H2020-SC6-GOVERNANCE-2018-2019-2020

DT-GOVERNANCE-05-2018-2019-2020

D3.1 Design of the ACROSS Data Governance framework for data sovereignty – Initial

Project Reference No	959157 — ACROSS — H2020-SC6-GOVERNANCE-2018-2019-2020
Deliverable	D3.1 Design of the ACROSS Data Governance framework for data sovereignty – Initial
Work package	WP3: ACROSS Data Governance framework
Nature	Report
Dissemination Level	Public
Date	31/10/2021
Status	Final v1.0
Editor(s)	Valentín Sánchez (TEC)
Contributor(s)	Idoia Murua, Urtza Iturraspe (TEC), Timo Behrmann (DAT), Max Kortlander (WAAG), Nikos Vasilakis (GRNET), Matīss Veigurs (VARAM)
Reviewer(s)	Vincenzo Savarino (ENG), Enrique Areizaga (TEC)
Document description	This report includes 1) the requirements, functional specification, the technical architecture, the design of the modules and a description its APIs; 2) the design of a generic data governance mechanism, such as a data governance data model for handling data access rights; 3) mock-ups of the user interface and 4) the relevant baseline technologies that will be used for the implementation of the data governance framework.

About

The project is co-funded by the European Commission's Horizon 2020 research and innovation framework programme. Spanning through three years, ACROSS consists of a consortium of 10 partners from 7 countries: Athens Technology Center (coordinator), Tecnalia, Dataport, Engineering, Fraunhofer, GRNET, TimeLex, The Lisbon Council, Waag and VARAM. The project kicked off its activities in February 2021, with an energising online meeting, where all partners took the floor to present their plans to make the project a great success.

Document Revision History

Version	Date	Modifications Introduced	
		Modification Reason	Modified by
V0.1	19/04/2021	Table of contents	TEC
V0.2	8/10/2021	GRNET contribution about Greek pilot and MyData hubs	GRNET
V0.3	8/10/2021	VARAM contribution about Latvian pilot	VARAM
V0.4	8/10/2021	VAAG Contribution about Attribute Based Encryption	VAAG
V0.5	19/10/2021	DATAPORT contribution about German pilot and the European Digital Identity	DATAPORT
V0.6	22/10/2021	First version for revision	TEC
V0.7	26/10/2021	Revision by Engineering	ENG
V1.0	28/10/2021	Final version	TEC

Executive Summary

The main objective of the ACROSS project is to provide the means (tools, methods and techniques) to enable user-centric design and implementation of interoperable cross-border (digital) public services compliant with the current European regulations (e.g. the Single Digital Gateway (SDG) and Once-Only principle (OOP), European Interoperability Framework (EIF)) where the private sector can also interconnect their services **while ensuring the data sovereignty of the citizens, who can set the privacy level that will allow the public and private sector to access to their data based on their requirements**.

In order to ensure the protection of personal data (and documents) and its compliance with GDPR and other relevant regulations, especially when shared between organizations, ACROSS will design and implement with **a data governance framework** where data subjects can control the use of their personal data empowering them**.**

The **data governance framework will** allow users to:

1) monitor which data are available and how they are used or how it has been accessed,
2) control: add, delete or change data, provide or block access to public bodies (where permitted or enforced by the law) businesses or data brokers, giving individuals the power to determine how their data can be used.

This report provides an accurate description of the ACROSS data governance framework requirements, along with a first draft of technical architecture, modules, and APIs.

It gathers the data governance, security and privacy requirements from the use cases, considering both the technical and operational perspectives (WP6), the final user expectations regarding data privacy (WP2) and the ACROSS platform integration strategy (WP4 and WP5).

Besides, the deliverable includes the design of a generic data governance mechanism, such as a data governance data model for handling data access rights.

Finally, a set of relevant baseline technologies that could be used for the implementation of the data governance framework has been analysed, including a short description of their scope and applicability to ACROSS. The technical requirements have been assessed and tested against the base technologies, MyData, Attribute Based Credentials and IDS, in order to identify which building blocks are available, and which modifications are required to ensure their fitness for purpose in the ACROSS project.

Table of Contents

List of Figures

List of Tables

List of Terms and Abbreviations

Abbreviation	Definition
IDSA	International Data Space Association
CPSV-AP	Core Public Service Vocabulary Application Profile
ABC	Attribute Based Credentials
GDPR	General Data Protection Regulation
DGA	Data Governance Act
PIMS	Personal Information Management System
SDGR	Single Digital Gateway Regulation
OOP	Once-only principle

1 Introduction

1.1 Purpose and Scope

One of the ACROSS objectives is **to ensure the protection of personal data (**and documents**) and its compliance with GDPR and other relevant regulations, especially when shared between organizations.** This objective will be fulfilled by designing and implementing a private/personal data governance framework where data subjects can control the use of their personal data empowering them.

ACROSS will offer the citizen the possibility of defining which public and private organization will be allowed to *access which data and for what purpose* trough the **ACROSS Data Governance Framework.** The main aim is to give the citizen the chance of **govern the access to** their data, profiting from a set of usage policies that implement levels of access and they can be the **sovereign owner** of such data.

The **data governance framework,** that allows users to

1) monitor which data are available and how they are used or how it has been accessed,
2) to control: add, delete or change data, provide or block access to public bodies (where permitted or enforced by the law) businesses or data brokers, giving individuals the power to determine how their data can be used.

From a technical point of view, the Data governance framework includes:

1) A "private/personal data" governance platform including a Personal data management site, which provides a user interface to define, manage and control the use of personal data. (Data portal)
2) A set of APIs/libraries to interact with the ACROSS platform

The governance framework will be based on existing solutions:

1. **MyData**[1] model for human-centered personal data management and processing
2. Built on experiences around **Attribute-Based Credentials** approaches in the DECODE[2] project,
3. Include generic ***data usage policies*** when the private data needs to be transferred among several stakeholders (IDSA Data Sovereignty)

This deliverable includes a first version of the ACROSS personal data governance framework requirements, functional specification, the technical architecture, the design of the modules and a description its APIs.

[1] https://mydata.org/
[2] http://decodeproject.eu/

Besides, it includes the design of a generic data governance mechanism, such as a **data governance data model** for handling data access rights.

Finally, a set of relevant baseline technologies that will be used for the implementation of the data governance framework is described.

1.2 Approach for Work Package and Relation to other Work Packages and Deliverables

The goal of WP3 is to design, implement and deploy a "private/personal data" governance framework that allows the citizens to control how their data and their activities are created or used by businesses, governments, or data brokers, giving individuals the power to determine how their data can be used. The governance framework will be based on existing solutions such as MyData model for human-centred personal data management and processing and built on experiences around Attribute-Based Credentials approaches in the DECODE project, but it will also include generic data usage policies when the private data needs to be transferred among several stakeholders.

The services from this WP will be integrated into the platform created in WP5 and will demonstrate the functionality of the use cases in WP6.

WP5 aims at providing the architectural and implementation aspects for the delivery of the ACROSS tools taking into account the full range of requirements for such service. The design of the ACROSS platform will drive the design and implementation of the various components produced in the context of WPs **WP3**, WP4 & WP5.

