionally, the problem of provider lock-in is also removed from cloud customers. The system resembles a virtual secure storage service on top of multiple less secure and less reliable storage offers. It supports interoperability by design, and the possibility for fragment renewal puts the end user in the position to effectively decide about moving shares from one provider to another, thus effectively preventing lock in with a provider. The same mechanism allows also for secure and effective deletion of data throughout all level of a provider’s storage architecture. The combination of multiple functionalities gives the operator of the PRISMACLOUD service more flexibility in achieving specific SLA attributes, currently not available in plain storage offers.

Solutions:

The solutions used on methodological level are the same as in Case Study 1. On a technological level, the case study builds on the:

- Data Sharing service (DSaaS) of the PRISMACLOUD services layer, which uses the
- Secure Object Storage tool (SECOSTOR) of the tools layer.

Both the DSaaS, as well as the SECOSTOR tool, will be provided as reference implementation by project end.

Case Study 3: Cloud backup and archiving service with location attestation (eGovernment)

Lombardia Informatica (LISPA) is an in house company of *Regione Lombardia*, providing IT services to regional governments and public bodies throughout the Lombardy region in Italy. LISPA is preparing a virtual data centre infrastructure based on cloud technology, which will later hosts services of public bodies of the Lombardy region. Other than centralising the infrastructure in a secure way, new possibilities and services that were not deployable in a decentralised setting need to be considered for further developments.

The PRISMACLOUD framework enables backup to the cloud and scale out scenarios in the e-Government context without sacrificing the confidentiality of data. A European wide distributed cloud storage service can be offered by LISPA by leveraging untrusted individual storage offers. The multi-cloud storage can easily be verified any time and further enables ubiquitous access and sharing capabilities.

FFD issues being addressed: The issues addressed are the ones of Case Study 2, but with another focus and delivery model used, i.e. especially long-term security is a focus in this case study and privacy-friendly possibilities to verify the integrity of archives. The use of long-term secure fragmentation algorithms also allows for computation of simple aggregated statistics over multiple customers. This enables the implementation of a privacy friendly open data strategy in e-Government scenarios.

Solutions:

The solutions used on methodological level are the same as in Case Study 2. On a technological level, the case study builds on the following results:

- The Secure Archiving service is used to store data in multi-cloud setting (SAaaS) which is based on cryptographic research results described in the Secure Object Storage tool (SECOSTOR).
- The Infrastructure Auditing service (IAaaS) can be used to assure geographical or topological properties of the data sharing network without revealing detailed information about inner connectivity or real locations. The service is based on the cryptographic tool called Topology Certification tool (TOPOCERT).

SPECS

SPECS proposes an innovative Platform-as-a-Service that offers a solution for the SPECS' Security-as-a-Service approach, based on SLA life cycle management. The SPECS platform enables the delivering of security services, described and guaranteed through Security SLAs. Cloud Service Customers are able to express at different grain-level the security features they need through a user-centric negotiation of Security SLA, that helps CSCs to effectively negotiate with a set of CSPs, by understanding the resulting trade-offs. Moreover, SPECS offers innovative Security Services to enforce SLA: when a cloud service does not grant the security features that a CSC has expressed in the negotiated SLA, SPECS provides additional security mechanisms that grant such specific feature.

In order to support CSCs to verify the correctness of the services offered by CSPs, SPECS offers innovative solutions for continuous Security Monitoring, it implements SLA-monitoring solutions dedicated to continuously control the security offered by CSP and to help ensuring the granted security service level objectives.

SPECS Framework, i.e. the software collection developed within the project, is open source and can be used by Cloud Service providers to offer their service offerings with Security SLAs and/or by developers in order to develop new (SPECS) applications that enhance the security of public CSPs.

Case Study 1: Secure Web Container

A web developer aims at acquiring a web container to run his/her own application, which fulfils some security requirements. The web container is represented by one or more Virtual Machines (VMs) provided by one or more IaaS CSPs. It is reasonable to suppose that the EU is not an expert in security field: she/he is aware of the technologies that may be involved (SSL, authentication and authorization protocols and so on), but she/he is not aware of the best practices and of how to protect her/his application from malicious attacks. For this reason, the acquisition of VMs and the enforcement of security features are accomplished through SPECS. The SPECS Solution:

- Offers a single interface to select among different offerings on different providers;
- Enables web developer to specify explicitly the needed security capabilities on the target web container, selecting the security controls.
- Automatically configure the VM(s) in order to enforce the security controls requested.
- Offers a set of security metrics in order to concretely monitor the respect of the security requests.
- Automatically remediate to (some of) alerts and violation that may occur to the SLA associated to the web container.
- Enables continuous monitoring of the security metrics negotiated.

FFD issues being addressed: *certification, contracts, switching of cloud services providers, interoperability.*

Solutions: The SPECS Web Container case study offers an example of cloud services able to support explicitly Security Service Level Agreements that enable (semi) automatic certification processes, and a clear support to comparison among different cloud services providers. Moreover, the adoption of machine readable formats to represent security enables secure interoperability process in service orchestration.

Case Study 2: End-to-End Encryption

When storing data with CSPs, cloud customers usually have to accept the risk of security incidents and failures related to modifications and loss of stored data. More than that, customers can never be sure that write-serializability (WS), i.e., consistency among updates, and read-freshness (RF), i.e., requested data always being fresh as of the last update, are always respected. And what is more important, even if customers are aware of data modifications or loss of data, they cannot prove to third parties when the cloud is to blame for WS or RF violations. On the other hand, the cloud provider itself cannot disprove false accusations.

In order to offer to customers secure storage solution and allow them to not only detect but also prove violations related to modification and loss of stored data, Secure Storage capability in SPECS is implemented by the E2EE security mechanism which provides the following functionalities:

- Client-side encryption enforcing confidentiality and integrity (I).
- Detection and proof of violations related to WS and RF.

- Backup of stored data.

While E2EE could be with some limited functionality used independently from SPECS, it is the SPECS platform that provides the functionality to:

- Automatically deploy E2EE components according to the SLA (deploying the virtual machines with all the required E2EE components according to the security properties selected by user).
- Detect violations out of the monitoring events generated by E2EE monitoring.
- Choose the remediation action (like moving the REST API form the primary storage site to the backup) after the analysis of the monitoring events.
- Trigger the remediation actions.

FFD issues being addressed: *ownership, interoperability, usability and access to data.*

Solutions: The SPECS End-2-End Encryption case study offers usable tools to grant protection of data even when shared and made available to others, protecting the ownership of data, granting at same time interoperability, usability and access to them.

Case Study 3: Next Generation Data Center

An End-User aims to acquire storage resources via a cloud provider. EMC's ViPR application enables users to create virtual storage resources built on top of physical resources (e.g. VMAX) via a web interface or using the REST API. Through ViPR, an organisations administrator can create a virtual array on which virtual pools of block or file storage can be built. These virtual pools can be configured with a number of options including RAID level, storage quotas, High Availability, Multi-Path options, disk drive type, etc. In this way, organisations can request the storage resources that meet their requirement without needing to manage the underlying physical resources.

In a typical scenario using ViPR, customers (e.g. department/team managers within an organisation) would submit requests to the ViPR admin staff. A member of the admin team then manually (via the web interface) configures the storage as per the requirements outlined in the request and makes it available to the customer (department/team members).

The limitation of this scenario is the lack of a clearly defined set of security requirements which creates a disparity between what the customer requests and how the administrator interprets that request into provisioned storage. In addition, the customer has no visibility into how the requirements are being enforced. The customer has not visibility on its resources status after the storage allocation.

Thanks to the adoption of the SPECS Framework:

- SPECS adds a layer of control and intelligence on top of the ViPR interface while providing a user friendly interface to the customer that further abstracts away the complex inner workings of ViPR.
- Using the SPECS web interface, the customer itself can specify their storage requirements and sign an SLA with SPECS confirming the enforcement of those requirements.

- SPECS automatically configures the storage and makes it available to the customer without the need for intervention from the admin staff. Both the administrator and customer maintain complete visibility of the process throughout the lifetime of the storage & SLA.
- External security features could be offered by the framework to the Storage Provider extending its native features.
- Additional security metrics are offered to users, like location of data when using bursting features.

FFD issues being addressed: *location of data ownership, interoperability, usability and access to data.*

Solutions: The SPECS Enhanced ViPR solution offers feasible tools to enable secure storage according to end users security needs: an end user which adopts the SPECS enhanced ViPR have clear grants regarding data location and can protect his ownership on data, grating at same time interoperability, usability and access to data.

Case Study 4: Star Watch

STAR Watch is CSA's response to the identified needs of automated decision-making tools to facilitate Cloud procurement processes, by delivering—in a database/machine readable format—the content of CSA's succinct yet comprehensive list of cloud-centric control objectives defined in the Cloud Controls Matrix (CCM) and the corresponding set of control assertion questions in the Consensus Assessments Initiative Questionnaire (CAIQ).

A Premium version of STAR Watch is in development now. It will leverage SPECS' security reasoning techniques to offer the ability to compare cloud service providers by assessing their control matrix responses, and to compare those responses against the enterprise's security requirements.

By adopting Cloud Watch, prospective and current Cloud customers can have a better level of transparency related to the CSPs delivering services to their organisation (even their own private clouds), to assure a consistent security baseline is maintained.

State of the practice lacks of automated tools to aid (prospective) Cloud customers in the process of comparing different CSPs from a security perspective. Thanks to the machine-readable information to be available in STAR Watch's repositories, it's possible to integrate SPECS' security reasoning techniques in order to allow the side-by-side comparison of CSPs based on a baseline set of end-user requirements. End-users don't need to be security experts in order to use this new functionality, because security requirements can be specified at different levels of granularity.

FFD issues being addressed: *switching of cloud services providers, interoperability.*

Solutions: Thanks to the data collected in the CSA repository and to the reasoning tools offered in SPECS it is possible to support switching of cloud services providers through (semi) automated decision processes.

SUNFISH

The emergence of diverse cloud-based solutions, as well as the related and ongoing adoption of cloud-based systems and platforms, has led business and public administrations to reconsider their IT strategies and their organisation of computing infrastructures and assets. Based on that, many organisations embraced these models and transformed their assets so that they increasingly rely on cloud-based approaches.

Considering the need of organisations to support interoperability and cooperation among existing cloud systems deployed at single or different organisations, SUNFISH proposes a new and innovative cloud federation approach based on the concept of *Federation-as-a-Service (FaaS).* The core characteristics of FaaS include it's strong, context-sensitive and transformative security mechanisms supported by a novel governance model that advocates distributed and democratic governance. The FaaS approach provides a secure-by-design federation of clouds and services with advanced management and orchestration capabilities of cloud federations in heterogeneous infrastructures. This is achieved by relying on distributed ledger technologies (blockchain and smart contracts) for strong integrity guarantees as part of a holistic security architecture. Among others features, the FaaS-based approach allows the integration of advanced access control, cryptographic data transformation services, and runtime monitoring facilities, fusing proactive and reactive security mechanisms for integrated security management and compliance.

The practical application of the platform provided by the SUNFISH project is evaluated based on the following case studies which allow integration of cloud infrastructures and services in a cross-organisational context with specific requirements concerning data security and privacy. Additional requirements of these case studies concern optimization models for resource allocation and load balancing, as well as the automated integration of clouds based on different vertical deployment models.

Case Study 1: On-line services for managing personnel salary accounts

Managing payroll systems involves access to highly sensitive data, such as health status, religious orientation, or information on duties performed in the scope of classified actions in the military or police. Some of these data are stored in different systems of diverse entities, including private and public companies, central and local public administrations, and military and police agencies. Each of these stakeholders may have different concerns and practices regarding privacy and security of these data. Additionally, the processes and procedures of these entities dealing with sensitive data or related infrastructure services may differ both on the technical and organisational level.

This study is based on a scenario where the DAG department of the Italian Ministry of Economy (MEF) is in charge of managing payroll functions for approximately two millions of Italian public sector employees. During the complex workflows, which deal with payslip management, MEF retrieves and processes data with different confidentiality requirements, hosted by various data providers. The ministry also provides processed or generated data to other entities, such as taxation agencies at various levels.

In this case, one of the involved stakeholders is the Ministry of Interior (MI), which has to provide MEF with the data needed to execute the payslip generation workflow. However, due to additional security requirements, classified data, and activities present at MI, additional restrictions have to be applied on data leaving the premises of MI. Another challenge in this process is the time-window in which all data exchange and processing has to be performed in order to be valid for each stakeholder. This assumes the availability of appropriate infrastructures with the resources to sustain high periodic loads while at the same time providing the required security level.

In order to address the service integration needs of this use case, the stakeholders federate their cloud infrastructures for the purpose of consolidating and sharing available resources and performing cross-organisational process flows in a timely correct and secure manner.

FFD issues being addressed: The secure cloud federation enables free movement of data, controlled data access, location-independent data access and interoperability through the Federation as a Service concept. This is especially relevant when dealing with sensitive data on a national level in conjunction with periodic peaks regarding required computing resources.

Solutions:

- The secure-by-design approach and integration of transformative controls, such as data masking, enables free movement of data within a cloud federation by balancing between security, privacy, and utility in a context-sensitive manner.
- Data access can be managed at a fine-grained level due to the direct integration of a policy enforcement framework based on the eXtensible Access Control Markup Language (XACML).
- FaaS components, including security and monitoring infrastructure, utilize blockchain technology to ensure high level of integrity and assure the execution of critical processes and evaluate its correctness.
- The FaaS concept ensures interoperability, as members of a cloud federation can change.
- The management framework allows monitoring of SLAs and automated optimization based on workflow scheduling and resource scaling.

Case Study 2: PaaS in public clouds processing sensitive personal information

The Maltese taxation departments within the Ministry of Finance (MFIN) require periodic data submission from taxpayers, employers and other institutions such as banks. The submitted information is highly sensitive as it encompasses payroll data, trading records, information about savings as well as personal information and accounting records. Submission of these data is required in order to compute possible tax deductions or back taxes.

Even though this process is carried out on a national level, there is currently no fully-automated and widely accepted solution in place to perform these tasks. One reason for this lack of automation is that some precomputations need to be done by each submitting party before transferring the data to the government institutions. Most larger administrations, as well as private entities of a certain size, already have the necessary infrastructure and processes in place to submit all financial data electronically. In these cases, the precomputation also occurs mostly automated. Smaller companies, however, often rely on manual paperwork and consequently submit the required data on paper or through spreadsheets. Naturally, the taxation department does not want to impose high costs, especially on small companies, but still aims to automate the whole process as much as possible. Consequently, it is unrealistic that small or medium businesses will invest in computing resources for fully-automated computation and submission of financial records. One way to tackle these issues in a cost-effective manner would be to utilize cloud resources. However, privacy concerns have prohibited the adoption of public clouds for such use cases.

The SUNFISH project can address these issues by providing a federation between MFIN's private cloud and public clouds. Since SUNFISH follows a secure-by-design approach and integrates data masking mechanisms, as well as fine-grained access control mechanisms based on the industry standard XACML, a federation between private and public clouds can be achieved while still satisfying all privacy requirements. This effectively leads to an efficient utilization of available resources as SUNFISH enables secure resource sharing between all participants of federated clouds. Confidentiality, integrity, and availability are ensured by SLA and data access policies directly supported by the SUNFISH framework. Such a cloud federation can be created and managed (including policy management) using a single coherent interface as desired, thus providing a federation as a service.

FFD issues being addressed: The secure cloud federation enables free movement of data, controlled data access, location-independent data access and interoperability through the Federation as a Service concept. Again, this is especially relevant when dealing with sensitive data on a national level. In this case, however, privacy is the primary issue to be tackled, due to the reliance on public cloud.