WP2 and WP6 together have defined the so-called user journeys based on the results several interviews with people from the three pilot countries. The aim of the interview process is form potential user journeys, building on initial ideas. User journeys can include actions, touch points, emotions, pain points, and phases. Eventually to result in concrete (socio-technical) requirements for the ACROSS platform modules. A specific section about Data privacy issues has been included in the questionnaire in order to gather requirements for the Data Governance Framework.

The Data Governance Framework will be designed as an independent platform, but it will share some components with the ACROSS platform (defined in WP4 and WP5). Furthermore, a set of APIs will be defined to interact with some other modules, as the User Journey Service Engine.

The decisions presented in this deliverable are a subject to refinements and modifications, based on the progress of the other work packages, as well as the validation and evaluation phases.

1.3 Methodology and Structure of the Deliverable

This deliverable has been structured in the following sections:

Section 2 describes the ACROSS context which includes the European Initiatives and legislation that affects the definition and deployment of the so called "ACROSS Personal Data Framework", and the related initiatives in the pilot countries.

Section 3 includes the data governance framework for data sovereignty requirements. First a section on the inputs gathered from other WPs and Deliverables applicable to the data governance framework is presented. Next, the initial version of the data governance framework functional requirements is defined. This version is considered initial because it will be refined and extended based on the final results of WP2 and WP6.

Section 4 "Data Governance Framework Initial Design" provides a summary of the architecture modules defined in deliverable 5.1 that are part of the data governance framework. This section presents the design of the data governance platform as an independent system.

Section 5 focuses on the Data governance model for handling data access rights describing the current the data model standards applicable to the data governance framework, analysing their suitability for ACROSS and the adaptations/extensions needed.

Section 6 includes a description of the existing and emerging technologies related to the ACROSS data governance framework.

Finally, some conclusions are drawn together with recommendations for future work.

2 Across Context

This section describes the European context that has influenced the definition of the ACROSS data governance framework. It includes the general European initiatives and legislation and the pilot's specific situation.

2.1 European initiatives and legislation

2.1.1 Data governance act

The **Data Governance Act** (**DGA**) [1] is a legislative proposal of the European Commission that aims to create a framework which will facilitate data-sharing, fostering the availability of data for use by increasing trust in **data intermediaries** and by strengthening data-sharing mechanisms across the EU. The instrument addresses the following situations:

- Making public sector data available for re-use, in situations where such data is subject to rights of others.
- Sharing of data among businesses, against remuneration in any form.
- Allowing personal data to be used with the help of a **'personal data-sharing intermediary'**, designed to help individuals exercise their rights under the General Data Protection Regulation (GDPR).
- Allowing data use on altruistic grounds.

According to the DGA **a Personal data sharing intermediary** is a specific category of data intermediaries which provides data sharing services to data subjects in the sense of Regulation (EU) 2016/679 (GDPR: General Data Protection Regulation).

Such providers focus exclusively on personal data and seek to enhance individual agency and the individuals' control over the data pertaining to them. They would assist individuals in exercising their rights under Regulation (EU) 2016/679, in particular managing their consent to data processing, the right of access to their own data, the right to the rectification of inaccurate personal data, the right of erasure or right 'to be forgotten', the right to restrict processing and the data portability right, which allows data subjects to move their personal data from one controller to the other.

In this context, it is important that their business model ensures that there are no misaligned incentives that encourage individuals to make more data available for processing than what is in the individuals' own interest. This could include advising individuals on uses of their data they could allow and making due diligence checks on data users before allowing them to contact data subjects, in order to avoid fraudulent

practices. In certain situations, it could be desirable to collate actual data within a personal data storage space, or 'personal data space' so that processing can happen within that space without personal data being transmitted to third parties in order to maximise the protection of personal data and privacy.

The ACROSS Personal Data Governance Framework is perfectly aligned with the **personal data-sharing intermediary** concept.

2.1.2 GDPR: General Data Protection Regulation

The GDPR [2] is one of the main drivers for defining the ACROSS Data Governance Framework. The analysis of the GDPR in relation with the ACROSS Data Governance Framework has been included in "D3.6 Legal requirements". The data governance framework technical requirements gathered from the GDPR analysis have been included in section 3.1.2.

As a summary, ACROSS Data Governance Framework can unambiguously be qualified as a Personal Information Management System (PIMS). As described in greater detail in a 2021 Tech Dispatch published by the European Data Protection Supervisor[4] "*PIMS are products and services that help individuals to have more control over their personal data. PIMS enable individuals themselves to manage and control their online identity.*

The PIMS concept offers a new approach in which individuals are the "holders" of their own personal information. PIMS allow individuals to manage their personal data in secure, local or online storage systems and share them when and with whom they choose. Individuals would be able to decide what services can use their data, and what third parties can share them. This allows for a human centric approach to personal data and to new business models, protecting against unlawful tracking and profiling techniques that aim at circumventing key data protection principles.

A basic feature of a PIMS is providing ***access control and an access trail****. Individuals, service providers and applications would need to authenticate to access a personal storage centre. This enables individuals to track back who has had access to their digital behaviour. Individuals are able to customize what categories of data they want to share and with whom. Other usually common elements of PIMS are secure data storage, secure data transfers (transporting data safely between systems and applications) and data-level interoperability and data portability".*

From a data protection compliance perspective, the Tech Dispatch highlighted seven key data protection priorities. The following will be covered by ACROSS:

- **Individual empowerment plus data protection by design and by default** - Help data controllers to implement the obligations of privacy and data protection by design and by default and to support them to demonstrate compliance with the GDPR.
- **Consent management** – ACROSS Data Governance Framework will rely on users' consent. Individuals would keep full control and would be free to share their personal data according to their own preference and delete them whenever they want.
- **Transparency and traceability –** ACROSS will allow for transparency both at the level of shared policies and by technical design, disclosing what services are processing which data for what specific purposes. Information can be given in real time. Personal data dashboards can help individuals to follow their data and their processing.
- **Exercise of individual's rights of access, to rectification and erasure or "right to be forgotten" –** This is a tentative functionality still to be analysed. It would provide features for individuals to be able to access their personal data, as well as to rectify or erase them, as provided for by the GDPR.
- **Data minimisation –** ACROSS will support data minimisation techniques (e.g. attribute-based credentials), to ensure that third parties can access only necessary pieces of information, thus avoiding the disclosure of the full identity of the individual.

These principles can also be found in the 2020 Opinion 9/2016 on Personal Information Management Systems from the European Data Protection Supervisor [5], which supervises data protection compliance and practices by EU institutions. The general priorities and requirements were largely the same, although the Opinion added one further component that will be analyzed in ACROSS:

- **Transfer controls –** PIMS may also help empower users to decide for themselves how far they wish to share their data geographically. Depending on the specifications of the individuals concerned, as gatekeepers, PIMS may help ensure that data will travel only insofar as the individual wishes it to do so. This kind of policy is covered by the IDS data usage policy model, one of the ACROSS base technologies.

2.1.3 SDGR: Single Digital Gateway Regulation and OOP: Once-only principle integration

The single digital gateway[3] facilitates online access to information, administrative procedures, and assistance services that EU citizens and businesses may need in another EU country.

[3] https://ec.europa.eu/growth/single-market/single-digital-gateway_en

By the end of 2023, Your Europe will offer access to 21 online procedures in all EU countries, with procedures such as registering a car or claiming a pension being fully digitalised and eliminating the need for paperwork. The most important administrative procedures for cross-border users will be fully available online in all EU countries.

In order to further facilitate the use of online procedures, SDGR will, in line with the 'once-only' principle, provide the basis for the creation and use of a fully operational, safe and secure technical system for the automated cross-border exchange of evidence between the actors involved in the procedure, where this is explicitly requested by citizens and businesses.

According to the SDGR, "*where the exchange of evidence includes **personal data**, the request should be considered to be explicit if it contains a freely given, specific, informed and unambiguous indication of the individual's wish to have the relevant personal data exchanged, either by statement or by affirmative action*". Therefore:

- Evidence may only be exchanged through the technical system based on the **prior request** of the user, i.e. the user must ask for evidence to be exchanged between competent authorities. It is thus not possible for authorities to exchange information without the user's consent, even if this would be in the public interest. Exceptions can exist where there is specific legislation that allows exchanges without any prior request.
- Evidence may only be exchanged through the technical system after the user has been able to **preview** the evidence, i.e. the user must be able to see evidence before it is sent to a competent authority, and can then decide whether they wish to proceed or not. In that way, the user can verify the accuracy of the evidence, and can also determine whether exchanging it is in their best interest.

2.1.4 Authentication and authorization: Digital identity

Identification and authentication are a cross cutting – i.e. non-sector specific and non-use case specific – requirement. In order for the ACROSS platform to be useful, the users have to be identifiable. The definition of a European digital identity and a common authentication and authorization framework applicable to public and private services in the EU is one of the main obstacles to the realization of the so-called "European Single Market". In the next sections, the current European initiatives regarding electronic IDentification, Authentication and trust Services are presented, along with their relation with the ACROSS approach.

2.1.4.1 eIDAS: electronic IDentification, Authentication and trust Services[4]

The aim of the eIDAS regulation [6] was to allow all EU citizens access to public services across the EU using means of electronic identification (eID) issued in their home country. It sought to enhance trust in electronic transactions in the internal market by providing a common foundation for secure and seamless electronic interaction between citizens, businesses and public authorities, thereby increasing the effectiveness of public and private online services, electronic business and electronic commerce in the EU.

Summarizing, the eIDAS Regulation:

- ensures that people and businesses can use their own national electronic identification schemes (eIDs) to access public services available online in other EU countries.
- creates a European internal market for trust services by ensuring that they will work across borders and have the same legal status as their traditional paper-based equivalents.

However, a recently published report from the European Commission[5] concludes that, even though, overall, the eIDAS Regulation has contributed positively to the further development of the Single Market and has provided the foundations for the development of an identity and trust services market in the EU, it would deserve a number of improvements in terms of effectiveness, efficiency, coherence and relevance.

Some of the weaknesses found in the eiDAS regulation implementation till now are directly related to the ACROSS Personal Data Governance Framework:

- The implementation of the current eIDAS system does not allow the user to actively enforce the GDPR principles of data minimisation and privacy by default and to control which data to share and with whom.
- Some key barriers to uptake by users and private sector service providers have prevented the regulatory framework to reach its full potential. Despite introducing references to eIDAS solutions in a number of sectoral EU legislation, the eIDAS Regulation has not yet replied to the needs of specific sectors (e.g. education, banking, travel, aviation).

In this context, a new European Initiative has been defined, the “European Digital Identity” which will extend and complement the current eIDAS regulation.

[4] https://digital-strategy.ec.europa.eu/en/policies/eidas-regulation
[5] https://digital-strategy.ec.europa.eu/en/library/trusted-and-secure-european-e-id-evaluation-regulation

2.1.4.2 European Digital Identity[6]

In order to identify and authenticate yourself, European citizen mostly need their personal ID. According to EU sources, the European Commission is convinced that a link is needed between the public sector (eIDAS) and private providers since existing eID solutions seem to be too inflexible in many countries. Furthermore, ID solutions "offered by social media operators and financial institutions [...] also raise privacy and data protection concerns"[7]. Ursula von der Leyen, current President of the European commission, underlines the necessity of such an identity that people trust, and its underlying technology is aimed to be used for any life occasion:

> *"Every time an App or website asks us to create a new digital identity or to easily log on via a big platform, we have no idea what happens to our data in reality. That is why the Commission will propose a secure European e-identity. One that we trust and that any citizen can use anywhere in Europe to do anything from paying your taxes to renting a bicycle. A technology where we can control ourselves what data is used and how."*
>
> **Ursula von der Leyen, President of the European Commission, in her State of the Union address, 16 September 2020**

[8]

The European Commission is proposing to amend the eIDAS Regulation to oblige Member States to issue digital identity wallets, defined as a service that will allow users to store data, credentials and attributes linked to their identity and either share them with third parties who request them, or use them to identify themselves online or offline, all at European level.

One of the main goals of this measure is to guarantee users' control over the data that identify them, in line with other European laws, such as the General Data Protection Regulation (GDPR).

The idea of this identity wallet is a very important step in creating a European ecosystem for the use and management of digital identities that may be used in all sectors and by multiple providers. The future wallet will also help to preserve users' privacy, giving them full control over the attributes that shape their identity and minimising the personal data to be exchanged. reducing dependence on third parties that could track user activity.

Hence, a European Digital Identity is supposed to be developed in order to be used by European citizen for a plenty of services, for example the following:

[6] European Digital Identity | European Commission (europa.eu)
[2] European Digital Identity | European Commission (europa.eu)
[8] European Digital Identity | European Commission (europa.eu)

- public services such as requesting birth certificates, medical certificates, reporting a change of address
- opening a bank account
- filing tax returns
- applying for a university, at home or in another Member State
- storing a medical prescription that can be used anywhere in Europe
- proving your age
- renting a car using a digital driving license
- checking in to a hotel

Figure 1 Example of European Digital Identity use case: Apply for a bank loan.[9]

The new European Identity initiative would complement the ACROSS data governance framework functionalities adding a global authentication and authorization framework within the EU, for both public services, a personal data secure storage platform and a data minimization framework.

2.1.4.2.1 German example

In Germany, ID wallets[10] are also taken into quite serious consideration. Various initiatives and projects deal with creating digital ways of digitally identifying and authenticating citizen. They also offer opportunities to store identities such as personal IDs as well as driver licences.

[9] European Digital Identity | European Commission (europa.eu)
[10] Digital Identities Ecosystem (digital-enabling.eu) (In German)

The very first version that went live and accessible for German citizen was the German ID wallet app on 23rd of September 2021. Immediately, there was a high level of public interest. After upcoming tests and further developments, the ID Wallet will be available again in the App Stores in a few weeks.

Both ID wallet approaches, the European as well as the German one, need to comply with the eIDAS regulation from 2016[11]. It contains regulations such as qualified electronic signature that is recognized in all other Member States. It lists minimum requirements (security levels, interoperability, etc.) for electronic identification services, so that they are recognized across countries.