Solutions:

- Free movement of data is inherently enabled by providing strong security guarantees and transformative data security enforcement for cross-organisational interactions.
- Free data movement is also fostered by providing standardized interfaces for data exchange and federation of diverse cloud infrastructures and hosted services.
- SLA is directly addressed and enforcement of cloud contracts is directly supported by data masking and access control mechanisms present in the SUNFISH federation. Furthermore, federated infrastructure allows the monitoring of different SLA metrics and automated alerting in case of contract breaches.
- Data access is controlled in such a way, that it is possible to outsource data into hostile environments without compromising confidentiality, integrity and availability while at the same time granting authorized entities full access to their data.
- The FaaS concept fosters interoperability, since from the FaaS users' point of view, it is irrelevant which cloud providers are combined into a federation.

Case Study 3: Secure Federated Cloud System for Cyber Intelligence Data Sharing

The South East Regional Organised Crime Unit (SEROCU) forms part of the UK response to Cyber crime and the threat that poses to UK infrastructures. The offences investigated focus on Cyber dependent crimes, whereas the victims range from members of public through small and medium sized businesses and large corporations or government agencies.

SEROCU obtains and stores large quantities of data that is potentially highly sensitive, including high level corporate information through to personal details about public persons. In an typical case, the data is stored on local premises within the operational unit, being not accessible to other entities thatn SEROCU staff. This imposes a range of limitations in a routine work, as the storage server is not open to all external stakeholders relevant to SEROCU operations. In this way, the unit is limited in regards to data sharing and collaboration with other entities.

The SUNFISH project addresses this issue by providing the framework for secure interconnection of distributed storage systems at different premises and organisations for the purpose of intelligence data sharing. This includes the ability to establish cloud-based federations and dynamically manage their lifecycle by allowing the members to join or leave on a permanent or temporary basis. The federated infrastructure should allow the definition and enforcement of different access levels to data, including the searches or data exchange among parties that support anonymized and masked data, using appropriate encryption channels for cross-entity communications.

FFD issues being addressed: The secure cloud federation enables interoperability by connecting different infrastructures, storage and data management systems using a unified and multi-purpose framework for service integration and controlled data exchange. The free flow of data and access to public data are supported by integrated indexing and search management platform that allows sharing only a partial subset of data, potentially anonymized or masked to conform to security requirements.

The framework supports enforcement based on data and user locality, allowing a high degree of usability for user that need to search through and evaluate a large sets of data. By integrating SLA monitoring and evaluation in the framework, the involved stakeholders are able to monitor the enforcement of contracts both in the term of performances, location-awareness and correctness of security policy enforcement.

Solutions:

- Cloud federation framework aimed at the federation of cloud storage infrastructures among different premises and organisations.
- Data indexing and search framework that integrates data security and privacy requirements and allows data owners and administrators to granularly define data indexing, searching and retrieving capabilities, including the visibility of data and application transformational measures (anonymization, encryption, masking) to facilitate its availability to a range of actors with different security clearance levels.
- Security management framework based on XACML that integrates other federation services using obligations.
- Supporting secure multiparty computation (SMC).

SWITCH

SWITCH addresses the urgent industrial need to develop and execute time critical applications in Clouds. Applications such as disaster early warnings, collaborative communication and live event broadcasting can only realize their expected business value when they meet critical requirements in terms of performance and user experience.

SWITCH targets:

- Software industry: to support software development and consultancy companies in delivering time-critical applications and services.
- Cloud service providers: to enable SLAs for time-critical services.
- Telecom service providers: for network providers and infrastructure operators.
- SMEs and entrepreneurs: for operating and developing their own applications with time critical requirements.
- Education organisations / Universities: for education/training purposes.
- For a wide collection of domains that require time critical services: Time critical applications in specific domains.
- Technology vendors including API management companies SDN and virtualization vendors, Telecom-managed service providers, and wireless/mobile infrastructure providers.

The very high requirements posed on network and computing services, particularly for well-tuned software architecture with sophisticated data communication optimization, implies that development of such time critical applications is often customized to dedicated infrastructure, and that system performance is difficult to maintain when the infrastructure changes.

This fatal shortcoming in the existing architecture and software tools yields very high development costs, and makes it difficult to fully utilize the virtualized, programmable services provided by networked Clouds to improve system productivity.

Furthermore, SWITCH aims to improve existing development and execution models of time critical applications by introducing a novel conceptual model (application-infrastructure co-programming and control model), in which application QoS/QoE, together with the programmability and controllability of the Cloud environments, can all be included in the complete lifecycle of applications.

Based on this conceptual model, SWITCH provides an interactive environment to develop applications and control their execution, a real-time infrastructure planner to deploy applications in Clouds, and an autonomous system adaptation platform to monitor and adapt system behavior.

Time-critical applications are required to respond immediately to a range of events that may occur at runtime. Often the quality of service (QoS) given directly impacts business value (e.g. for multimedia platforms) or public safety (e.g. for disaster response). Many such applications are distributed and highly demanding. Cloud environments provide on-demand virtualized infrastructure that could support such applications, but there is a lack of tools for exerting fine-grained control over software-defined infrastructure and applications at runtime.

Case Study 1: A collaborative business communication platform

Collaborative real-time business communication platform, a platform which gathers all communication needed for real time business in most companies. One of the main requirements for this service is the adaptability of the service on the traffic demand while maintaining the quality of the service. The SWITCH software workbench helps companies which work with real time communications in several ways. First the SIDE subsystem allows developers to define the system, at container level with QoS requirements to describe the system. This user interface establishes a common ontology which can be used for different subsystems inside the service or even different services. Second, the DRIP subsystem will be able to check resources needed for the service before starting execution. Moreover, if application must be scaled up, DRIP will provision new resources in a suitable cloud to host new containers while maintaining QoS. Finally, ASAP is responsible to monitor metrics and resources remaining as well as QoS of the service by means of probes which will be deployed in the same host as containers. Also, ASAP will control the deployment of new subservices (running in containers) scaling up/down the service according to demand or QoS requirements.

FFD issues being addressed: Data flow monitoring represents one of the main challenges, as real-time communication plays an increasingly important role for many business applications, videoconferences, cooperative working environment, and remote diagnosis. The service must meet the requirements of QoS from the beginning which makes necessary a test of resources before starting the application.

Solution:

Running on the Cloud, the (Unified Communication) UC platform becomes a service usually called UCaaS – Unified Communication as a Service – and to provide cloud interoperability among different Clouds the micro services comprising the service make use of containers for execution. Moreover, SWITCH enables free flow of voice communication, as the use of containers favours fast adaptability, because the time needed for deployment of new containers in any geographical region is much lower than the time to deploy new virtual machines.

Case Study 2: An elastic disaster early warning system

Early warning for natural disasters is an important challenge for many countries. An early warning system often collects data from real-time sensors, processes the information using tools such as predictive simulation, and provides warning services or interactive facilities for the public to obtain more information.

BEIA provides an IoT Telemetry platform (http://eng.beia-telemetrie.ro), built on Java and Time Series Databases (TSB), that gathers data from different Remote Telemetry Units (RTUs) and facilitates visualization, predictions and notifications in case of various disasters (floods, phyto-sanitary, droughts, air pollution, nuclear radiation, etc.).

FFD issues being addressed: One major challenge which is faced by IoT businesses is to securely transmit data from sensor/edge nodes to the cloud, as well as securely transmitting commands to actuating components. Furthermore, for time critical early warning applications there is a need for resilient and secure storage of data in cloud, especially in case of disasters. Using SWITCH components we envision to handle IoT data across heterogeneous cloud infrastructures.

Solution:

By integrating the SIDE (SWITCH interactive development environment), the developers will be able to use tools for securely developing, deploying and controlling the execution of time-critical applications, supporting every stage of the application lifecycle. BEIA will benefit of an application-infrastructure co-programming and control model that relates application logic, QoS constraints, and developments in programmable infrastructure. Also, BEIA will integrate in its solution the authentication, authorization and auditing mechanisms of SWITCH, while using a search based application for free flow of data, which allows access to partitions of data stored in containers.

Case Study 3: A cloud studio for directing and broadcasting live events

MOG Technologies provides a cloud studio for directing and broadcasting live events that manages the streaming of video feeds and the production of the broadcast stream virtually rather than on-site. The production of live TV events by its very nature requires very strict requirements: delivering video and audio with as little delay as possible while maintaining the quality and security requirements that the television industry requires to ensure the maximum quality of experience (QoE) to viewers.

FFD issues being addressed: In a live event, the broadcaster or production company has to deploy a large number of personnel and many items of equipment to fully cover it. Multiple cameras are placed around the event venue to cover all the different angles that the director considers relevant.

Live TV production, due to its distributed nature, requires broadcasters to deploy equipment and human resources to several different places, increasing production's costs and complexity of data transmission, including encryption. As the performance and cost-effectiveness of IP networks grow, broadcasters can use this to handle real-time production-quality video and audio that is expected in a professional environment.

Similarly, the evolution of virtualization's technologies on the cloud and the efforts developed in order to provide solutions of virtualized, elastic, controllable cloud environments leverages the adoption of this technologic stack for time-critical applications, such as a live TV production. However the usage of such type of technologies is still in its infancy since the requirements for developing high quality live production media workflows and time critical applications are still very high.

Solution:

The SWITCH components will provide to MOG tools for managing the complete lifecycle of time-critical applications within the Cloud, explicitly linking user-level QoS with securely programmable infrastructure and autonomous runtime monitoring and control. MOG will be changing the actual paradigm of professional media production and data flow for broadcasting, which is based on baseband digital media connections, outside trucks/vans and satellite connections. Given that the time critical requirements (e.g latency) are very demanding, SWITCH will be used to validate the tool workbench for the use case.

TREDISEC

Case Study 1: Storage efficiency with security

This first use-case describes the upload, storage and deletion of user data using the ~okeanos cloud storage service in such a way that (a) data confidentiality is guaranteed throughout the data life-cycle and (b) storage and computational efficiency are preserved. Since ~okeanos relies on a block-based storage backend, the objective is to enable block-based deduplication over encrypted data.

- Data protection

- Security and data privacy in a holistic way
- Safeguard personal & business data in the cloud
- Protect the data persistency layer
- Data storage efficiency
- Confidentiality of data

Case Study 2: Multi-tenancy and access control

The second use case, envisaged by GRNET, concerns the management of multiple users that share resources (that could be infrastructure or data) on the Cloud. In such an environment, the cloud is required to support multi-tenancy and to deploy mechanisms for fine-grained access control and low-level resource isolation (physical, VM, OS layers). Such a service should not of course undermine the confidentiality and integrity of the outsourced data. We, again, assume the usage of a block-based deduplication policy for the storage backend.

- Data protection
- Security and data privacy in a holistic way
- Safeguard personal & business data in the cloud
- Protect the data persistency layer
- Facilitate context-aware access to encrypted and physically distributed data
- Multi-tenancy
- Access control
- Confidentiality of data

Case Study 3: Optimised WebDav service for confidential storage

The third use case provided by Arsys is focused on a shared storage service that wants to provide multi-tenancy access control through WebDav access. The use case aims to build a WebDav service upon a confidential storage whereby even ISP administrators cannot violate the confidentiality of users’ data. The service will allow tenants (customers) to manage more than one user with different data access and permissions while data isolation among tenants is guaranteed. Finally, original services or tasks such as user data sharing and storage optimization should not be sacrificed for the security; neither should be the performance of the provided service by allowing whole-file-based data deduplication over encrypted data.

- Data protection
- Security and data privacy in a holistic way
- Safeguard personal & business data in the cloud
- Protect the data persistency layer
- Data-centric security
- Data storage efficiency
- Multi-tenancy
- Access control

- Confidentiality of data

Case Study 4: Enforcement of biometric-based access control

This first use-case, supplied by Morpho, considers biometric-based online authentication. It assumes that a user has to perform some authentication before accessing a service. The biometric-based authentication is delegated to a third party (called Cloud Authentication Server) who performs the authentication and, in addition, provides a proof that the authentication has been correctly performed. Verifiable computation techniques ensure the security of the authentication process when part of this process is outsourced. The objective is to supply efficient proofs for computation correctness.

- Data protection
- Security and data privacy in a holistic way
- Safeguard personal & business data in the cloud
- Data storage efficiency
- Access control
- Confidentiality of data
- Secure data migration to a cloud
- Trusted authentication

Case Study 5: Secure upgrade of biometric systems

The second use-case, also supplied by Morpho, considers major upgrades of biometric systems. Due to the evolution of algorithms and/or biometric data format, updates of existing biometric data should be performed, and sometimes a large amount of biometric data should be processed. In this use-case, these updates are outsourced to a cloud service. For privacy reasons, biometric data must be encrypted. The functionality offered by the cloud is then to generate a new database of encrypted biometric templates from a large amount of encrypted biometric images. The objective is to come up with efficient and scalable encryption techniques compatible with the signal-processing algorithms used to process the biometric images.

- Data protection
- Security and data privacy in a holistic way
- Safeguard personal & business data in the cloud
- Protect the data persistency layer
- Data storage efficiency
- Confidentiality of data
- Secure data migration to a cloud
- Trusted authentication

Case Study 6: Database migration into a secure cloud

This use case describes the migration of a company's legacy data into a secure cloud environment, requiring the data to be encrypted; yet stored in such a way that SQL queries can be executed over it. Encrypting large sets of legacy data (potentially multiple gigabytes of data) could take several months, potentially has impact on the running business, and results in a larger storage footprint so that to enable efficient processing over encrypted data at the cloud.

- Data protection
- Security and data privacy in a holistic way
- Safeguard personal & business data in the cloud
- Protect the data persistency layer
- Data storage efficiency
- Multi-tenancy
- Access control
- Confidentiality of data
- Secure data migration to a cloud
- Trusted authentication

WITDOM

Case Study 1: Genomic sequences operations (eHealth)

Next Generation Sequencing technologies exploit massive parallelization allowing to process millions sequence reads of the biological sample in parallel. This means that they produce a huge amount of short sequences of DNA (200GB[28] for an haploid DNA sequence). This set of data has to be subsequently manipulated according to the aims of different experiments. In WITDOM, these experiments consist in:

- The reconstructions of the DNA contained in some selected genes through an alignment operation, which involves the alignment of the aforementioned short sequences of DNA according to a reference gene sequence;

- The detection of well-known gene sequence variations that are linked to medical conditions (e.g. detect a BRCA1 or BRCA2 mutations which increases a woman's risk of breast and ovarian cancer by factors that range up to x30, for the case of ovarian cancer with the BRCA1mutation)

All these operations are computationally expensive and citizens will largely benefit when organisations are able to outsource them to the cloud, as they dramatically decrease their cost.

[28] The haploid DNA sequence is made of 3 billion of nucleobases. Our encoding requires a quality character for each nucleobase. Assuming a 30x coverage, i.e., the average number of times each base has been read by the sequencer, and supposing that every character requires one byte for being stored, we have at least 180 GB of data.

This use case focuses of effectively protecting genomic data, which is challenging as it is not clear to the medical community what personal data may be inferred from the a DNA sequence (it is clear the relation of some specific genes and physical characteristics such as the colour of the eyes). The genomic data must be protected in two complementary ways. The first is by using anonymization, and data masking techniques (using the Data Masking protection components) to protect the metadata related to the patient to whom the DNA belongs and the second one is to protect the DNA sequences themselves using the Secure Signal Processing protection component which implements the desired operations through Private Information Retrieval (PIR) mechanisms and Secure Multiparty Computation (SMC) protocols. The Protection Orchestrator is the component that coordinates the different protection components and may assess the need and the existence of the patient's consent or the level of the protection of the data achieved before outsourcing it to the cloud. Finally and due to the legal requirements regarding the storage of the data, a combination of an End to End Encryption (E2EE) and an integrity and consistency verification (ICV) protection components are used to create a protected backup of genomic information. WITDOM architecture also includes the concept of transformation services, focused on interoperability and which is demonstrated in this case study. Given that WITDOM will deal with data from different sources and different formats, WITDOM provides its own generic data structure that may accommodate most of existing data formats and will rely on simple transformation services which will move the data from one format to another. As one of the objectives of WITDOM is to minimize the number of modifications to be performed on the end-users' systems and to make transparent the consumption of these privacy-preserving services, WITDOM includes the concept of the Broker. The Broker is the actual element that decides when (according to existing policies and end-users' security preferences) the services are to be provisioned by a trusted version of the service, deployed in the trusted environment, and when it may be executed in the untrusted environment, by a secured service, and only after the adequate protection of the data.