2.1.4.2.2 How does it relate to ACROSS?

The idea of having an ID solution which is recognized across countries applies to this project, ACROSS. In case ACROSS is going to offer a central platform or a single point of contact for citizen, who tend to study or work abroad within the European Union, ACROSS might serve as gateway to all relevant services and information – supported by a central identity and access management like an identity provider such an European wallet. Within that wallet or data governance framework (that is how ACROSS names it) a citizen possibly can see and control their data. That means, which authority does forward the personal data to whom, which specific dataset is sent around and most importantly, which data should not be shared to certain institutions. The following promoted attribute and benefit of the European Digital Identity supports this: "A simple and safe way to control how much information you want to share with services that require sharing of information"[12].

2.1.5 Interoperability solutions for public administrations, businesses and citizens: ISA²

The Interoperability solutions for public administrations, businesses and citizens [13] (ISA²) Programme supported the development of digital solutions that enable public administrations, businesses and citizens in Europe to benefit from **interoperable cross-border and cross-sector public services** between 2016 and 2020.

ISA² main achievements consist in the support to the implementation of EU policies and actions through interoperability solutions, the facilitation of the re-use of interoperability solutions, and the contribution to the promotion of a holistic approach to interoperability in the EU.

Another important result is JOINUP, a collaborative platform created by the European Commission and funded by the European Union via the ISA² programme. It offers several services that aim to help e-

[11] EUR-Lex - 32014R0910 - EN - EUR-Lex (europa.eu)
[12] European Digital Identity | European Commission (europa.eu)
[13] https://ec.europa.eu/isa2/

Government professionals share their experience with each other and to support them to find, choose, re-use, develop and implement interoperability solutions.

Even though ISA[2] does not deal explicitly with personal data treatment some of the data models, APIs and tools can be applicable to the ACROSS Personal data framework, specifically those included in the **Catalogue of services** action.

The objective of this action is enhancing **the interoperability and the exchange of information about public services** within and across EU countries and at the European level by:

- Helping public administrations sharing machine-readable description of public services by maintaining a **common data model CPSV-AP**.
- Maintaining and improving reusable **tools** for the creation, validation, management and exchange of public service descriptions.
- **Piloting** the adoption and use of the CPSV-AP and tools together with EU countries and European Portals.
- Providing technical support to EU countries in setting up user-centric **one-stop-shops** for public services.

The **Service Catalogue** is one of the modules needed by the Personal Data Framework and the core vocabularies and the Common Public Service Vocabulary Application Profile will be used as a base for the service model of the Personal Data Governance Framework.

2.2 Personal data governance initiatives by country

2.2.1 MyData hubs concept

MyData[14] is a global network with the mission to achieve the vision for a type of personal data management that is human, people centric, ethical and fair for society as a whole. Local MyData data hubs at the national, regional or city level are instrumental vehicles to realise the MyData vision. In practice a Local Hub is a group of MyData Global members operating in the city, region or country level with the purpose to promote MyData mission locally.

MyData local hubs are currently operational around the world, engaging to varying degrees with promoting the principles of the MyData declaration in public events, working with technical and legal stakeholders.

[14] https://mydata.org/

In the context of the ACROSS project the active local MyData hubs in Germany and Greece could work either at the project level or with individual pilot country partners, in order to use the IHAN project test-bed[15] as a pilot. IHAN is a project that enables the development of services based on data sovereignty, sharing the objectives of Gaia-X[16].

2.2.2 Germany

The local hub of MyData in Germany operates as an informal community hub aiming to promote the principles of the declaration with awareness raising and knowledge exchange activities as well as with international networking and joint activities with the global MyData community.

Local representatives of the Germany MyData hub have developed a trusted MyData operator called Comuny. Comuny simplifies identity management by delivering verified data as-a-Service into any customer services of companies in regulated markets. It makes important customer information available for any fraud-free digital use across processes and providers, in real time, securely and conveniently.

In addition to the MyData hub, Germany also offers one central initiative that aims at displaying of which data flows from where to where + query of data. Through this, the German registries are supposed to become modernized and interconnected. The “Datenschutzcockpit” (English: Data protection cockpit) is a meant to be a nationwide service, which is legally determined by the ‘Register Modernization Act’ by the Register Modernization Authority (Federal Administrative Authority). With this initiative, 50 registries will be modernized. The goal is to build transparency and trust among all participants.

Through a front-end application, the users can access and read all their data that were transferred from one authority to another (within the last two years). The registration works by using the identity and access management system of the eID (server). Most of the end users utilize it via the mobile app “AusweisApp2” by Governikus GmbH & Co. KG. Within this system, the German tax ID serves as one unique and central number that links all personal information back to the person who the data belong to. In order to make this technically possible, an own standard was established for this which is called ‘xBasisdaten’ that builds on ‘XÖV’ (standard within the public sector). The routing takes place via the DVDV (“Diensteverzeichnis der öffentlichen Verwaltung” = Service directory of the public administration – the ‘telephone directory of the German administration’) and its protocol OSCI (technical protocol standard for public administration). A description of the planned architecture in Germany is illustrated below (only in German):

[15] https://www.sitra.fi/en/projects/testbed-for-fair-data-economy-ihanfi/
[16] https://www.data-infrastructure.eu/GAIAX/Navigation/EN/Home/home.html

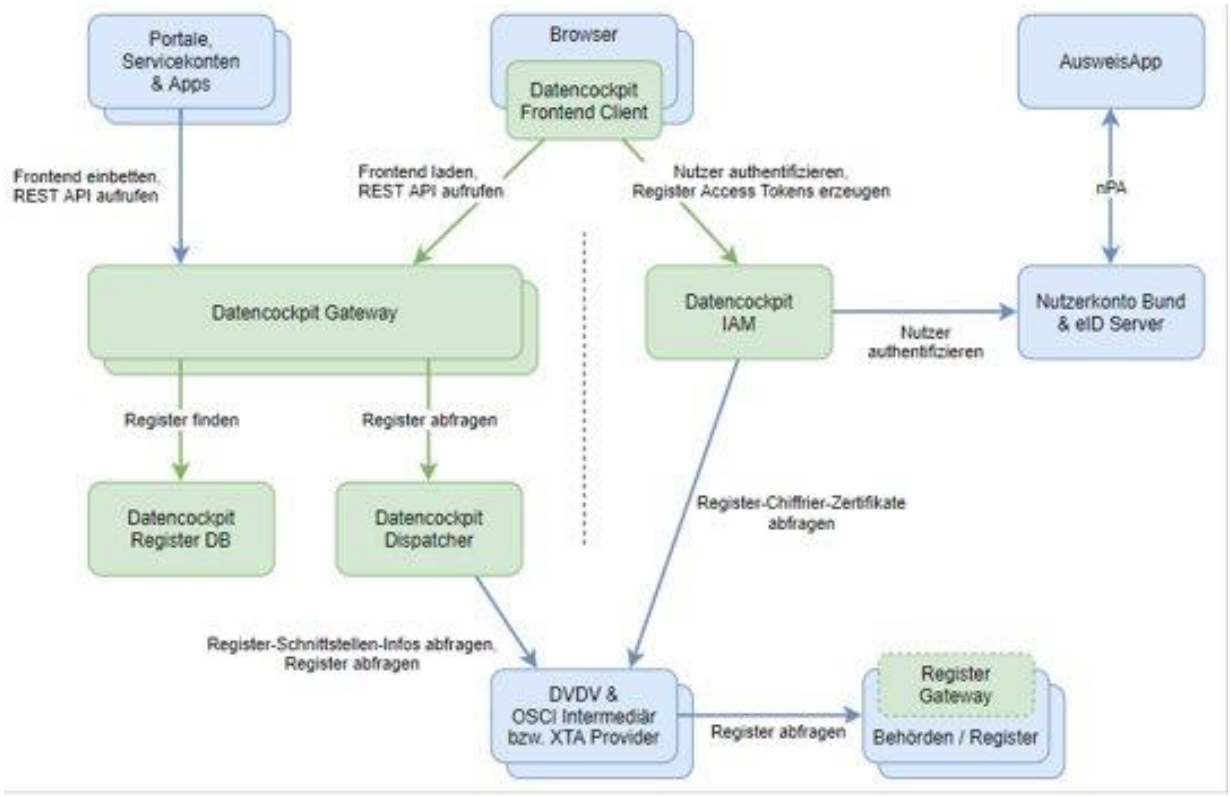

Figure 2 German Data Protection Cockpit architecture

However, the data protection cockpit does not fully deal with the once-only-principle (OOP) yet. The ideal way of entering an address only once and then reuse it is not possible here. In order to do so, the data protection cockpit most likely will be complemented with a consent module in order to enable users to control who receives their data once they have entered them into any public system.

The roadmap states that there will be a version live in 2022 which will be firstly piloted in Bremen, one of the 16 federal states in Germany. Finally, the German “Nachnutzung” (‘after use’) determination ensures that the service will be rolled out to other federal states afterwards. The following roadmap shows the planned milestones and its dates:

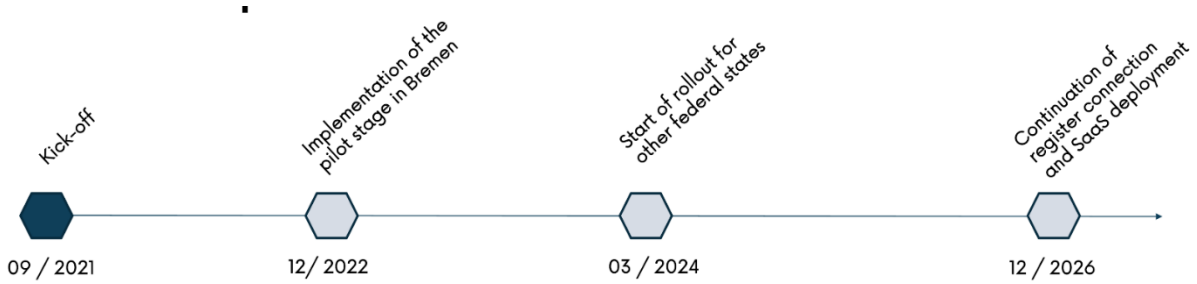

Figure 3 German Data Protection Cockpit roadmap

2.2.3 Latvia

Currently there are no significant public (citizen groups, think tanks, NGOs, and other) initiatives regarding the definition of a central point for the citizens to define the personal data treatment preferences in Latvia. However, there is already ongoing work on government data management architecture which will support data management tools for users. Data management system (DAGR) is supposed to be initially released until the end of 2023 and will be further developed to support user control over sharing their personal data, especially in cases when legislature does not set regulatory principles on sharing data between public institutions. Even more significant this tool will be when private institutions (for example, commercial organisations) will join this infrastructure which is planned in the future. Data sharing audit will be provided to users as journaling mechanism with all data usage records. It is also planned to integrate an option to transfer personal data from one system to another with user's consent.

2.2.4 Greece

In Greece the local MyData hub also operates as an informal community, aiming to foster and accelerate the needed changes towards establishing a human-centric approach to personal data in Greece. The ACROSS partner GFOSS as a member of MyData coordinates the activities of the local hub. Local Initiatives of the hub bring together experts from business, legal, tech and society sectors in order to discuss themes such as rebuilding trust for human-centered data economy in Greece, putting individuals in control of their data, compliance with the principles laid out both in the GDPR and by the MyData movement and foster entrepreneurship and customer value through the ethical use of personal data.

Other topics tackled within the local activities of MyData Greece include personal data protection, democracy and data, ID and authentication, system architecture for the data reuse, AI and ethics of data use, overseas policy developments, data portability.

3 Data Governance framework for data sovereignty requirements

3.1 Inputs from other WPs

3.1.1 Pilot and co-creation process inputs: user journeys

3.1.1.1 Methodology

WP2 and WP6 together have defined the so-called user journeys based on the results several interviews with people from the three pilot countries. The aim of the interview process is form potential user journeys, building on initial ideas. User journeys can include actions, touch points, emotions, pain points, and phases. Eventually to result in concrete (socio-technical) requirements for the ACROSS platform modules. A specific section about Data privacy issues has been included in the questionnaire in order to gather requirements for the Data Governance Framework.

This is the Questionnaire used for gathering the requirements regarding personal data and GDPR.

> *One of the ACROSS objectives is to develop a Private/Personal data management web application that allows citizens to:*
>
> - ***Control** how their private data are collected, created or used by businesses, governments, or data brokers, giving individuals greater power to determine how their data can be used.*
> - ***Monitor** which data are available and how they are used or how it has been accessed*
>
> *Are you interested in knowing and controlling who and how your personal data is used in each of the public and private services that you are using?*
>
> *Would you mind if public administrations share your personal data with other public administrations, inside your own country or internationally? And could they also share your data with non-profit private services, like schools, universities, etc, or even with commercial companies?*
>
> *The European legislation on data privacy grants the citizen a series of rights regarding the use of their personal data. What functionalities would be the most important for you, i.e. which of the following rights are the most important to support in a data management application?*
>
> 1. ***The Right to Information**: obtain information about the processing of your personal data*
> 2. ***The Right of Access**: Obtain access to personal data concerning you*
> 3. ***The Right to Rectification**: request that incorrect, inaccurate or incomplete personal data be corrected*
> 4. ***The Right to Erasure**: request that personal data be erased when they are no longer necessary*

5. ***The Right to Restriction of Processing****: request the limitation of the processing of your personal data in certain cases*
6. ***The Right to Data Portability****: receive your personal data in a machine readable format and send it to another data controller.*
7. ***The Right to Object****: oppose the processing of your personal data for marketing purposes or for reasons related to your particular situation*
8. ***The Right to Avoid Automated Decision-Making****: request that decisions based on automated processing that concern you or significantly affect you and based on your personal data be made by individuals and not only by computers; Likewise, you have the right, in this case, to express your point of view and to challenge the decision.*

In addition, should the use of your data only be allowed:

- *With your consent (i.e. when you actively allow your data to be accessed and used, on a case by case basis)?*
- *When the law requires this (i.e. when there is legislation that allows someone, such as a public administration, to access and use your data)?*
- *When this is in the public interest (i.e. when there is a clear benefit to society, such as access and use of your data for scientific research)?*

Next sections present the results and conclusions gathered from the survey in the three pilots.