Case Study 2: Outsourcing customers' financial data for cost-effective analysis

For their normal and daily operation, banks have to collect and reflect all customers' financial transactions, including money transfers, ATM withdrawals, credit and debit payments, namely almost any non-cash transaction. This is a huge amount of data: just in the EU, during 2013, 100 billion noncash payments have been executed. By applying advanced analytical techniques such as data mining, streaming analytics, geospatial analytics, predictive modelling, financial organisations are now able to leverage these datasets to create more robust and realistic financial and risk models, with the purpose of improving their decision capabilities, both in terms of speed and quality. The size of the datasets and the kind of analytics performed make this scenario a perfect candidate to be outsourced to the cloud. However, these data is considered personal and due to general and sector-specific legal framework, it must be adequately protected before its outsourcing. WITDOM will demonstrate three different types of analysis within this case study:

- *Fraud detection in credit card transactions*: a large dataset of credit card transactions classified as fraudulent or non-fraudulent will be protected by having its identifiers adequately masked (Data masking protection component), with statistical-based anonymization (Anonymization protection component) and by encrypting some of the data using homomorphic encryption

(Secure computation protection component). Machine learning classification algorithms will be applied to the outsourced data in order to develop a fraud-detection classifier that can be real-time applied to new transactions.

- *Credit risk scoring*: credit risk scoring models are used to quantify customers or operations' risk. Credit risk scoring models are usually based on the creation regression models to estimate credit risk on new operations according to their similarity to previous ones (and if they were defaulted or not). The personal data involved in this analysis is more sensible than the one involved in the fraud detection one, however, mostly the same protection components will be involved (Data Masking, Anonymization). The Secured signal processing protection component will, in this case, be responsible of performing the actual analysis on the protected data in order to build the regression model.
- *Cash flow forecasting*: looking at the past of a customer's transactions it is easy to find periodic patterns such as its monthly incomes and fixed expenditures (i.e. mortgages, water and electricity bills). These transactions can be the used to forecast the future and propose to the customer products to address potential cash shortages or to suggest saving products whenever an excess of cash is detected. In this case, the secure signal processing protection component will apply ARIMA models to forecast customers' accounts balance.

In all cases, at a technical level, the Protection Orchestrator will coordinate the different protection components and ensure the adequate level of data protection is met before its outsourcing. The decision of what "adequate level" is the outcome of the application of a requirement elicitation methodology (SPACE) and WITDOM's privacy framework which supports the formalization of technical requirements through the definition of privacy properties, metrics and thresholds in a close dialogue with the stakeholders. As in WITDOM's Genomic sequences operations case study, the Broker is used to decide when, according to the bank's preferences the bank's private cloud or a public cloud should be used to perform the different analytics.

The following table summarises the topics described in the case studies description above.

Table 4a. Topics addressed in project case studies related to FFD working areas

Project	Case Study No.	GDPR Issue/FFD working area	Case study topic	Priority
CLARUS	**CS1**	**Free Movement of data**	Geospatial data services in the Cloud with trust in security enforcement thanks to CLARUS Solution.	1
		Location of data	The CLARUS privacy-preserving mechanisms are key innovations. Examples include: Data splitting – only CLARUS knows cloud locations for a given dataset. Homomorphic encryption – secret is stored in CLARUS. Data encryption – CSP never has access to data or keys.	1
		Ownership		3
		Access to data	The overriding goal of CLARUS is to give users increased control over data. **Access Control** functionality (storage, request and search).	1
		Access to public data	Geospatial data of both public and sensitive data must be handled with CLARUS Solution with low overhead in configuration of services.	1
		Usability	Performance assessment for computation services, especially on Geospatial Data.	3
		Interoperability	Support of main OGC Web Services standards	1
		Switch of CSPs	No CSP Lock-in with demonstration of interaction between CLARUS Proxies and Clouds.	2
		Cloud certifications	N/A	
		Cloud contracts (SLA)	Performance assessment for computation services, especially on Geospatial Data with analysis also on SLA-Ready Common Reference Model planned.	4
	CS2	**Free Movement of data**	Statistics computation use cases require share of sensible Health data, possibly among several organisations using CLARUS solution.	1
		Location of data		5
		Ownership	Patient files remain the property of each organisation.	1
		Access to data		2
		Access to public data	N/A	
		Usability	Data base management systems and	3

			File systems are used.	
		Interoperability	Standard data models support	2
		Switch of CSPs		
		Cloud certifications	Trust by end-users (hospital services)	2
		Cloud contracts (SLA)		
Coco Cloud	**CS1-CS2-CS3**	**Free Movement of data**	· Data anonymization · Data sharing agreement · Data-centric security · User-centric consent management	5
		Location of data		4
		Ownership		4
		Access to data		4
		Access to public data		4
		Usability		4
		Interoperability		4
		Switch of CSPs		1
		Cloud certifications		2
		Cloud contracts (SLA)		3
CREDENTIAL	**CS1-CS2-CS3**	**Free Movement of data**	• High data and meta-data privacy guarantees.	3
		Location of data	N/A	1
		Ownership	• Secure (i.e., encrypted) storage with fine-granular option for users to grant access rights.	5
		Access to data	• Secure (i.e., encrypted) storage with fine-granular option for users to grant access rights	5
		Access to public data	N/A	1
		Usability	• Privacy by design approach. • Advanced user interfaces and dedicated applications available on a broad range of devices.	5
		Interoperability	• Integration of privacy-preserving aspects into existing identity and access management systems.	3
		Switch of CSPs	N/A	1
		Cloud certifications	N/A	1
		Cloud contracts (SLA)	• Development of a certification catalogue for privacy-preserving IAM and data sharing.	2
ESCUDO-CLOUD	**CS1**	**Free Movement of data**	Protection of the confidentiality and the authenticity of the stored data. Key management solutions.	
		Location of data		

		Ownership		5
		Access to data	Access control for regulating information sharing.	5
		Access to public data		
		Usability	Integrability of ESCUDO-CLOUD solutions with existing cloud solutions. Secure data deletion.	5
		Interoperability		
		Switch of CSPs		
		Cloud certifications		
		Cloud contracts (SLA)		
	CS2	**Free Movement of data**		
		Location of data		
		Ownership	Protection of the confidentiality of the stored data and computations	5
		Access to data	Access control for regulating information sharing. Fine-grained data access through encryption.	5
		Access to public data		
		Usability		
		Interoperability		
		Switch of CSPs		
		Cloud certifications		
		Cloud contracts (SLA)		
	CS3	**Free Movement of data**		
		Location of data		
		Ownership	Protection of the confidentiality of the stored data	5
		Access to data	Access control for regulating information sharing	5
		Access to public data		
		Usability		
		Interoperability	Multi cloud deployment	5
		Switch of CSPs		
		Cloud certifications		
		Cloud contracts (SLA)		
	CS4	**Free Movement of data**		
		Location of data		

		Ownership	Protection of the confidentiality and the authenticity of the stored data	5
		Access to data	Access control for regulating information sharing	
		Access to public data		
		Usability		
		Interoperability	Multi cloud deployment	5
		Switch of CSPs	Use of 3rd party services for protecting data	5
		Cloud certifications		
		Cloud contracts (SLA)	Security and privacy aspects in SLAs	
MUSA	**CS1**	**Free Movement of data**	• Continuous monitoring of SLA fulfilment.	5
		Location of data	• Design location-aware multi-cloud applications. • CSP selection based on their location. • Continuous monitoring of SLA fulfilment.	5
		Ownership	N/A	1
		Access to data	• Access control enforcement at component level. • Continuous monitoring of SLA fulfilment.	5
		Access to public data	N/A	1
		Usability	• DevOps and agile engineering. • Notifications to cloud consumer of SLA violations.	3
		Interoperability	• Interoperability of CSPs as CSP selection criteria. • Multi-cloud deployment.	3
		Switch of CSPs	• CSP Decision Support for multi-cloud based on security, performance and business aspects.	4
		Cloud certifications	• Continuous monitoring of SLA fulfilment.	4
		Cloud contracts (SLA)	• DevOps Risk Analysis. • Security and privacy-aware Composite SLA Generation. • Continuous monitoring of SLA fulfilment.	5
	CS2	**Free Movement of data**	• Continuous monitoring of SLA fulfilment.	5
		Location of data	• Design location-aware multi-cloud applications. • CSP selection based on their location. • Continuous monitoring of SLA fulfilment.	5

		Ownership	N/A	1
		Access to data	• Personal data access control enforcement. • Continuous monitoring of SLA fulfilment.	5
		Access to public data	• Composition of services available in ITS Factory.	1
		Usability	• DevOps and agile engineering. • Notifications to cloud consumer of SLA violations.	3
		Interoperability	• Interoperability of CSPs as CSP selection criteria. • Multi-cloud deployment.	3
		Switch of CSPs	• CSP Decision Support for multi-cloud based on security, performance and business aspects.	4
		Cloud certifications	• Continuous monitoring of SLA fulfilment.	4
		Cloud contracts (SLA)	• DevOps Risk Analysis. • Security and privacy-aware Composite SLA Generation. • Continuous monitoring of SLA fulfilment.	5
OPERANDO	**CS1**	**Free Movement of data**	• Continuous monitoring of changes in privacy settings by Online Service Providers • Anonymization tool	5
		Location of data	• Continuous monitoring of changes in privacy settings by Online Service Providers	5
		Ownership	N/A	5
		Access to data		5
		Access to public data		
		Usability		4
		Interoperability		
		Switch of CSPs		
		Cloud certifications		
		Cloud contracts (SLA)		
	CS2	**Free Movement of data**	• Continuous monitoring of changes in privacy settings by Online Service Providers • Anonymization tool	5
		Location of data	• Continuous monitoring of changes in privacy settings by Online Service Providers	5
		Ownership	N/A	5
		Access to data		5
		Access to public data		

		Usability		4
		Interoperability		
		Switch of CSPs		
		Cloud certifications		
		Cloud contracts (SLA)		
	CS3	**Free Movement of data**	• Continuous monitoring of changes in privacy settings by Online Service Providers • Anonymization tool • Regulator API which facilitates compliance with privacy laws as well as auditing and pro-active supervision of OSPs by privacy regulators	5
		Location of data	• Continuous monitoring of changes in privacy settings by Online Service Providers • Regulator API which facilitates compliance with privacy laws as well as auditing and pro-active supervision of OSPs by privacy regulators	5
		Ownership	• Privacy Policy Computation • Big data • Privacy for benefit • User device enforcement that supports delivery of privacy enforcing services by client applications. • Regulator API which facilitates compliance with privacy laws as well as auditing and pro-active supervision of OSPs by privacy regulators	5
		Access to data	• Rights Management • Authorization •Web Browser and Application: Identify Management	5
		Access to public data		
		Usability	• Dedicated web browser plug in and application	4
		Interoperability		
		Switch of CSPs		
		Cloud certifications		
		Cloud contracts (SLA)		
PaaSword	**CS1-CS2-CS3-CS4**	**Free Movement of data**	Secure storage incl access control	4
		Location of data	Secure storage incl access control	5

		Ownership	Secure storage incl access control and tenant control of encryption keys	5
		Access to data	Secure storage incl access control and tenant control of encryption keys	5
		Access to public data	N/A	1
		Usability	Holistic security by design, annotation framework and IDE plug in	4
		Interoperability	Multi-cloud deployment	2
		Switch of CSPs		1
		Cloud certifications		1
		Cloud contracts (SLA)		1
PRISMACLOUD	**CS1**	**Free Movement of data**	SAEaaS, VSaaS services in PRISMACLOUD platform	
		Location of data		
		Ownership	SAEaaS, VSaaS services in PRISMACLOUD platform	
		Access to data	SAEaaS, VSaaS services in PRISMACLOUD platform	
		Access to public data	BDAaaS service in PRISMACLOUD platform	
		Usability		
		Interoperability		
		Switch of CSPs		
		Cloud certifications		
		Cloud contracts (SLA)		
	CS2	**Free Movement of data**	DAaaS service in PRISMACLOUD platform	
		Location of data		
		Ownership		
		Access to data	DSaaS service in PRISMACLOUD platform	
		Access to public data		
		Usability	DSaaS service in PRISMACLOUD platform	
		Interoperability		
		Switch of CSPs	DSaaS service in PRISMACLOUD platform	
		Cloud certifications	DSaaS service in PRISMACLOUD platform	
		Cloud contracts (SLA)	DSaaS service in PRISMACLOUD platform	
	CS3	**Free Movement of data**	SAaaS service in PRISMACLOUD platform	

		Location of data	IAaaS service in PRISMACLOUD platform	
		Ownership		
		Access to data	SAaaS service in PRISMACLOUD platform	
		Access to public data		
		Usability	SAaaS service in PRISMACLOUD platform	
		Interoperability		
		Switch of CSPs	SAaaS service in PRISMACLOUD platform	
		Cloud certifications	IAaaS service in PRISMACLOUD platform	
		Cloud contracts (SLA)	SAaaS service in PRISMACLOUD platform	
SPECS	**CS1**	**Free Movement of data**		
		Location of data		
		Ownership		
		Access to data		
		Access to public data		
		Usability		
		Interoperability	Common model to express security SLA	5
		Switch of CSPs	Common model to express security SLA	5
		Cloud certifications	Common model to express and evaluate security SLA	5
		Cloud contracts (SLA)	Common model to express security SLA	5
	CS2	**Free Movement of data**		
		Location of data		
		Ownership	Tools to grant ownership and access to data	5
		Access to data	Tools to grant ownership and access to data	5
		Access to public data		
		Usability	Tools to grant interoperability and usability of data	5
		Interoperability	Tools to grant interoperability and usability of data	5
		Switch of CSPs		
		Cloud certifications		
		Cloud contracts (SLA)	Common model to express security SLA	5
	CS3	**Free Movement of data**		

		Location of data	Dedicated Security metric to grant and monitor data location	5
		Ownership	Tools to grant ownership and access to data	4
		Access to data	Tools to grant ownership and access to data	4
		Access to public data		
		Usability	Tools to grant interoperability and usability of data	4
		Interoperability	Tools to grant interoperability and usability of data	4
		Switch of CSPs		
		Cloud certifications	Common model to express security SLA	4
		Cloud contracts (SLA)	Common model to express security SLA	4
	CS4	**Free Movement of data**		
		Location of data		
		Ownership		
		Access to data		
		Access to public data		
		Usability		
		Interoperability		
		Switch of CSPs	Common model to compare CPS security	4
		Cloud certifications		
		Cloud contracts (SLA)	Common model to express security SLA	4
SUNFISH	**CS1**	**Free Movement of data**	Applying secure-by-design FaaS concept with integrated transformative controls	2
		Location of data	Applying secure-by-design FaaS concept with integrated advanced blockchain-based logging and anomaly detection	2
		Ownership	Framework for federated security policy specification, evaluation and distributed enforcement	1
		Access to data	Framework for federated security policy specification, evaluation and distributed enforcement and integrate transformational data controls	1
		Access to public data	Data transformation service using anonymization and data masking	3
		Usability	Web-based centralized administrative console	4
		Interoperability	Federation framework supporting different technologies and deployments	3
		Switch of CSPs	-	
		Cloud certifications	-	