3.1.1.2 German pilot

In general, the participating interviewees within the German user journey research did state that they feel good when they know the homepage provider, where they type in their personal data. For instance, if the homepage is hosted by an official and trust-worthy institution like public authorities. "As long as I know that it is a public page provided by EU or the state, I feel safe".

Once there occurs an unknown third-party provider it needs to be credited by “a certain reason”. Otherwise, users won’t trust this service and do not give the data.

DATAPORT User/Customer Journey for WORKING ABROAD

Driven by the fact that ACROSS might be officially represented by the European Commission or the European Union, the "single-point of access” or the “central platform” that interviewees thoroughly desired, has a valid chance to be accepted.

Especially interviewees that went for working abroad strived for a “transparent data cockpit where you see who has accessed your data, you sent your data to whom etc.”. The “data sovereignty & transparency is very important” to the interviewees.

DATAPORT User/Customer Journey for STUDYING ABROAD

As long as the student know “what happens” to their data, it seems to be adequate - “as long as I know it”. If this would be combined with their active consent, according to the interviewees, the data exchange, especially between authorities and universities, is widely accepted. "Data Governance should include an active decision to know what you release your data for, so that you can control who uses your data”. This general perspective goes in line with the approach of ACROSS to offer a central tool where the data exchange is not only displayed transparently, but can also be controlled by the citizen. In one case it happened that personal data such as the address even got published on the internet accidentally – years after studying abroad - which felt very disappointing and uncertain for the student.

One fact that also stands out immediately is the question why the cross-border systems are not better connected yet. If so, the citizen would appreciate it – the interviewees are commonly sure about this. “I was not sure if health data from abroad was transferred to my home health insurance. In the end, it worked. But no idea, how. Such a pity, I’d have liked to know that.”

Another distinct topic is not only the overview of all data, but rather “an overview of the process status”. According to the research, it was highly annoying for students to have submitted a form, but not to be informed about the upcoming steps and the status quo.

DATAPORT Conclusion

To sum up, ACROSS does have the chance to serve as a tool that addresses those requirements. When ACROSS provides a data governance that gives the power to the citizen who can see as well as actively confirm the data exchange between authorities across countries, ACROSS might add much more value to real life scenarios of studying and working abroad within the European Union. But it needs to fulfill requirements of transparency, interconnection, central access and trust-worthy provision.

3.1.1.3 Latvian pilot

Respondents from questionnaire used in creation of user stories for ACROSS were asked to evaluate importance of various rights regarding data processing. Interviews revealed that overall people care about data processing, and the main issue is trust in institutions. General trust in institutions facilitates overall acceptance of data processing practices regardless their type. However, respondents noted that data sharing with third parties, especially with private (commercial) sector, must happen only with consent, thus calling for convenient data management tools and practices.

Evaluating different rights provided by GDPR regulation the most important rights for respondents surveyed by Latvian pilot was right to rectification and right to object. Respondents were motivated to

name these rights as the most important because incorrect data cannot be considered as actual personal data, thus affecting access to digital services. Also, respondents were strongly against data usage for commercial purposes. Respondents cared the least about right to erasure. Some of them motivated it with argument that they do not believe in full data erasure once data have been in digital environment, thus caring less for such right and also actual need to use it in their personal situation.

Overall, data sharing among institutions is welcome, especially while accessing digital public services and receiving health services. One-stop personal authorization option to access wide array of services is preferable over case-by-case authorization. Respondents are keen to share personal data in cases when they need to access digital services, thus extra barriers on data sharing among institutions or through one platform are unnecessary as long as users are given control over personal data portfolio.

3.1.1.4 Greek pilot (GRNET)

The participants of the Greek co-creation process have been interviewed and questioned about their beliefs on the management of their data in terms of the pilot and their expectations from a Data Governance framework. As a bottom line their main requirement relies on being aware of who and how is using their personal data in each service and for requesting their consent for getting access to their personal data. They highly appreciate the Right of Information and the Right of Access, as well as the Right to Rectify personal information that is inaccurate or incomplete. They also appreciate the right to erase their data when they are no longer required. It is also important for them to be able to avoid automated decision making from their data and ask for the right to express their own point of view or objection in each individual decision. The table that follows shows the distribution of answers in each question that has been set to them.

Table 1 The answers to the DG framework requirements from the Greek co-creation lab

Questions on Privacy	No	Yes	Only with consent	N/A
Are you interested in knowing and controlling who and how your personal data is used in each of the public and private services that you are using?	10%	60%	0%	30%
Would you mind public administrations share your personal data with other private services, like schools, universities, etc?	10%	20%	30%	40%

The European law on data privacy grants the citizen a series of rights regarding the use of personal data. <u>What are the most important for you?</u>				
The Right to Information: obtain information about the processing of your personal data	0%	70%	0%	30%
The Right of Access: Obtain access to personal data concerning you	0%	70%	0%	30%
The Right to Rectification: request that incorrect, inaccurate or incomplete personal data be corrected	0%	60%	0%	40%
The Right to Erasure: request that personal data be erased when they are No longer necessary	0%	50%	10%	40%
The Right to Restriction of Processing: request the limitation of the processing of your personal data in certain cases	20%	40%	0%	40%
The Right to Data Portability: receive your personal data in a machine readable format and send it to another data controller	40%	30%	0%	30%
The Right to Object: oppose to the processing of your personal data for marketing purposes or for reasons related to your particular situation	10%	60%	0%	30%
The Right to Avoid Automated Decision-Making: request that decisions based on automated processing that concern you or significantly affect you and based on your personal data be made by individuals and Not only by computers; Likewise, you have the right, in this case, to express your point of view and to challenge the decision.	0%	70%	0%	30%
In addition, <u>should the use of your data only be allowed</u>:				
With your consent (i.e. when you actively allow your data to be accessed and used, on a case by case basis)?	0%	60%	10%	30%

When the law requires this (i.e. when there is legislation that allows someone, such as a public administration, to access and use your data)?	10%	30%	0%	60%
When this is in the public interest (i.e. when there is a clear benefit to society, such as access and use of your data for scientific research)?	10%	10%	20%	60%

3.1.2 Analysis of GDRP issues

As a result of the effort of the work that has been carried out in deliverable "***D3.6 Legal requirements***", it has been possible to extract a summary of principles that has been carried out in the different policy areas such as:

- Privacy and data
- e-Government and public services
- Identification and authentication
- Governance and sovereignty

The information presented in the following table has been obtained from the section "5.2 Statement of legal compliance principles in ACROSS" and in the next section of this deliverable, "3.2 Data Governance Framework functional requirements", are going to identify them as a functional requirement.

Identifier	Description
DPP-01	ACROSS is built on the primacy of consent of the individual user. Users should be able to choose which data is available through the ACROSS infrastructure, to whom, and for what purpose. No exploitation (including commercialisation or direct marketing) of user data should occur without their consent.
DPP-02	ACROSS supports **privacy enhancing technologies**, including by supporting and promoting pseudonymous information exchanges where this meets the objectives of a specific use case.
DPP-03	ACROSS supports **access management and controls**. Users should be to grant, deny or terminate access to their data with equal ease, in a sufficiently granular manner to enable effective control.
DPP-04	ACROSS provides **transparency.** Users should be able to see who has (had) access to their data at all times, and who is responsible for complying with data protection laws. Contact information should be easily available to them at all times.
DPP-05	ACROSS supports **data subject rights.** Users should be able to access, amend, correct and/or delete their data at all times, and to be able to obtain a copy of it. Wherever possible,

	exercising these rights must be built into the architecture; users should not just be referred to a third party if ACROSS is capable of helping.
DPP-06	ACROSS supports **secure storage and exchange.** User data should not be exposed to third parties without user consent, and it should be protected against loss or corruption.
SDGP-01	ACROSS supports **user control, also towards public authorities.** User data should not be shared with public authorities without the explicit request of the user, and users should always be able to review information before it is shared with public authorities.
SDGP-01	ACROSS supports **free choice.** For that reason, the ACROSS platform should never be the only option for citizens, since that would make ACROSS mandatory in some situations, and thus no longer based on consent. An alternative should therefore always exist.
IAP-01	ACROSS supports regulated **European electronic identification schemes and European electronic identification infrastructure, in order to enable cross border transactions.** Use of such schemes is however not required to use ACROSS.
IAP-02	ACROSS supports regulated **European electronic signatures and seals, in order to enable cross border verification of the integrity and authenticity of exchanged information.** Use of such signatures and seals is however not required to use ACROSS.
GSP-01	ACROSS is governed **independently**, prioritising the interests of citizens at all times. Other stakeholder should be consulted, but ACROSS it should not be controlled or unduly influenced by the interests of service providers
GSP-02	ACROSS is governed **on a not-for-profit basis,** meaning that the platform itself should not aim to gain commercial profits or benefits from the data that it holds without consent of the citizens.
GSP-03	ACROSS supports **accessible complaints resolution**. The platform will facilitate alternative dispute resolution mechanisms, and will provide a complaints handling mechanism.

DPP: Data Privacy Principle

SDGP: Single Digital Gateway Principle

IAP: Identification and Authentication Principle

GSP: Governance and Sovereignty Principle

3.1.3 WP5 requirements

WP5 has gathered (D5.1 System Architecture & Implementation Plan – Initial) a first set of requirements for the whole ACROSS platform and ICT modules. Some of these requirements are directly related to the Data Governance Framework while others are generic IT requirements or non-functional requirements. The applicable requirements from WP5 are included in Annex I – WP5 Requirements.

3.2 Data Governance Framework functional requirements

This section describes the functional requirements of the Data Governance Framework. These requirements have been obtained considering the requirements derived from the User Journey Questionaries of WP2, the architecture of ACROSS described in D5.1 and the analysis of the GDPR issues carried out and reflected in D3.6. Regarding the latter, in the following subsections we will refer to the legal compliance principles specified in section 3.1.2 that are fulfilled by each described functional requirement.

In overall, the Data Governance Framework fulfil the following legal compliance principles described in section 3.1.2: SDGP-02, IAP-01, IAP-02, GSP-01 and GSP-02.

3.2.1 Users management

The data Governance Framework must allow an end user to create a new account or to remove it. The first time a user enters the framework, the end user will have the opportunity to create an account in the framework.

3.2.2 Data catalogue management

When defining a new Service, apart from providing its description, the description of the type and structure of the set of personal data processed by the Service must be also provided. Each dataset will be associated to a specific use purpose and specific processing ways e.g.:

- Dataset A contains: Name, Last name, gender, nationality.
- Dataset B contains: Gender, date of birth, nationality, address, phone.

Each of these datasets is used by the service for a specific purpose, is shared with specific organizations and is processed in a specific way.

The framework must allow to create, to read, to update and to remove these datasets associated to a specific service. That is, it must allow to invoke CRUD operations on these datasets.

3.2.3 Service management

A service provider will be able to:

- Create and Edit all descriptions of Services that will be integrated with the framework, according to the Service Description Data Model defined (See section 5.2).
- Get an overview and manage the lifecycle of Services Descriptions (Create, Import, Export, Register, DeRegister, Delete and so on).
- Get an overview and details of the Consents that End Users have given at corresponding registered Services, in particular:
    - Processing and Purpose details.
    - Consents history.
    - Consents hash and notarization.
    - Consents raw data (JSON).

3.2.4 Service consent management

The end user will be able to:

- Get an overview of his personal data being processed by the Services he is linked to.
- Get an overview of previously registered Services by Service Providers, and ready to be linked to his account, and of already linked Services:
- Link his account to an available Service.
- Disabling a linked Service. This will put all its active Consents (if any) in Disabled state
- Get an overview of given or pending Consents, where the following information will be provided:
    - Processed personal data
    - With which Organization data can be shared
    - Other info
- Manage the lifecycle of given Consents by changing its status:
    - Disable: disable the Consent
    - Activate: enables the previously disabled Consent or pending Consent.
    - Withdraw: revoke the Consent, a new one must be given.

  The aforementioned actions will involve the creation/modification/removal of the policy rules that will describe how the personal data should be used by the service.

- Enable or disable each single Data Concept contained in the Resource Set regulated by that Consent (e.g.: his age). This action will involve the modification of the corresponding data usage control policy rule.

These functional requirements fulfil the following legal compliance principles described in section 3.1.2: DPP-01, DPP-03 and DPP-04.

3.2.5 External APIs

The framework will expose a set of APIs to be used by the Service Provider services that are going to be integrated in the framework. This API will allow the service to check the service linking status, the consents associated to end users and to inform the framework the usage of personal data.

4 Data governance Framework initial design

4.1 Architecture

The following figure provides the overall view of the main components of the Data Governance Framework.