		Cloud contracts (SLA)	Federated SLA monitoring	4
	CS2	**Free Movement of data**	Applying secure-by-design FaaS concept with integrated transformative controls	1
		Location of data	Applying secure-by-design FaaS concept with integrated advanced blockchain-based logging and anomaly detection	2
		Ownership	Framework for federated security policy specification, evaluation and distributed enforcement	2
		Access to data	Framework for federated security policy specification, evaluation and distributed enforcement and integrate transformational data controls	1
		Access to public data	Data transformation service using anonymization and data masking	5
		Usability	Web-based centralized administrative console	4
		Interoperability	Establishing federations using arbitrary cloud providers	3
		Switch of CSPs	-	
		Cloud certifications	-	
		Cloud contracts (SLA)	Federated SLA monitoring	5
	CS3	**Free Movement of data**	Applying secure-by-design FaaS concept with compartment-based security monitoring and enforcement	2
		Location of data	Integrated policy management with data-at-rest encryption	2
		Ownership	Integrated policy management with data-at-rest encryption	3
		Access to data	Integrated policy management with data-at-rest encryption	1
		Access to public data	-	
		Usability	Web-based centralized administrative console	3
		Interoperability	Data indexing and search framework with restricted visibility and anonymity controls	2
		Switch of CSPs	-	
		Cloud certifications	-	
		Cloud contracts (SLA)	Federated SLA monitoring	5
SWITCH	**CS1-CS2-CS3**	**Free Movement of data**	Secure storage in containers	3
		Location of data	Planning aware multi-cloud deployment	4
		Ownership	Use case owners	5
		Access to data	Access control at component level	5
		Access to public data	N/A	2
		Usability	IDE framework	4

		Interoperability	DevOps using multi-cloud approach	3
		Switch of CSPs	Knowledge base for multi-cloud using use case relevant metrics.	1
		Cloud certifications	Model to evaluate cloud SLA	2
		Cloud contracts (SLA)	Continuous monitoring of SLA fulfilment	5
TREDISEC	**CS1-CS2-CS3-CS4-CS5-CS6**	**Free Movement of data**		
		Location of data		
		Ownership	Verifiable Ownership	3
		Access to data	Access Control models for multi-tenancy Resource Isolation in Multi-Tenant Systems Data Provisioning Secure Enforcement of Policies in Clouds	5
		Access to public data	Optimizing Encryption for Data Outsourcing Privacy Preserving Primitives for Data Processing	3
		Usability		3
		Interoperability	Data Provisioning Optimizing Encryption for Data Outsourcing	3
		Switch of CSPs	Data Provisioning Optimizing Encryption for Data Outsourcing	3
		Cloud certifications	Processing Verifiability	3
		Cloud contracts (SLA)	Processing Verifiability	2
WITDOM	**CS1, CS2**	**Free Movement of data**	- Outsourcing of data to from trusted (private cloud) to untrusted environments (public cloud). - Protection of data based on combination of crypto and non-crypto tools: Anonymization, Secure signal processing, Secure computation, Integrity and consistency verification, Data masking and desensitization, and end-to-end encryption (E2EE). - Protection of data based on trade-offs between privacy and utility. Compliance with GDPR	4
		Location of data		1

		Ownership	- Compliance with GDPR - Access control and authentication enforcement Data protection based on combination of crypto and non-crypto tools	3
		Access to data	- Integration of privacy-preserving aspects into the solution. - Compliance with GDPR Access control and authentication enforcement	2
		Access to public data		1
		Usability	- Privacy and security by design approach. - SPACE methodology uses co-creation to foster end-users' participation in the solution design. Usability validation	4
		Interoperability	- Common model to express protection policies. - Common data model for scenarios. Agnostic from cloud provider.	4
		Switch of CSPs		1
		Cloud certifications	N/A	1
		Cloud contracts (SLA)	N/A	1

b. Projects' Methodological Results towards Free Flow of Data

In this section we provide a catalogue of formalisms, models, methods, techniques, procedures, standards, etc. resulting from the clustered projects that are already paving the path towards making Free Flow of Data a reality. The summary of this section is provided together with the summary of section c in a Table at the end of section c.

CLARUS

The initial version of the web-based guidelines and checklists for both CLARUS end users and service providers from legal, technical and standardization perspectives will be available on http://www.clarussecure.eu website in July 2017.

CLARUS provides a privacy-by-design approach to data processed and stored in the cloud. In relation to the free flow of data (FFD), data anonymization and data splitting are among the main innovations of CLARUS to enhance the privacy in cloud services. They significantly outperform standard cryptographic techniques in terms of efficiency, flexibility of operations and of data access, and utility for CSPs.

Coco Cloud

All Coco Cloud results are available at the http://coco-cloud.eu website. Most of the resources available in the website are rather technological than methodological.

CREDENTIAL

All CREDENTIAL results are available at the https://credential.eu website. As methodological results we include:

- Privacy requirements elicitation and monitoring methods.
- Adaption of advanced cryptographic schemes (proxy re-encryption, redactable signature schemes, etc.).
- Support of fine-granular and dynamic access control policies.
- Development of dedicated mobile apps, browser plugins, etc. to reduce the knowledge, understanding, and actions required by users to a minimum.
- End-user involvement at all stages of the development.
- Analysis and potential extension of existing IAM solutions like SAML to support the used cryptographic primitives.
- Transfer of project results into a dedicated new certification catalogue of the StarAudit cloud certification scheme.

CloudWatch2

All CloudWatch2 results regarding the transfer of personal data by means of the use of cloud providers, together with other legal recommendations for SMEs and Public Administrations (PAs), are available at the following link: http://www.cloudwatchhub.eu/legal-services

ESCUDO-CLOUD

All ESCUDO-CLOUD results are available at the www.escudocloud.eu website.

MUSA

All MUSA results are open access and will be available at the www.musa-project.eu website.

- Extensions to CAMEL language for better addressing multi-cloud security and deployment requirements.
- DevOps oriented Risk Analysis methodology.
- Cloud Service Providers (CSP) selection supporting mechanisms.
- Security and privacy-aware SLA model that includes SLOs for security and privacy, expressing security controls and security metrics.
- Cloud Security metrics Catalogue.
- Multi-cloud Threat Catalogue.
- Composition rules for creation of composite SLA that take into account component level and overall level SLOs.
- Multi-cloud deployment model.
- Continuous monitoring techniques for composite SLA fulfilment assurance.
- Security enforcement mechanisms for multi-cloud (such as access control and scalability).

OPERANDO

All OPERANDO results are available at the project website http://www.operando.eu/servizi/notizie/notizie_homepage.aspx.

- Privacy Authority Architecture.
- Privacy Enforcement Platform for G2C, including policy computation algorithms, anonymization, privacy watchdog and APIs for regulatory aspects.
- PlusPrivacy.com and PlusPrivacy App, for B2C.

PaaSword

All PaaSword results are available at the www.paasword.eu website.

- The PaaSword Reference Architecture aims to satisfy the different types of requirements following a use-case driven approach. The overall goal was to identify all stakeholders and as many as possible functionalities that would be required towards the formulation of a secure PaaS framework.

- PaaSword Context-aware Security Model
    - PaaSword Context Element
    - PaaSword Permission Context Element
    - PaaSword Context Pattern Element
    - PaaSword Context DDE (Data Distribution and Encryption) Element
- PaaSword Security Policy Models, where three main types of security policies are considered:
    - Data encryption policies
    - Data fragmentation and distribution policies
    - Access control policies
    - Secure and Searchable Encryption
        - transparent and secure symmetric encryption of outbound sensitive data,
        - data distribution to different server or clouds to limit the knowledge about metadata.

PRISMACLOUD

All methodological results of PRISMACLOUD will be published on the project homepage https://prismacloud.eu. In particular, the project is putting forward and extending security and privacy by design methods for cryptographic service design and applies these methods on all layers of the PRISMACLOUD architecture. All technological developments will be fully documented to facilitate later reuse and are accompanied with usage guidelines and privacy patterns.

- PRISMACLOUD Architecture: A new reference architecture comprising four layers: cryptographic primitives and protocols, tools layer, services layer, applications layer. The tools layer completely encapsulates the cryptographic functionality;
- A new Cryptographic Service Development Lifecycle (CryptSDL), an extension of secure software development lifecycles based on the PRISMACLOUD architecture which helps to deal with the complexity involved in the integration of cryptographic protocols;
- PRISMACLOUD Toolbox: A specification of flexible and reusable cryptographic tools for security and privacy friendly cloud services.
- PRISMACLOUD Service modelling: Security and privacy SLA models for the developed PRISMACLOUD services
- Security and privacy patterns as well as HCI recommendations for the development of cryptographic services

SPECS

All SPECS results are available on the project website http://www.specs-project.eu and on the SPECS source code repository https://bitbucket.org/specs-team/. In particular the project proposes a Security SLA Model that can be used to automate Security SLA Management and delivery of services based on Security SLAs. The main methodological results are:

- SPECS Security SLA Model: A WS-Agreement extension to address security in Service Level Agreements, using standard security control frameworks (NIST SP-800-53, CSA CCM);
- A machine readable representation of security SLA
- SPECS Security Metric Catalogue: A catalogue of security metrics and their link to standard security controls in order to enable measurement and Monitoring of security properties in SLAs;
- Security evaluation techniques
- Monitoring Policy (MoniPoli) for SLA-related events definition;
- Innovative approaches to manage the SLA remediation

SUNFISH

The main methodological result is the following:

- SUNFISH federation architecture, governance model and processes.

It will be published on the project webpage http://sunfishproject.eu.

SWITCH

All SWITCH results (papers, deliverables, Git source code) are available at the http://www.switch-project.eu website.

TREDISEC

The **Security Primitives** describe a security solution in a cloud system together with performance capabilities. Due to the constraints and different characteristics that this artefact has to provide, represent and the different implementations it can have (together with the specific configurations) we have designed an architecture that covers all the requirements defined in the project. There are different phases associated to the security primitives from their initial design and definition till they are deployed into a cloud system: the so-called Security Primitives Lifecycle. Two roles interact mainly in this lifecycle: the security expert engineer and the security technology provider.

The **TREDISEC Framework** is a component that allows the creation, use, management and deployment of security primitives in a target cloud. It provides an online packaging of security primitives to be used by the different roles identified (End-User, Security Expert Engineer, TREDISEC Framework Administrator and the Security Technology Providers) together with tools for specific functionalities (e.g. user interface for managing it, testing and deployment component for testing the security primitives and do their deployment, etc.).

The framework offers three operational modes: development, maintenance and provisioning. The first covers the design, development and testing of the security primitives along its lifecycle (from security primitive pattern to TREDISEC Recipe), the maintenance mode covers the functionality and lifecycle of the update, refining, extension, etc. of the different artefacts of TREDISEC. Finally, the deployment phase covers the functionality and lifecycle of the TREDISEC Recipes and their deployment into the target cloud.

WITDOM

All methodological results of WITDOM will be published on the project homepage: http://www.witdom.eu.

In particular it is worth mentioning the following:

- The **SPACE methodology** is the combination of PRIPARE[29], a privacy and security by design methodology, with co-creation, methodologies, that follow a collaborative and multi-disciplinary approach, involving the users in the creation process.
- The **WITDOM Privacy framework** is a holistic privacy framework that comprises a set of guidelines that allow developing complete privacy-by-design architecture. It is based on four steps:
    - Identify the scenario high level privacy requirements (to be carried out using the SPACE methodology).
    - Mapping high-level requirements to privacy properties (e.g., anonymity, unlinkability, cryptographic confidentiality, etc.). A dialogue with the stakeholders from the scenarios is essential to successfully carry out this translation.
    - Establish which metric or metrics will be used to measure to which extent the privacy properties are fulfilled (i.e. entropy, anonymity sets…).
    - Choose a threshold for this metric for which the stakeholder considers that privacy is preserved.
- The **WITDOM architecture** is a generic architecture for processing data in untrusted domains (mainly the public cloud).

[29] http://pripareproject.eu/

c. Projects' Technological Results towards Free Flow of Data

In this section we provide a catalogue of tools, frameworks, platforms, technologies, etc. available or under work in the clustered projects that address the issues identified in Section a.

CLARUS

All CLARUS results are available at the http://www.clarussecure.eu website.

CLARUS solutions are built on standards to provide a solution as general as possible. The wide spectrum of solutions provides CLARUS with the ability to cope with diverse needs regarding security, efficiency, functionality, access, interoperability, etc., and with different scenarios, such as standalone users, collaboration between users located at different companies, data spread through different CSPs, among others.

- Privacy-preserving mechanisms for proper protection of sensitive and personal data outsourced to the cloud, innovating over the current state of the art. Data operations include:
    - Two types of anonymization mechanisms have been designed:
        - *Data coarsening*: it systematically generalizes input records (independently, *one at a time*) according to a user-defined coarsening level. Since coarsened data are less detailed than the original ones, disclosure risk is minimized.
        - *Data microaggregation*: it clusters a fixed number *k* of similar records together and replaces them with average values; thus, it transforms the whole dataset in a *monolithic, global* way (it cannot be applied independently to each record). Since the *k* microaggregated records within each cluster are indistinguishable, the re-identification probability is lowered to *1/k*.
    - **Data splitting** makes a local partition of the sensitive data and separately stores data fragments in different CSPs, in a way that each individual fragment does not cause privacy risks; data fragments are stored in the clear without any modification; thus preserving data accuracy and the analytical interest. Only CLARUS knows the exact cloud locations for a given dataset.
    - **Data encryption** is a method to protect data in a secure and reversible way. Encryption is performed by CLARUS once at the storage stage, and the decryption is performed after recovering the encrypted data from the CSP. The keys are stored at the proxy. In this way, the CSP never has access to the plaintext data or to the keys.
    - **Homomorphic encryption** is used to store data in a secure way that allows performing certain computations directly on the encrypted data. The encryption scheme is a public-key one, and the secret key is stored in CLARUS. The secret key is used to decrypt data. The encrypted data can only be decrypted by the users owning the secret key.
    - **Searchable encryption** is used to store data in a secure way that allows performing queries on the encrypted data. The encryption is reversible. Encryption is performed by CLARUS once at the storage stage, and the

decryption is performed after recovering the encrypted data from the CSP. The user can perform queries on the encrypted data. The answer of the query is a link to the encrypted document containing the keywords in the query.

- Development of a **secure and attack-tolerant framework** for the storage and processing of data outsourced to the cloud, enabling users to monitor, audit, and control the stored data without impairing functionality, including functionality provided by high-level services such as data storage, management, retrieval, transformation, as well as cost-saving benefits of cloud services. Attack-tolerant framework has a variety of security mechanisms under the control of cloud users without reducing benefits of a cloud service. Very few of the current solutions have the capability of managing intrusions and attacks. Very few provide countermeasures to protect the system and guarantee expected behaviour in a hostile context.
- To enhance privacy, security and trust vis-à-vis CSPs, the location of the **CLARUS proxy** is in a domain trusted by the end user, (e.g., a server in her/his company's intranet or a plug-in in the user's device) that implements security and privacy-enabling features towards the CSP. The aim is to have an extensible technology that can be configured by the end user based on the type of data and type of policy required.
- CLARUS also provides monitoring; access control (storage, request and search); verifiability.

Coco Cloud

The Coco Cloud architecture defines two main subsystems: the Coco Cloud Engine and the DSA Subsystem, which together form the reference Coco Cloud platform.

Coco Cloud Engine: The Coco Cloud Engine is the component responsible for the enforcement of the policies to be applied to the protected data (defined in its corresponding DSA) and for the publishing (storing in a secure way) of Coco Cloud Objects (CCOs). The Coco Cloud Engine builds on the following components:

- Coco Cloud API: this component provides the functionalities of Coco Cloud to external applications (Coco Cloud-aware). The API communicates internally with the rest of elements of the Coco Cloud Engine in a transparent way.
- Enforcement Subsystem: this component is in charge of evaluating the requests and usage of the Coco Cloud data against the policies defined in its corresponding DSA. It provides its functionality for both Cloud-based and mobile systems.
- Publishing Service: it is in charge of creating Coco Cloud Objects using as input the data and the DSA that specifies its security and privacy policies. It communicates internally with the DSA Subsystem, Enforcement Subsystem, and plugins (e.g. Key and Encryption Manager).
- Storage Adapter: this component is in charge of providing access to the storage (and accessing) of the Coco Cloud objects (generated by the Publishing Service), by abstracting the specific Cloud storage provider implementation.

- Plugin Adapter: the adapter defines the abstract interface for the integration of the different plugins that can be used in Coco Cloud. We currently use this component for exploiting the features of the Identity Management, the Encryption & Key Management, Audit Management and the Integrity Management.

DSA Subsystem: The DSA Subsystem is the component responsible for creating and managing the DSAs, i.e. the policies to be attached to data created under Coco Cloud. The DSA Subsystem is composed of the following components:

- DSA Authoring Tool: this component creates and manages DSAs by providing a framework for defining the access and usage policies using a Controlled Natural Language (CNL). This language is created using as basis ontologies of specific domains.
- DSA Analyser: this component is responsible of checking that the policies rules defined in the DSA are correct and have no conflicts between them. This component provides information to the policy authors about how to better define the rules so they have no conflicts or can be better processed.
- DSA Mapper: the DSA Mapper tool transforms the policies from Controlled Natural Language to the enforceable language UPOL (Usage POLicies). This transformation can be either carried out after using the DSA Analyser in order to obtain information of the conflicts, or directly as it can apply semi-automatic conflicts solving strategies.
- DSA Repository: this component stores the DSAs to be used by the different services and elements of Coco Cloud. It is accessed and managed via the DSA API.
- DSA Lifecycle Manager: the DSA Lifecycle Manager is in charge of managing and communicating all the tools and components that compose the DSA Subsystem (explained previously: DSA Authoring Tool, DSA Analyser, and DSA Mapper). This component provides an independent platform that can be used by three different roles for specific goals and functionalities.

CREDENTIAL

All CREDENTIAL results are available at the https://credential.eu website. As technological results we include:

- Cryptographic primitives such as attribute-based credentials on encrypted data.
- CREDENTIAL generic and pilots' specific mobile apps.
- Cloud certification catalogue.

ESCUDO-CLOUD

All ESCUDO-CLOUD results are available at the www.escudocloud.eu website.

ESCUDO-CLOUD Framework:

- Provides *modular family* of solutions for securing outsourced data.

- Supports *multi-dimensional* security goals.
- *Multiple systems* that complement each other.
- Combination through use of *open standards* and APIs.
- *Open-source contributions* of selected results.

MUSA

All tools in MUSA are currently prototypes under work and the first tested version will be available at the www.musa-project.eu website, under Tools menu. All tools are open source and will be available in Bitbucket.

- **MUSA Integrated framework** for the DevOps and agile based engineering and operation of the multi-cloud application taking into account the security requirements of the application and the security offerings of the cloud providers to use. The integrated framework is a Kanban style IDE that allows a DevOps team to work with the different tools for designing, deploying and operating (monitoring and enforcement) multi-cloud applications.

  The parts of the framework, which also work as independent tools are:
- **MUSA IDE Modeller**: Web Modeller of multi-cloud applications (based on CAMEL language) that allows the specification of the deployment and security requirements.
- **MUSA Risk Analysis tool** (DevOps oriented) that allows identifying and ranking the risks of the multi-cloud application components and overall application.
- **MUSA DST - Cloud Service Providers (CSP) selection Decision Support Tool** for multi-cloud, including security features in selection process.
- **MUSA Security and privacy-aware SLA Generator**. This tool allows the creation of the composite SLA for the multi-cloud application based on the SLAs of the individual components and the SLAs of the CSPs being used by the components.
- **MUSA Multi-cloud Deployer** which includes a GUI for refining the deployment plan.
- **MUSA Security Assurance Platform** for continuous runtime monitoring of composite SLA fulfilment and security features enforcement. The Platform will be provided as a SaaS in pay-per-use model.
- **MUSA Security monitoring agents** for multi-event monitoring in distributed applications. These agents are the probes used by the MUSA Security Assurance Platform but could be adopted in other platforms.
- **MUSA Security enforcement agents** (such as for access control and scalability) actionable at runtime operation. These agents are the enforcement mechanisms used by the MUSA Security Assurance Platform but could be integrated with multi-cloud applications independently of the MUSA Platform if needed.

OPERANDO

All results / components of the OPERANDO platform, Web browser and mobile application are open source and distributed mostly under the MIT license. They are available at: https://github.com/OPERANDOH2020

- Integrated Platform to implement the "Privacy Authority" concept including components for Rights Management, authorization, anonymization, big data analytics, and so on;
- Monitoring of changes in the privacy settings of OSPs via a Privacy Watchdog;
- Provision of a policy computation engine, which acts as decision support engine for providing privacy aware services.
- Web Browser plug-ins for enforcing anti-trackers, manage identities, protect data leakage and scan reputation;
- Mobile Application supporting the services to be delivered by the privacy service provider, including an application permission scanner, a reputation scanner, and the corresponding identity manager;
- Several APIs, such as the Regulator API, which facilitates compliance with privacy laws as well as auditing and pro-active supervision of OSPs by privacy regulators.

PaaSword

The framework and tools from PaaSword are currently prototypes under work and the first tested version will be available at the www.paasword.eu website. The PaaSword framework and corresponding tools will be provided non-commercially under an open-source license, which is expected to facilitate their wider adoption by the industrial and the ICT communities.

- **PaaSword Framework** is designed to help the PaaSword Administrator, the Product Manager or even the Cloud Application developer instantiate and customize the semantic background (i.e. PaaSword Context Model (CM)) based on which PaaSword Access Policies can be defined. The instantiation and customization of the semantic background is performed using the web-based Context Model Editor. (www.paasword.eu/results/paasword-context-model-editor/).
- **PaaSword Annotations Governance And Validity Control Mechanism (AGVC)**. The purpose of the Annotations Governance and Validity Control (AGVC) mechanism is to deliver the following main kinds of functionality:
    - Determination of policy well-formedness.
    - Policy lifecycle management.
    - Evaluation of policy validity.
- **PaaSword Annotation Interpretation Mechanism**. Three annotation types have been defined that serve a specific functional purpose under the PaaSword framework. These XACML based annotations are:
    - @PaaSwordPEP annotation
    - @PaaSwordDDE annotation

  - @PaaSwordEntity annotation

  The PaaSword Annotation Interpretation Mechanism is used to efficiently interpret the annotations into XACML-based enforceable Access Control Policies.

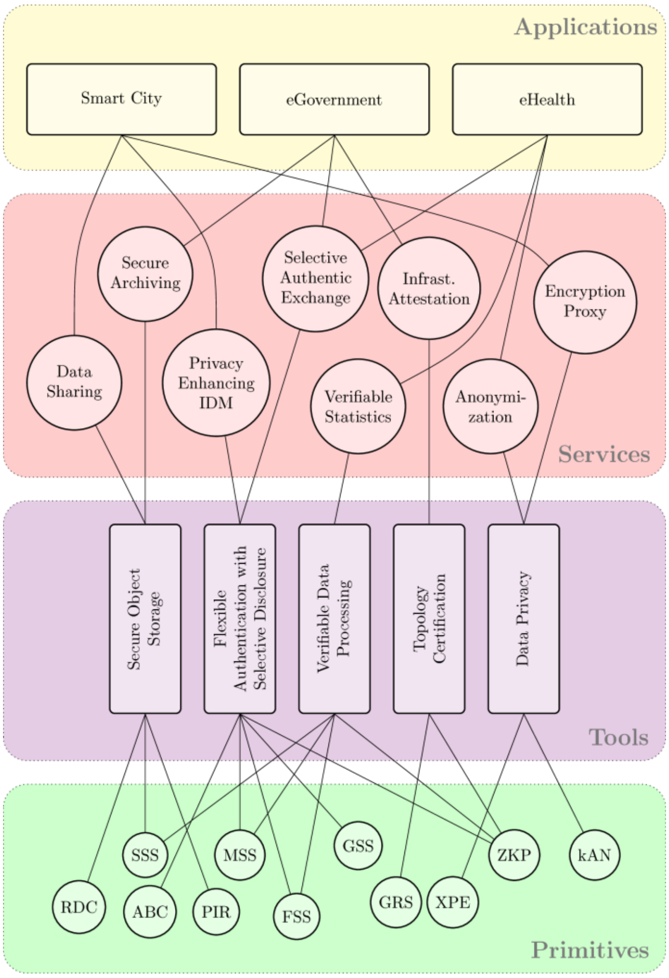
- **PaaSword Distributed Key Management Mechanism**. If database encryption is set, data may only be stored encrypted to prevent unauthorized access. Authorized access should only be possible on behalf of the data owner, the tenant. To share the data of a tenant within its employees you encrypt the database with one tenant key (TK) which should be solely under the control of the tenant.

  The PaaSword Key Management Mechanism is responsible for preparation, distribution and usage of keys, which enable database access. The PaaSword approach is based on an architecture that separates the Application (A) where the data is processed from the DB-Proxy (P) whose task is to store and access the data in a cloud database. The key is distributed among the User (Ui), Application (A) and DB-Proxy (P), who store solely their individual part of the key (TKui, TKai, TKpi). To recompute the key, needed for database access during runtime, all three parts (from Ui, A, P) are needed.

- **PaaSword Database Proxy**. It resides between the application and the untrusted cloud to secure outgoing sensitive data before it is uploaded to the cloud. It is designed to work transparently to the application and user. The Proxy handles:
  - transparent and secure symmetric encryption of outbound sensitive data,
  - decryption of inbound data for further processing,
  - automatic data distribution to different server or clouds to limit the knowledge about metadata,
  - handling of the reverse index for fast determination which data is relevant to the query and support of sublinear search time,
  - automatic query rewriting to deal with encryption and data distribution.

PRISMACLOUD

The technological results of PRISMACLOUD are grouped according the architecture, which is organized in 4 tiers as shown below. The architecture distinguishes between *Primitives*, *Tools*, *Services* and *Applications* whereby the latter three are generate technological results. A tool can therefore be regarded as an abstract concept or piece of software, e.g., a library, which is composed of various cryptographic primitives and which can be parametrized in various different ways. From the tools of the toolbox, the services of the next layer are built. A service is seen as a customization of a particular tool for one specific application. It is a way to deliver the tool to system and application developers, the users of the tools, in a pre-configured and accessible way. The applications developed in PRISMACLOUD are based on the case studies provided (see CS1 to CS3 for FFD relevant applications) and provide a way to validate the research output on the tools and services in real world applications.



In PRISMACLOUD we have chosen to specify a selection of services which we will develop during the project and which are suitable for showcasing the suitability of the chosen primitives and the tools constructed from them within the selected use cases. In particular, we will specify, implement and demonstrate 5 different tools and 8 specific services within the project. The services developed are the following:

- Data Sharing Service (DSaaS): The data sharing service allows multiple parties to securely store data in a cloud-of-clouds network such that no single storage node learns plaintext data, while still enabling the owner to share the data with other users of the system, i.e., the data sharing service supports secure collaboration without the need to trust one single storage provider.
- Secure Archiving Service (SAaaS): This service is a generic infrastructure service which can easily be integrated into cloud based backup scenarios while providing a demonstrable higher level of data privacy and availability than current cloud-based archiving solutions.
- Selective Authentic Exchange Service (SAEaaS): This service enables users to move their authentic documents to a cloud service and then delegate the selective sharing of parts of these documents to another party, while maintaining the authenticity of the selected parts. The other party can then verify the authenticity of the received data.

- Privacy Enhancing IDM Service (PIDMaaS): This service offers the capability of a privacy enhanced identity management. In particular, it allows users to store their attribute credentials obtained from some entity (e.g, a service provider or an authority) in this component and to realize a selective attribute disclosure functionality.
- Verifiable Statistics Service (VSaaS): This service provides the functionality to delegate the computation of verifiable statistics on authenticated data in a secure way. The computations have the feature of being public verifiability, i.e., any verifier can check whether an outsourced computation has been performed correctly, or not.
- Infrastructure Auditing Service (IAaaS): The infrastructure auditing service offers the capability to certify and prove properties of the topology of a cloud infrastructure without disclosing sensitive information about the actual infrastructure's blueprint.
- Encryption Proxy Service (EPaaS): The service supports moving legacy applications to the cloud by encrypting sensitive information identified within
- HTTP traffic in a format and/or order preserving way.
- Anonymization Service (BDAaaS): This service enables users to anonymize large data sets, and in particular database tables. The service allows users to identify private and sensitive information in the data sets and produce an anonymized version of the data set.

The services are based on following tools from the cryptographic toolbox:

- Secure Object Storage Tool (SECOSTOR)
- Flexible Authentication with Selective Disclosure Tool (FLEXAUTH)
- Verifiable Data Processing Tool (VERIDAP)
- Topology Certification Tool (TOPOCERT)
- Data Privacy Tool (DATPRIV)

Besides the basic building blocks, the case studies are also realized as applications which will be demonstrated in the pilots. Therefore, piloting applications are also technological output and provide first versions of exploitable products for PRISMACLOUD.

SPECS

All SPECS results are available on the project website http://www.specs-project.eu and on the SPECS source code repository https://bitbucket.org/specs-team/. In particular the following technological outputs can be reused:

- SPECS Metric Catalogue Application: A catalogue of security metrics and their link to standard security controls in order to enable measurement and Monitoring of security properties in SLAs.
- SPECS Framework: A framework to develop cloud applications able to provide services granted by Security SLAs, including all the phases of the SLA life cycle (Negotiation, Implementation, Monitoring, Reaction, Renegotiation);
- SLA Platform and its API to manage the SLA life cycle and support the SPECS Framework;

- SPECS Secure Web Container: An example of SPECS application able to deliver secure web servers according to negotiated SLAs.
- SPECS End-to-End encryption: A SPECS application able to offer Data-as-a-Service, protected with end-to-end encryption techniques and able to guarantee security SLAs
- SPECS ViPR Enhancement: An integration of SPECS and ViPR (or its opensource version CopperHead) able to deliver storage-as-a-service in next generation data centers according to Security SLAs
- SPECS Reasoners: a set of tools able to compare and reason over Security SLA, comparing quantitatively CSP offerings in terms of security, reusing informations like the ones offered by CSA in STAR registry and/or STARWatch.

SUNFISH

All results are open access and will be available at the http://sunfishproject.eu website. They include:

- FaaS – Federation as a Service, a governance model and framework implementing secure-by-design model to establish and manage cloud federations.
- A holistic security management framework that integrates preventive and reactive mechanisms with the goal to achieve an advanced and context-sensitive definition and enforcement of security requirements for cross-organisational data sharing, service and process integrations.
- Federation architecture, governance model and processes.
- Integrated security management of Web-API integration processes supporting semantic controls, high level of expressiveness and interoperable (cross-entity) policies that allow high level of abstraction and granularity.
- Out-of-the-box solution aimed at public administrations and federation of private cloud and public cloud services.

Federation components:

- Federated Administration and Monitoring: UI-based entry point for administrators that allows following functionalities:
    - Graphical entry-points for cloud federation administration functionalities.
    - Reporting on SLA policy violations.
    - Collection and management of security alerts received by the FRM and FSA components.
    - Visualization of security alerts by means of an integrated dashboard.
    - Subscription to federated services for service consumer.
- Identity Management: a component that allows the integration of different organisational identity management components and processes for the purpose of definition and enforcement of federation-level security policies. Allows the following functionalities:
    - Providing an identity to any service consumer and provider.

  - Defining a SSO authentication mechanism.
  - Enabling the federation of the identity managers of the individual clouds.
  - Providing endpoints for the generation of authenticating crypto-tokens.
  - Defining an identity management compliant with eIDAS.
- Registry Interface (RI) is a component that manages storage and retrieval operations of the blockchain-based registry. Supports the following functionalities:
  - Offering a set of APIs to store and retrieve the governance data to and from the blockchain-based registry.
  - Defining authorization controls on the API invocation based on crypto-tokens.
- Data security component supporting Attribute-based access control and enforcing the access policies associated with federated services. This component consists of a range of sub-components deployed, providing the following functionality:
  - Supporting the evaluation and distributed enforcement of ABAC policies.
  - Defining invocation mechanisms for the DTS components.
  - Defining access controls for different operational processes in the federation.
  - Defining administrative controls on the modification actions on access control policies.
- Data transformation service (DTS) that includes a range of components seamlessly integrated in the framework, supporting real-time data transformation on the fly. Includes SMC, Data masking and anonymization components.
- Secure multi-party computation (SMC) is one of data transformation service (DTS) components that allows secure computation without revealing information about private inputs.
- Data Masking, another DTS component that provides generic service to selectively masking personal and sensitive information. It supports following functionalities:
  - Providing the processes for masking and unmasking of personal and sensitive data.
  - Integrating the (un)masking processes with the SUNFISH platform.
  - Managing the masking service state (tokenization table) via the RI.
- Anonymization component supports k-anonymity and differential privacy, provides through two complementary services:
  - Micro data anonymization.
  - Macro data anonymization.
- Intelligent Workload Manager (IWM) is a component acting as the service broker in the federation, based on following functionality goals:
  - Providing the computation of optimised federation-based workload deployment targets.
  - Supporting different optimisation parameters for the calculation of the workload deployment targets.
  - Providing support for interaction with APIs of the federated clouds, both private and public.
  - Integrating DS policies with the optimisation path to filter out requestor-specific targets.

- Federated Runtime Monitoring component allows a distributed infrastructure to intercept access control requests, responses and related events on the level of SUNFISH infrastructure with the goal to detect policy violations
    - Providing distributed probes to monitor the access control system.
    - Detecting access control policy violations by analysing the collected access control data.
    - Raising alerts to the FAM to signal access control violations.
    - Providing to the FSA the access logs to perform its reasoning tasks.
- Federated Security Audit component implements automated detection service against vulnerabilities in distributed access control system and against potential security breaches occurring in the federation. It is based on data aggregation and role mining to support retrieval, analysis and transformation of low-level activities into higher-level business transactions that are easier to audit and monitor.

SWITCH

All results are available at the http://www.switch-project.eu website. The workbench has three subsystems:

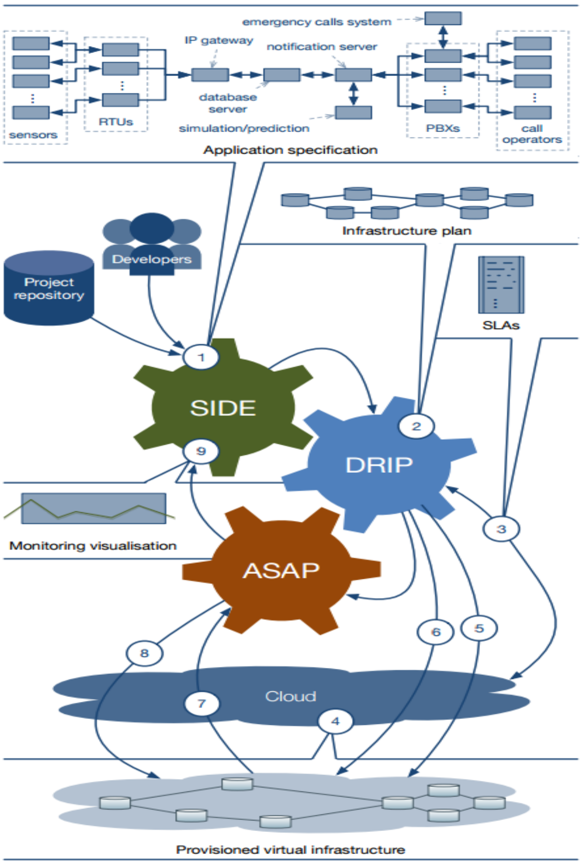
1. The SWITCH Interactive Development Environment (SIDE), to specify applications for deployment on Cloud.

2. The Dynamic Real-time Infrastructure Planner (DRIP), to plan and provision applications on virtual infrastructure.

3. The Autonomous System Adaptation Platform (ASAP), to monitor and intercede in the execution of applications.

The modularity of SWITCH allows components to be replaced as new Cloud standards and technologies come into existence.

The SWITCH application lifecycle is split into a number of interlinked phases, as presented in the figure below:



The following phases are followed during the lifecycle of an application:

1. Application composition and verification.
2. Resource selection and infrastructure planning.
3. SLA negotiation.
4. Infrastructure provisioning.
5. Application deployment.
6. Application execution and runtime management.
7. Runtime monitoring and diagnosis.
8. Runtime adaptation.
9. Runtime visualisation and feedback.

The SWITCH project will make an impact on:

- Improving the development productivity of time critical Cloud applications;
- Upgrading industrial technologies of time critical applications to use Cloud infrastructure;
- Improving deployment efficiency of time critical applications;
- Reducing operational cost of time critical services;

- Promoting business competitiveness of Clouds.

TREDISEC

The **TREDISEC Framework**, as aforementioned provides the functionalities for creating, managing, testing and deploying security primitives. All these functionalities are supported by the different components that compose the Framework.

The TREDISEC Framework is composed of different components that are responsible of the specific functionalities described above. Following, we describe each component along with the interactions with the rest of the framework:

- UI: this component represents the elements (e.g. user interface, API, etc.) and functionalities for interacting and managing the security primitives of TREDISEC. The UI provides specific functionalities according to the role that is using it, achieving separation of responsibilities that improves and facilitates the work with the TREDISEC Framework as each role will have defined very clearly its actions. The UI interacts with the rest of components of the TREDISEC Framework as each role interact one way or the other with them. Therefore, the UI is the main entry point for the different roles of TREDISEC (explained in Section 4.2) to interact and work with the security primitives (e.g. creating, logging, deploying, testing, etc.). Finally, the UI will allow the different roles to configure TREDISEC recipes and instances, to build and deploy them, etc.
- Security Primitives Component: this component is responsible for creating and managing the security primitives of TREDISEC. The component is composed of APIs and an interface that makes available all the functionalities for working with the security primitives and accessing them. The repository contains the security primitives (which can be in different phases or states according to their level of completeness or validation state) and provides artefacts (e.g. API) for managing and working with them.
- Testing Component: the testing component is in charge of providing a testing environment for security technology providers to check the functionality of their security solutions into the cloud-specific domain it is created for. That way they can analyse and identify the problems of their solutions and check if it fulfils the requirements and necessities specified in the security primitive in the specific cloud environment where it is applied. This component will also provide logs and other type of information (e.g. metrics, performance, etc.)
- TREDISEC Recipes Component: it contains and provides functionalities and methods for working with the TREDISEC Recipes. The recipes are pre-defined deployment packages that have been tested and validated in a specific cloud infrastructure. A recipe can have one or more security primitives (specified at design time so the implementation by the security technology provider covers their expected functionality) and also includes configuration information for the part of the deployment.

A catalogue of **Security Primitives** implemented in the project is available in the project website: http://www.tredisec.eu/primitives-catalogue

The catalogue includes the following primitives:

- Secure file deduplication
- Processing Verifiability
- Privacy-preserving data processing and word search
- Data provisioning and optimized encryption
- Data confidentiality & Deduplication
- Storage integrity with Proof of Retrievability (PoR)
- Attack surface reduction
- Fuzz testing
- Proof of Ownership
- Key Management
- Resource Isolation (container privacy)
- Remote attestation for Platform Integrity
- Deduplication, replication and PoR
- Secure Enforcement of Policies
- Secure Deletion in the Cloud

WITDOM

WITDOM framework relies on a flexible end-to-end secure architecture which forms the basis for securing use-case applications. The WITDOM architecture uses a **service-orientation design paradigm** (SOA), isolating applications from changes in the architecture, such as changes in implementations or locations of elements. The services provided by the components of the architecture are abstract, reusable, only loosely coupled, and can be composed with each other.

The framework is **secure-by-design** and covers distributed processing and data storage in order to raise the level of data privacy and security with respect to the levels offered by the respective service providers.

Services provided to end-users may run in trusted or untrusted domains and can range from infrastructure type services, such as data storage, to sophisticated methods that are specific to use-cases. Business-related applications are always hosted in a trusted domain, while other application functionalities are outsourced to the untrusted domain in order to benefit from better utilization of resources. Since they are located in an untrusted environment, the service providers that run those functionalities use data protected by so-called **protection components.**

Besides service-specific elements, the components of the architecture are composed of component of generic nature which serve as the basis for the secure functioning of the architecture, and specific protection components, which are the main features of WITDOM and serve to protect data before they are sent to untrusted domains.

- **Protection Orchestrator** (PO): Responsible for interpreting a protection configuration (BBMN xml file) that is defined according to the results of WITDOM privacy framework and which invokes the different services and protection components in the right order and with the adequate parameters to achieve the required level of privacy.
- **Broker:** is the component responsible for keeping the trail of WITDOM's infrastructure (where are all the components deployed) and to forward all external and internal calls to the adequate services. One of the key points of this component is that it can trigger some special data operations (protect and unprotect) whenever it decides or detect that the invoked service will be executed in an untrusted environment.
- **Protection components:** Each of these protection components shares the purpose of applying some technique to the data in order to achieve some level of protection.
    - Secure Signal Processing (SSP): Applies some pre-processing to the data before its outsourcing, e.g. encryption, secret sharing and data splitting, generation of garbled circuits). After its protection, the encrypted/obfuscated data is outsourced to one or more clouds (assumed to be non-colluding) where the processing takes place. This processing may be interactive, where the cloud have to collaborate with the data outsourcer or other clouds to perform it or non-interactive where each cloud can independently operate over the data. Once the cloud operations over the data are performed, some extra processing must take place at the client side in order to join and/or unprotect the results.
    - Secure Computation (SC): Uses homomorphic encryption techniques that allow performing some computations over encrypted data. The end user encrypts data using his private key and outsources this data to the cloud, indicating some operation to be performed. The result of this computation in the cloud is still in an encrypted form that can be only decrypted by the owner of the original private key.
    - Anonymization: These techniques specifically focus on privacy and do not guarantee confidentiality or integrity. The main idea is using statistic metrics and some very basic methods apply the methods until some privacy metric criteria are met. Examples of methods are:
        - Generalization: i.e. remove the level of granularity of the data (rounding amounts, grouping data in ranges)
        - Noise addition: i.e. add or subtract values following some statistic symmetric distribution will modify individual values but would not affect the mean value of the data
        - Attribute removal: directly remove one attribute that does not have too much utility but may help to identify data
        - One way functions: Apply a hash function to unique identifiers (e.g. Social Security Number)
    - Data Masking (DM): The main idea is to use a key-based one way function to protect identifiers that are necessary to maintain the integrity of references when outsourcing data.

- Integrity and Consistency Verification (ICV): The ICV protocol implemented and deployed in any cloud will immediately detect any discrepancy related to the outsourced data (e.g. modifications not applied by the cloud, presentation of outdated version, etc.)
- End-2-End Encryption (E2EE): It provides encryption and integrity to outsourced data. It requires a trusted third party to monitor the data and metadata in order to detect anomalies (i.e. tampering)

- **Transformation Services:** Simple services that transform data between domain-specific formats and WITDOM's common tabular format.

The following table summarises the methodological and technological outcomes of the clustered projects that already are solutions addressing free flow of data issues.

Table 4c. Methodologies and technologies addressing FFD working areas

Project	GDPR Issue/FFD working area	Methodology	Technology (supporting tool)
CLARUS	Free Movement of data	No CSP lock-in is a key requirement in the definition of the CLARUS solution. Use cases and demonstrators are oriented on data sharing between several users and possibly several organisations. Geo-Publication Services with CLARUS inside are compliant with European INSPIRE Recommendations that enforce publication and share of public data in the environmental field.	Support of most popular RDBMS protocol (Postgresql), standards protocols for Geodata services, S3 compatibility, demonstrators with CSP-agnostic cloud services.
	Location of data	End-User control on security policies and technical solutions is enabled by CLARUS Solution	Specific Data and Security Policy Viewer application has been developed - Web-based and possibly native Android too.
	Ownership	End-User control on security policies and technical solutions is enabled by CLARUS Solution	Specific Data and Security Policy Viewer application has been developed - Web-based and possibly native Android too.
	Access to data	End-User control on security policies and technical solutions is enabled by CLARUS Solution	Specific Data and Security Policy Viewer application has been developed - Web-based and possibly native Android too.
	Access to public data	CLARUS Proxies can be used in pass-thru mode for non-	Various types of security policies can be considered

		sensible data / Security Policies Management.	through CLARUS settings: -Data encryption policies -Data fragmentation and distribution policies -Access control policies Specific Data and Security Policy Viewer application has been developed - Web-based and possibly native Android too.
	Usability	Knowledge of Security techniques and most relevant cases to which they applied is made explicit in MetaData Database. Easy-to-Use Web tools are made available in order to facilitate management and visualisation of data both in organisation's trusted zone and in the Cloud.	Specific Data and Security Policy Viewer application has been developed - Web-based and possibly native Android too.
	Interoperability	Data Operations supported by CLARUS can be applied to wide range of datasets types.	Large set of Data Operations Modules combined with Security Techniques such as Data Anonymisation, Data Splitting, Homomorphic Encryption, Encryption, Searchable Encryption, Verifiable Search
	Switch of CSPs	Use of Multiple CSPs is possible through multiple CLARUS Proxies settlement. Switch for one CSP to another is always possible since no CLARUS deployment is necessary at the CLOUD level (CLARUS acts in organisations trusted areas).	CLARUS Proxies
	Cloud certifications	Certifications of compatibility with CLARUS Solution can be given to different Clouds for specific combination of Security Techniques and Data Operations for the different application that will be used with CLARUS: - Storage, Update, Retrieval, Search, Compute, Verification Data Operations - Data Anonymisation, Data Splitting, Homomorphic Encryption, Encryption, Searchable Encryption,	CLARUS-CSP Protocol Module + Data Operations Modules embedded in CLARUS Proxies

		Verifiable Search Techniques	
	Cloud contracts (SLA)	Guidelines and checklists for both CLARUS end users and service provider	Documentation Management in OpenSource Project repositories
Coco Cloud	Free Movement of data		Coco Cloud Engine
	Location of data		Coco Cloud Engine DSA Subsystem
	Ownership		Coco Cloud Engine DSA Subsystem
	Access to data		Coco Cloud Engine
	Access to public data		Coco Cloud Engine DSA Subsystem
	Usability		Coco Cloud Engine
	Interoperability		Coco Cloud Engine
	Switch of CSPs		
	Cloud certifications		
	Cloud contracts (SLA)		DSA Subsystem
CREDENTIAL	Free Movement of data	• Privacy requirements elicitation and monitoring	
	Location of data	N/A	N/A
	Ownership	• Adaption of advanced cryptographic schemes (proxy re-encryption, redactable signature schemes, etc.)	Cryptographic primitives such as attribute-based credentials on encrypted data
	Access to data	• Adaption of advanced cryptographic schemes (proxy re-encryption, redactable signature schemes, etc.) • Support of fine-granular and dynamic access control policies	Cryptographic primitives such as attribute-based credentials on encrypted data
	Access to public data	N/A	N/A
	Usability	• Development of dedicated mobile apps, browser plugins, etc. to reduce the knowledge, understanding, and actions required by users to a minimum • End-user involvement at all stages of the development	CREDENTIAL generic and pilots specific mobile apps
	Interoperability	• Analysis and potential extension of existing IAM solutions like SAML to support the used cryptographic primitives	
	Switch of CSPs	N/A	N/A
	Cloud certifications	N/A	N/A

	Cloud contracts (SLA)	• Transfer of project results into a dedicated new certification catalogue of the StarAudit cloud certification scheme.	Cloud certification catalogue
ESCUDO-CLOUD	Free Movement of data	Development of techniques for providing self-protection of data and support of access sharing restrictions	Tools enforcing self-protection over data based on encryption and over-encryption for policy management
	Location of data	Consideration of multi-authority and multi-providers scenarios Consideration of selective sharing restrictions	Tools enforcing selective access and sharing
	Ownership	Empowerment of data owners with control over their data (in storage, sharing, processing)	Tools for enforcing self-protection of data Tools enabling specification of access control policies and selecting sharing restrictions
	Access to data	Consideration of access control policies and access requirements over data	Tools for enforcing access restrictions and enabling authorized access and processing
	Access to public data		
	Usability	Consideration is being given to the integrability of the techniques and tools with existing cloud solutions	
	Interoperability	Consideration of multi-authority, multi-provider, and federated scenarios Support for collaborative queries	Design and implementation of federated object storage based on requirements of the Data Protection as a Service for multi cloud environments. Technique to protect access confidentiality in multi-cloud environments. Techniques supporting execution of collaborative queries involving different data authorities and providers
	Switch of CSPs	Consideration of SLA and providers guarantees	Approaches to reason about SLAs and satisfaction of requirements by different

			providers
	Cloud certifications		
	Cloud contracts (SLA)	Consideration of security and privacy aspects in SLAs	Technique based on security metrics for the evaluation of cloud providers compliant with the requirements of a cloud storage service. Approaches to reason about SLAs and satisfaction of requirements by different providers
MUSA	Free Movement of data	• Continuous monitoring techniques for composite SLA fulfilment assurance.	• MUSA Security Assurance Platform
	Location of data	• Extensions to CAMEL language for better addressing multi-cloud security and deployment requirements. • Cloud Service Providers (CSP) selection supporting mechanisms. • Continuous monitoring techniques for composite SLA fulfilment assurance.	• CSP Decision Support tool. • MUSA Security Assurance Platform
	Ownership	N/A	N/A
	Access to data	• Security enforcement mechanisms for multi-cloud (such as access control and scalability).	• MUSA Security Assurance Platform • Access control enforcement agents.
	Access to public data	Composition rules for creation of composite SLA that take into account component level and overall level SLOs.	N/A
	Usability	MUSA framework as integrated toolset.	• MUSA DevOps framework
	Interoperability	• Cloud Service Providers (CSP) selection supporting mechanisms. • Multi-cloud deployment model.	• Multi-cloud Deployer
	Switch of CSPs	• DevOps oriented Risk Analysis methodology. • Cloud Service Providers (CSP) selection supporting mechanisms.	• DevOps Risk Analysis tool. • CSP Decision Support tool.
	Cloud certifications	• Continuous monitoring techniques for composite SLA	• MUSA Security Assurance Platform

		fulfilment assurance.	
	Cloud contracts (SLA)	• DevOps oriented Risk Analysis methodology. • Security and privacy-aware SLA model that includes SLOs for security and privacy, expressing security controls and security metrics. • Cloud Security metrics Catalogue. • Composition rules for creation of composite SLA that take into account component level and overall level SLOs. • Multi-cloud Threat Catalogue.	• DevOps Risk Analysis tool. • SLA Generator for composite security and privacy-aware SLAs.
OPERANDO	Free Movement of data		•Continuous monitoring of changes in privacy settings by Online Service Providers • Anonymization tool • Regulator API which facilitates compliance with privacy laws as well as auditing and pro-active supervision of OSPs by privacy regulators.
	Location of data		• Regulator API which facilitates compliance with privacy laws as well as auditing and pro-active supervision of OSPs by privacy regulators.
	Ownership		• OPERANDO Platform that implements the “Privacy Authority”. • Privacy Policy Computation engine • User device enforcement that supports delivery of privacy enforcing services by client applications. • Regulator API which facilitates compliance with privacy laws as well as auditing and pro-active supervision of OSPs by privacy regulators.
	Access to data		• OPERANDO Platform that implements the “Privacy Authority”. • Privacy Policy Computation engine

			• Web Browser and Application: Identify Management
	Access to public data		
	Usability		Dedicated web browser plug ins and application.
	Interoperability		
	Switch of CSPs		
	Cloud certifications		Several APIs to facilitate compliance with privacy laws as well as auditing.
	Cloud contracts (SLA)		
PaaSword	Free Movement of data	PaaSword Security Policy Models, where three main types of security policies are considered: -Data encryption policies -Data fragmentation and distribution policies -Access control policies	PaaSword Framework
	Location of data	PaaSword Security Policy Models, where three main types of security policies are considered: -Data encryption policies -Data fragmentation and distribution policies -Access control policies	PaaSword Database Proxy.
	Ownership	PaaSword Security Policy Models, where three main types of security policies are considered: -Data encryption policies -Data fragmentation and distribution policies -Access control policies	PaaSword Key Management Mechanism
	Access to data	PaaSword Security Policy Models, where three main types of security policies are considered: -Data encryption policies -Data fragmentation and distribution policies -Access control policies	The PaaSword Annotation Interpretation Mechanism. is the used to efficiently interpret the annotations into XACML-based enforceable Access Control Policies.
	Access to public data		
	Usability	PaaSword Context-aware Security Model	PaaSword Annotations Governance And Validity Control Mechanism (AGVC)
	Interoperability		
	Switch of CSPs		

	Cloud certifications		
	Cloud contracts (SLA)		
PRISMACLOUD	Free Movement of data	Security and privacy patterns for PC services	• Protect confidentiality, integrity and availability of data. • Protect authenticity of data for agile cloud based data sharing
	Location of data		Tools and services for infrastructure auditing enable privacy friendly geolocation audits
	Ownership		• Methods for end-to-end authenticity for cloud based data sharing guarantee quality and also ownership over trust boundaries • Verifiable computings enable aggregation of authentic data without destroying ownership information
	Access to data		• Privacy friendly sharing of authentic data is enabled by built-in selective disclosure • Verifiable data processing based data sharing enables to give only access to statistics of authentic data
	Access to public data		Anonymization techniques for large datasets
	Usability	• HCI guidelines for usage of cryptographic services • End user guidance for secure and privacy friendly cloud usage	• Keyless secure data sharing services. • Transparent encryption and anonymization services
	Interoperability		
	Switch of CSPs	Guidelines for secure multi-cloud storage deployment	• Development of multi-cloud storage services which prevent from lock-in. • Usage of agile digital signatures to protect authenticity of data for more flexible provider switch
	Cloud certifications	Security and privacy patterns for cloud usage	Tools and services for infrastructure auditing
	Cloud contracts (SLA)	• SLA models and policies for multi-cloud storage • Capability models for PRISMACLOUD services	• Multi-cloud storage solutions on the basis of fragmentation increase more flexibility in SLA configurations

SPECS	Free Movement of data	Offers a security SLA models able to specify security criteria related to data management and access.	SPECS Framework helps to automatize enforcement of security SLA, even for data management. SPECS ViPR+ explicitly focuses on security SLA for data storage.
	Location of data	Dedicated Security metric to grant and monitor data location	Tools to monitor data location
	Ownership	Data Encryption techniques; access control policies.	Tools to grant ownership and access to data
	Access to data	Data Encryption techniques; access control policies.	Tools to grant ownership and access to data
	Access to public data		
	Usability	Simple web-based GUI and user guides	Tools to grant interoperability and usability of data
	Interoperability	Common model to express security SLA and techniques to compare CSP	Tools to grant interoperability and usability of data
	Switch of CSPs	Common model to express security SLA and techniques to compare CSP	Machine readable format to express security SLA and evaluation techniques
	Cloud certifications	Common model to express security SLA	Machine readable format to express security SLA
	Cloud contracts (SLA)	Common model to express security SLA	Machine readable format to express security SLA
SUNFISH	Free Movement of data	Dynamic data transformation using cryptographic functions and redaction Selective transformational sharing restrictions	• Data transformation service. • Data masking, data encryption, format preserving encryption, secure multi-party computation and anonymization.
	Location of data	Federation-level unified security policies are employed to express data location requirements Secure multiparty logging and monitoring of transactions to ensure accountability	• Administrative console for policy management. • Blockchain-based policy store and transaction log. • Federated policy enforcement infrastructure • Federated monitoring and anomaly detection.
	Ownership	Specification of data access restrictions using federated security policies Secure multiparty logging and monitoring of transactions to ensure accountability	• Administrative console for policy management. • Blockchain-based policy store and transaction log. • Federated policy enforcement infrastructure • Federated monitoring and anomaly detection.

	Access to data	Security enforcement mechanisms for multi-cloud , distributed policy enforcement Selective and transformative data disclosure Specification of data access restrictions using federated security policies Secure multiparty logging and monitoring of transactions to ensure accountability	• Administrative console for policy management. • Blockchain-based policy store and transaction log. • Federated policy enforcement infrastructure. • Federated monitoring and anomaly detection.
	Access to public data	Transforming sensitive data prior to its publication	• Administrative console for policy management. • Data transformation service. • Federated policy enforcement infrastructure. • Data masking, data encryption, format preserving encryption, secure multi-party computation and anonymization.
	Usability	Managing configurations using web interface	• Federated administrative console.
	Interoperability	Dynamic data transformation Federating heterogeneous infrastructures Identity management federation	• Federate security policies. • FaaS - Cloud Federation as a Service. • Support for SAML and OpenID Connect.
	Switch of CSPs	-	-
	Cloud certifications		
	Cloud contracts (SLA)	Federated SLA metrics Continuous SLA monitoring	• Blockchain-based logging
SWITCH	Free Movement of data	Secure storage in containers	SIDE and DRIP subsystems
	Location of data	Planning aware multi-cloud deployment	SIDE and DRIP subsystems
	Ownership	Use case owners	SIDE subsystem
	Access to data	Access control at component level	SIDE subsystem
	Access to public data	N/A	N/A
	Usability	IDE framework	SIDE
	Interoperability	DevOps using multi-cloud approach	DRIP subsystem
	Switch of CSPs	Knowledge base for multi-cloud using use case relevant	ASAP subsystem

		metrics.	
	Cloud certifications	Model to evaluate cloud SLA	ASAP subsystem
	Cloud contracts (SLA)	Continuous monitoring of SLA fulfilment	ASAP subsystem
TREDISEC	Free Movement of data		
	Location of data		
	Ownership	Security primitives	Security Primitives
	Access to data	Security primitives	Security Primitives
	Access to public data	Security primitives	Security Primitives
	Usability	Security primitives TREDISEC Framework	TREDISEC Framework. Security Primitives
	Interoperability	Security primitives TREDISEC Framework	TREDISEC Framework. Security Primitives
	Switch of CSPs	Security primitives TREDISEC Framework	TREDISEC Framework. Security primitives
	Cloud certifications	Security primitives	Security Primitives
	Cloud contracts (SLA)	Security primitives	Security Primitives
WITDOM	Free Movement of data	SPACE, WITDOM Privacy framework	PO, SSP, SC, ICV, Anonymization, DM, E2EE
	Location of data		
	Ownership		
	Access to data		
	Access to public data		
	Usability	SPACE	
	Interoperability		Transformation Services
	Switch of CSPs		Broker
	Cloud certifications		
	Cloud contracts (SLA)		

5. Technology options to address Free Flow of Data Issues

This section explains the technologies and solutions that address the different problems that Free Flow of Data initiative would raise. We provide the explanation of the technology options grouped in two major categories corresponding to WGs of the Cluster:

- WG1: Advanced security and data protection mechanisms
- WG2: Trust & Interoperability

a. Advanced Security and data protection technologies for Free Flow of Data

This section summarises the technical solutions that each project offers to address the topics related to Free Flow of Data initiative. The following analysis founds and re-elaborates the data already presented by the projects in previous section, outlining, for each topic or area of work in FFD, the contributions offered by different projects.

Free Movement of data

The following table summarises all the techniques and tools offered by the clustered projects that can be used in order to simplify the free movement of data. The most part of project solutions focuses on support to policy specification and specification into Service Level Agreement of security objectives related to data. Many of the tools are reported also in other topics (like access to data and access to public data).

Project	Free Movement of data
CLARUS	CLARUS Support of most popular RDBMS protocol (Postgresql), standards protocols for Geodata services, S3 compatibility, demonstrators with CSP-agnostic cloud services. No CSP lock-in is a key requirement in the definition of the CLARUS solution.
Coco Cloud	Coco Cloud Engine offers solutions to manage privacy information on data collected in Cloud.
ESCUDO –CLOUD	ESCUDO-Cloud provides tools enforcing self-protection over data based on encryption and over-encryption for policy management. ESCUDO-Cloud develops techniques for providing self-protection of data and support of access sharing restrictions
MUSA	MUSA Security Assurance Platform: enable Continuous monitoring techniques for composite SLA fulfilment assurance.

OPERANDO	Continuous monitoring of changes in privacy settings by Online Service Providers. Anonymization tool. Regulator API for compliance and auditing.
PaaSword	PaaSword Framework supports Security Policy Models, where three main types of security policies are considered: -Data encryption policies -Data fragmentation and distribution policies -Access control policies
PRISMACLOUD	PRISMACLOUD offers tools to protect confidentiality, integrity and availability of data, focusing on Security and privacy patterns for PC services.
SPECS	SPECS offers a security SLA model able to capture security criteria related to data management and access. SPECS Framework helps to automatize enforcement of security SLA, even for data management. SPECS ViPR+ explicitly focuses on security SLA for data storage.
SWITCH	ASAP subsystem of SWITCH allows a SLA model to capture metrics related to data management and access.
SUNFISH	SUNFISH offers integrated cloud federation framework with federation-level security management components that allow specification and federated enforcement of security policies. Beside access control functionalities, these may deal with context-based transformational goals that allow online data redaction using data masking, data encryption, format-preserving encryption and anonymization. Accountability in the framework is ensured by using blockchain based federated monitoring and anomaly detection architecture.
WITDOM	The WITDOM Privacy framework offers the following technologies that may apply to free movement of data. Involved WITDOM components are: PO, SSP, SC, ICV, Anonymization, DM, E2EE

Location of data

Data localization is often considered as one of the main inhibitors for the cloud adoption: CSP often refuses to clearly declare where data are stored and managed, when reported in the cloud, in order to be freely able to manage them according to their internal policy and internal management systems.

Technologies to keep track of data localisation, however, are already available and the following clustered projects offer some possible solutions to address explicitly the issue.

Project	**Location of Data**
CLARUS	End-User control on security policies and technical solutions is enabled by CLARUS Solution. CLARUS offers a Specific Data and Security Policy Viewer

	application, Web-based and possibly native Android too.
Coco Cloud	Coco Cloud Engine DSA Subsystem.
ESCUDO-CLOUD	ESCUDO-CLOUD has taken into consideration of multi-authority and multi-providers scenarios and selective sharing restrictions. ESCUDO-CLOUD offers Tools enforcing selective access and sharing.
MUSA	MUSA offers: • Extensions to CAMEL language for better addressing multi-cloud security and deployment requirements. • Cloud Service Providers (CSP) selection supporting mechanisms using the MUSA CSP Decision Support Tool. • Continuous monitoring techniques for composite SLA fulfilment assurance using the MUSA Security Assurance Platform.
OPERANDO	Regulator API for compliance and auditing.
PaaSword	PaaSword Framework supports Security Policy Models, where three main types of security policies are considered: -Data encryption policies. -Data fragmentation and distribution policies. -Access control policies.
PRISMACLOUD	Tools and services for infrastructure auditing enable privacy friendly geolocation audits.
SPECS	SPECS proposed a dedicated Security metric to grant and monitor data location. SPECS ViPR extensions enable the tools to monitor data location.
SUNFISH	SUNFISH integrates specification and enforcement of data location requirements within federated security management infrastructure. The data location enforcement optionally allows balancing between security and utility of the enforcement, enabling the online transformation of data to allow its use outside the intended infrastructure in security and privacy conform way.
SWITCH	SIDE subsystem of SWITCH allows specifying geo-location requirements for the infrastructure.

Ownership

Ownership of data is one of the key issues in the context of Free Flow of Data as it will allow data sovereignty and data economies. The following clustered projects offer solutions to support the management of ownership of data.

Project	Ownership
CLARUS	Specific Data and Security Policy Viewer application - Web-based and possibly native Android too.
Coco Cloud	Coco Cloud Engine. DSA Subsystem.
CREDENTIAL	Cryptographic primitives such as attribute-based credentials on encrypted data.
ESCUDO-CLOUD	Tools for enforcing self-protection of data. Tools enabling specification of access control policies and selecting sharing restrictions.
OPERANDO	OPERANDO Platform that implements the "Privacy Authority". Privacy Policy Computation engine. User device enforcement that supports delivery of privacy enforcing services by client applications. Regulator API for compliance and auditing.
PaaSword	PaaSword Key Management Mechanism.
PRISMACLOUD	Methods for end-to-end authenticity for cloud based data sharing guarantee quality and also ownership over trust boundaries. Verifiable computings enable aggregation of authentic data without destroying ownership information.
SPECS	Tools to grant ownership and access to data.
SUNFISH	SUNFISH includes the tools to specify ownership and data access constraints in federated cloud environment using federated administration console and an extension that supports the creation of administrative policies on a level of participating clouds.

TREDISEC	Proof of Ownership security primitive deals with the assurance that a Cloud client indeed possesses a given file

Access to data

Access control is one of the most explored fields in the security context. Free Flow of Data implies adoption of fine grained solutions to control the accesses to data. The following clustered projects offer solutions to monitor, control, define and enforce policies to manage access to data.

Project	**Access to Data**
CLARUS	End-User control on security policies and technical solutions is enabled by CLARUS Solution. Specific Data and Security Policy Viewer application has been developed - Web-based and possibly native Android too.
CREDENTIAL	Adaption of advanced cryptographic schemes (proxy re-encryption, redactable signature schemes, etc.). Support of fine-granular and dynamic access control policies. Cryptographic primitives such as attribute-based credentials on encrypted data.
ESCUDO-CLOUD	ESCUDO-CLOUD address the problem of access control policies and access requirements over data. ESCUDO-CLOUD offers tools for enforcing access restrictions and enabling authorized access and processing.
MUSA	MUSA addressed issues related to Security enforcement mechanisms for multi-cloud (such as access control and scalability). MUSA Security Assurance Platform and MUSA Enforcement agents enforce techniques to control access to data.
OPERANDO	OPERANDO Platform that implements the “Privacy Authority”. Privacy Policy Computation engine. Web Browser and Application: Identify Management.
PaaSword	PaaSword Framework supports Security Policy Models, where three main types of security policies are considered: -Data encryption policies. -Data fragmentation and distribution policies. -Access control policies.

	The PaaSword Annotation Interpretation Mechanism. is the used to efficiently interpret the annotations into XACML-based enforceable Access Control Policies.
PRISMACLOUD	Privacy friendly sharing of authentic data is enabled by built-in selective disclosure. Verifiable data processing based data sharing enables to give only access to statistics of authentic data.
SPECS	Data Encryption techniques; access control policies. Tools to grant ownership and access to data.
SUNFISH	SUNFISH includes the federated access control enforcement that depends on federated administrative console (defining constraints), distributed policy enforcement infrastructure (to evaluate and enforce policies), distributed transformation services (to provide online transformation of data in move) and federated monitoring and anomaly detection architecture (backed by blockchain to ensure federation wide accountability and non-repudiation).
SWITCH	SIDE subsystem
TREDISEC	Several Security Primitives enable to control access to data. Proof of Retrievability (PoR) deal with the dual problem of ensuring - at the client-side - that a server still stores the files it ought to, and he requirement of data integrity (ensuring that data has not undergone malicious modifications) and availability (ensuring that data is still available in its entirety and can be downloaded if needed). Efficient Shared Ownership, Secure Data Deletion primitives and Multi-tenant Secure Enforcement of Policies ensure access to data stored in the Cloud is guaranteed in the conditions that the client states.

Access to public data

Access to public data is a subset of the Access to data problem faced above. It outlines the need of publicly disclosing or sharing some part(s) of the data, even maintaining other data reserved. The following clustered projects offer solutions to monitor, control, define and enforce policies to manage access to data.

Project	Access to public data
CLARUS	Geo-Publication Services with CLARUS inside are compliant with European INSPIRE Recommendations that enforce publication and share of public data in the environmental field.

	CLARUS Proxies can be used in pass-thru mode for non-sensible data / Security Policies Management. Various types of security policies can be considered through CLARUS settings: • Data encryption policies • Data fragmentation and distribution policies • Access control policies Specific Data and Security Policy Viewer application has been developed - Web-based and possibly native Android too.
Coco Cloud	DSA Subsystem.
MUSA	MUSA SLA Generator enables to manage composition rules for creation of composite SLAs that take into account component level and overall application level SLOs, which may affect the public access to data.
PRISMACLOUD	Anonymization techniques for large datasets.
SUNFISH	Data transformation service, anonymization, format-preserving encryption
TREDISEC	Resource Isolation Security Primitives (for containers e.g. Docker, at Hypervisor level, Fuzz testing) and Software hardening techniques ensure the Cloud infrastructure and Platform improve resistance to attacks and vulnerabilities, contributing to a secure and reliable Access to Public data.

Usability

The trade-off between usability and security is a well-known problem. Introduction of complex regulation can be inhibited in its concrete application, if the technical solutions are too hard to be used by common users. The following clustered projects offer solutions that help in improving usability of security solutions.

Project	Usability
CLARUS	Knowledge of Security techniques and most relevant cases to which they applied is made explicit in MetaData Database. Easy-to-Use Web tools are made available in order to facilitate management and visualisation of data both in organisation's trusted zone and in the Cloud.

Coco Cloud	Coco Cloud Engine
CREDENTIAL	Development of dedicated mobile apps, browser plugins, etc. to reduce the knowledge, understanding, and actions required by users to a minimum End-user involvement at all stages of the development CREDENTIAL generic and pilots specific mobile apps
ESCUDO-CLOUD	Consideration is being given to the integrability of the techniques and tools with existing cloud solutions
MUSA	MUSA framework as integrated toolset. MUSA DevOps framework
OPERANDO	Dedicated web browser plug ins and application.
PaaSword	PaaSword Context-aware Security Model PaaSword Annotations Governance And Validity Control Mechanism (AGVC)
PRISMACLOUD	HCI guidelines for usage of cryptographic services End user guidance for secure and privacy friendly cloud usage Keyless secure data sharing services Transparent encryption and anonymization services
SPECS	Simple web-based GUI and user guides Tools to grant interoperability and usability of data
SUNFISH	Federated administrative console for definition of security policies, configuration management and event monitoring
SWITCH	SIDE subsystem
TREDISEC	Security primitives in TREDISEC Framework

Interoperability

Interoperability is one of the key factors to enable free flow of data. The following clustered projects offer solutions devoted to improve systems interoperability.

Project	Interoperability
CLARUS	Data Operations supported by CLARUS can be applied to wide range of datasets types. Large set of Data Operations Modules combined with Security Techniques such as Data Anonymization, Data Splitting, Homomorphic Encryption, Encryption, Searchable Encryption, Verifiable Search.
Coco Cloud	Coco Cloud Engine.
CREDENTIAL	Analysis and potential extension of existing IAM solutions like SAML to support the used cryptographic primitives.
ESCUDO-CLOUD	Consideration of multi-authority, multi-provider, and federated scenarios; Support for collaborative queries. Design and implementation of federated object storage based on requirements of the Data Protection as a Service for multi cloud environments. Technique to protect access confidentiality in multi-cloud environments. Techniques supporting execution of collaborative queries involving different data authorities and providers.
MUSA	Cloud Service Providers (CSP) selection supporting mechanisms. Multi-cloud deployment model. Multi-cloud Deployer.
SPECS	Common model to express security SLA and techniques to compare CSP. Tools to grant interoperability and usability of data.
SUNFISH	SUNFISH offers integrated dynamic data transformation service (to support different formats), federated identity management (to support different IdM subsystems and protocols) and FaaS establishment and management tools (to support federating heterogeneous infrastructures and cloud managements frameworks)
SWITCH	DRIP subsystem

TREDISEC	The TREDISEC framework allows for different implementations of the Security Primitives, to support different scientific approaches to solve the same original problem (primitive pattern) but also different technologies used for the solution and the deployment in different target cloud environments (TREDISEC recipes). The testing component and the deployment component of the framework are the particular elements that contribute to this support.
WITDOM	Transformation Services.

Switch of CSPs

“Vendor Lock-in”, “Closed Gardens”, “proprietary solutions”, “closed APIs” are just some of the terms that outline the problem of solutions that work only with one Cloud Service Provider. Free flow of data requires the capability of easily move data (and in some cases applications too) among different CSPs. The following clustered projects are addressing the issues related to such problem:

Project	Switch of CSPs
CLARUS	Use of Multiple CSPs is possible through multiple CLARUS Proxies settlement. Switch for one CSP to another is always possible since no CLARUS deployment is necessary at the CLOUD level (CLARUS acts in organisations trusted areas). CLARUS Proxies can be used as a tool that implements such techniques.
ESCUDO-CLOUD	Consideration of SLA and providers guarantees. Approaches to reason about SLAs and satisfaction of requirements by different providers.
MUSA	DevOps oriented Risk Analysis methodology. Cloud Service Providers (CSP) selection supporting mechanisms. DevOps Risk Analysis tool. CSP Decision Support tool.
PRISMACLOUD	Guidelines for secure multi-cloud storage deployment. Development of multi-cloud storage services which prevent from lock-in. Usage of agile digital signatures to protect authenticity of data for more flexible provider switch.

SPECS	Common model to express security SLA and techniques to compare CSP. Machine readable format to express security SLA and evaluation techniques.
TREDISEC	The TREDISEC framework allows for different implementations of the Security Primitives but also permit the deployment in different target cloud environments (TREDISEC Recipes). The testing component and the deployment component of the framework are the particular elements that contribute to this support
WITDOM	WITDOM Broker.

Cloud Services Certification

Adoption of Certification techniques is often considered a reliable way to delegate to a third party the verification of quality of products. Such verification may play a key role in the free flow of data. The following clustered projects offer solutions related to certification.

Project	Cloud Services Certification
CLARUS	Certifications of compatibility with CLARUS Solution can be given to different Clouds for specific combination of Security Techniques and Data Operations for the different application that will be used with CLARUS: - Storage, Update, Retrieval, Search, Compute, Verification Data Operations -Data Anonymisation, Data Splitting, Homomorphic Encryption, Encryption, Searchable Encryption, Verifiable Search Techniques The tools involved are CLARUS-CSP Protocol Module + Data Operations Modules embedded in CLARUS Proxies.
MUSA	Continuous monitoring techniques for composite SLA fulfilment assurance. MUSA Security Assurance Platform.
PRISMACLOUD	Security and privacy patterns for cloud usage. Tools and services for infrastructure auditing.
SPECS	Common model to express security SLA Machine readable format to express security SLA.
SWITCH	ASAP subsystem

TREDISEC	Many Security primitives (Verifiable Computation, Verifiable Storage and Verifiable Ownership) contribute to Cloud Services Certification by assessing the correctness of the outsourced computation, and allowing a cloud customer to check whether her (Big) data is stored correctly at the cloud server provider.

Cloud contracts (SLA)

Adoption of Service Level Agreement (SLA) enables to offer a clear statement of the grants each CSP offers to the others. Similarly, the SLA of a (multi-)cloud-based application reflects the grants the application can offer to its customers, which is dependent on the Cloud contracts signed with the CSPs it uses. This is considered a key enabling factor to free flow of data and in general in Cloud adoption for system with security and/or critical requirements. The following clustered projects offer technical solutions to manage SLAs in Cloud.

Project	Cloud contracts (SLA)
CLARUS	Guidelines and checklists for both CLARUS end users and service provider. Documentation Management in Open Source Project repositories.
Coco Cloud	DSA Subsystem.
CREDENTIAL	Transfer of project results into a dedicated new certification catalogue of the StarAudit cloud certification scheme. Cloud certification catalogue.
ESCUDO-CLOUD	Consideration of security and privacy aspects in SLAs. Technique based on security metrics for the evaluation of cloud providers compliant with the requirements of a cloud storage service. Approaches to reason about SLAs and satisfaction of requirements by different providers.
MUSA	DevOps oriented Risk Analysis methodology. Security and privacy-aware SLA model that includes SLOs for security and privacy, expressing security controls and security metrics. Cloud Security Metrics Catalogue. Composition rules for creation of composite SLA that take into account component level and overall level SLOs.

	Multi-cloud Threat Catalogue. DevOps Risk Analysis tool. SLA Generator for composite security and privacy-aware SLAs.
PRISMACLOUD	SLA models and policies for multi-cloud storage. Capability models for PRISMACLOUD services. Multi-cloud storage solutions on the basis of fragmentation increase more flexibility in SLA configurations.
SPECS	Common model to express security SLA Machine readable format to express security SLA.
SUNFISH	Integrated tools for continuous monitoring of SLA metrics. Federated monitoring and logging based on blockchain technology.
SWITCH	ASAP subsystem
TREDISEC	The Remote Attestation Security Primitive provides a proof or evidence of the integrity of the cloud platform and the software where the client outsources their business application/service. This evidence can be used by an SLA monitoring component to compute the trust value associated to a particular cloud environment at any given time.

b. Trust and interoperability technologies for Free Flow of Data

The following paragraphs summarise the conclusions of CloudWatch2 project research on trust and data portability of personal data.

Trust is instrumental for a safe and reliable free flow of data (whether personal data or not) and for the protection of the four fundamental freedoms of the EU single market enshrined in the Treaties (goods, workers, service provision and capital). Reliable technologies and legal instruments are both necessary for allowing cloud service providers and their users (e.g., undertakings, public authorities, natural persons) to transfer data in the cloud environment.

The provision of cloud services very often entail that personal data are processed in servers and infrastructures located outside the European Union. It is unavoidable, in such cases, that personal data are transferred outside the EU. The utmost attention must be paid to the rules governing the flow of personal data from the European legal space to the outer world.

In the forthcoming General Data Protection Regulation, the EU legislators translated the need of “trust” for the free flow of data in a series of legal instruments permitting a lawful transfer of personal data. According to the Regulation, personal data can be transferred outside the European Union on the basis of an adequacy decision related to the country where the recipient of the transfer is located, pursuant to Article 45 thereof, or where specific safeguards have been put in place, in accordance with Article 46 thereof (one particular example of such safeguards are the Binding Corporate Rules, which are specifically addressed in Article 47) or if one of the derogations contained in Article 49 applies. Therefore, for a personal data transfer to be valid, one of the provisions contained in Articles 45, 46, 47 or 49 must apply. Given that only a few countries have been awarded an adequacy decision by the European Commission (full updated list is available here: http://ec.europa.eu/justice/data-protection/international-transfers/adequacy/index_en.htm), and that the derogations in Article 49 apply in limited cases, most situations will require adequate safeguards to be adopted, pursuant to Articles 46-47.

On the other hand, “interoperability” was codified in Article 20 of the General Data Protection Regulation. Cloud service providers using proprietary data formats and service interfaces may lead to a lock-in effect which renders the interoperability and portability of data from a cloud service provider to another difficult if not impossible.

The Regulation introduces a brand new right to data portability that should enable the data subject to receive the personal data provided to the service provider, in a structured, commonly used machine- readable format, and the right to transmit it to another provider.

The right to data portability is surely enforceable against private data controllers, whereas according to Recital (68) of the Regulation “by its very nature, that right should not be exercised against controllers processing personal data in the exercise of their public duties”. Data portability however depends on the availability of standards which lead to interoperability. Therefore, preference should be given to interoperability standards that would make the data portable at the request of the data subjects; as clarified above, the latter obligation does not apply to public authorities (in their role of data controllers) processing personal data in the cloud in the exercise of their public duties.

The lock-in effect might also hurdle the migration of services that the client developed on a platform offered by the original cloud service provider (PaaS).

By May 2018 (when the Regulation will come into force) the European Union will have the implemented suitable mechanisms enforcing trust and interoperability for the free flow of data within and outside the Union.

6. Conclusions

This document is the first attempt to provide a collection of solutions to address issues in FFD that are available as a result of the work of the clustered projects.

The document does not intend to provide an exhaustive and detailed description of all the available resources from the clustered projects that do solve some of the technology challenges brought by the FFD. The aim of the document is to serve as reference to all the technologies, methodologies, tools, prototypes, etc. that the projects have developed or are developing and that are intended to solve some of the aspects of the working areas of the FFD, such as free movement of data, ownership, access to data, etc. The collection is described per project and per area of work within the FFD.

For further details and better understanding of the solutions proposed, please visit the referenced websites of the projects and open source repositories.

The whitepaper is the result of the collaborative work between the clustered projects. More information on the Cluster, the projects within and the joint activities carried out can be found in the cluster website[30].

[30] https://eucloudclusters.wordpress.com/data-protection-security-and-privacy-in-the-cloud/