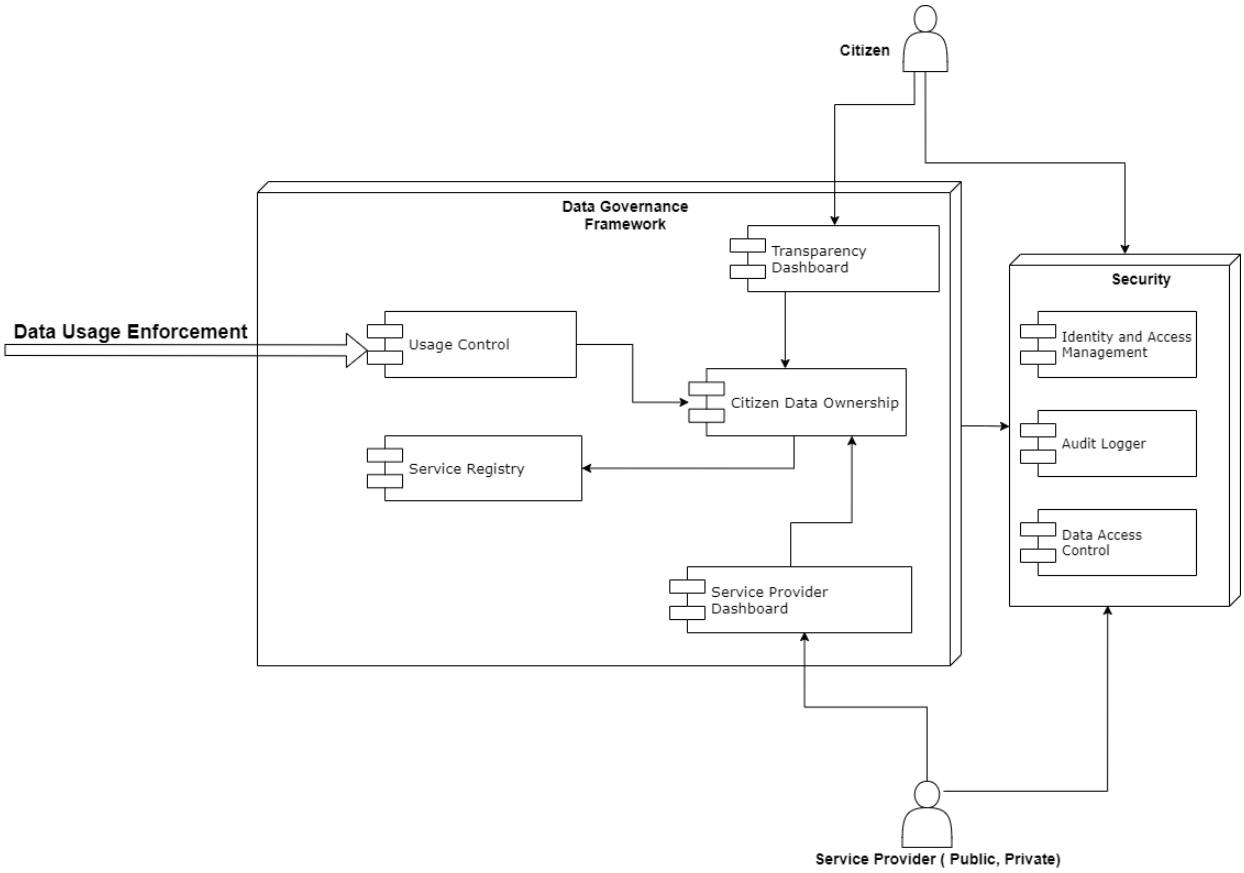

Figure 4 - Component View of Data Governance Framework

The Data Governance Framework will allow citizens / users to register in a series of services (Service Registry) and allow the use of their data based on consents that should be approved by them. To carry out this transfer of information in a secure way, the Usage Control module will be used, which will allow the usage of data based on previously defined usage policies.

The components in the Security layer will be used by all the components in the Data Governance Framework. This layer provides all the security features needed for a citizen and a service provider to be authenticated and authorized, and for logging all the interactions among all components of the framework.

4.2 Modules

4.2.1 Citizen Data Ownership

This component allows the citizens to manage their personal data and allows the organizations/services to fulfil the requirements in line with the GDPR. It will expose several interfaces for the Transparency Dashboard, so that the individuals can grant and withdraw their consents and receive notifications about how their data is being used. On the other hand, it will expose several interfaces for the services, so that they can be informed about the consents of the citizens, and they can send notifications about the data that is being used.

4.2.2 Usage Control

This component provides the enforcement mechanism to apply usage policies according to previously defined consents.

The available formats of data usage policies include GDPR consents and IDS data policies enforcement.

4.2.3 Service Registry

This component provides human and machine-readable description of services that will be available in ACROSS platform for user journey services provisioning. The registry enables the storage and publishing of service by providing general, technical and data processing information based on standard models (e.g. ISA[2]).

The component provides the following functionalities:

- Publishing, searching, and retrieving of an already available service in the platform
- Service Description versioning
- API for programmatically interaction with the registry
- User Dashboard and service editor

4.2.4 Transparency Dashboard

This module is a web application that uses a human centric approach to liberate the potential of personal data and to facilitate its controlled flow from multiple data sources to applications and services. Citizens must be able to opt-in and out from the use of their personal data, in line with the requirements of the GDPR. The main objective is to give the individual control of their own data.

The component provides the following functionalities:

- Monitor which data are available and how they are used or how it can be accessed. It provides individual's linked services, and data use related policies and consents.
- Notify users about realised data processing at services.
- Give users control over their data allowing them to add as well as delete or modify information.

4.2.5 Service Provider Dashboard

This module is a web application that allows the service providers to:

- manage the Semantic Descriptions and registrations of its own provided Services, so that it is available for the citizens through the Transparency Dashboard.
- view and manage the Service Linking and Consents status given by all the Users of its registered services.

4.3 APIs

4.3.1 Citizen Data Ownership

The Citizen Data Ownership module will expose the following interfaces:

- SearchConsent: Seach consents by different criteria.
- ModifyConsent: Modify consent status (e.g.: withdraw), enable or disable specific data to which consent applies, change organizations to which data is shared.
- ViewLogs: Show information about the events that have happened related to the linked services and the consents given/withdrawn.
- SearchServices: Search services by different criteria.
- LinkUserToService: Link a user to a service, so that he can manage the consents given to that service.

4.3.2 Usage Control

The Citizen Data Ownership module will expose the following interface:

- UsageControlEnforcement: Apply usage policies so that data is used accordingly.

4.3.3 Service registry

The Service Registry module will expose the following interfaces:

- Store/delete: Storing or delete a service description.
- Search: Search a service description according to several metadata in accordance to the adopted service model.
- Publish: The service description is active and available and searchable.

5 Personal data governance model for handling data access rights

In order to implement the Personal data governance framework a set of data models has to be defined to capture the data governance related concepts and the relationship among them.

In this section a set of existing data models and a first analysis of their applicability in the context of the ACROSS personal data governance framework is presented. The main objective is to provide an initial starting point to make possible interoperability among European Public Administrations services and also the private services regarding personal data.

5.1 ISA² Core models

Some of these models are called e-Government Core Vocabularies[17] that have been developed by ISA² for public administrations in an open process where stakeholders like e-Government Core Vocabularies Working Group, Directorate-General for Informatics: DG DIGIT - in particular the SEMIC action of the ISA² programme and the Publications Office of the EU work together.

Next, a summary of the main ISA² standard core models is presented. The information has been taken from the ISA² core vocabularies web site.

Core Vocabularies are simplified, re-usable and **extensible data models** that capture the fundamental characteristics of an entity in a **context-neutral** fashion. Public administrations can use and extend the Core Vocabularies. They can be useful in the following contexts related to ACROSS project:

- **Information exchange between systems**: The Core Vocabularies can become the basis of a context-specific data model used to exchange data among existing information systems.
- **Data integration**: The Core Vocabularies can be used to integrate data that comes from disparate data sources and create a data mesh-up.

The following core vocabularies have been selected to be part of the Personal data governance model:

- The Core Business Vocabulary
- The Core Location Vocabulary
- The Core Person Vocabulary
- The Core Public Service Vocabulary

[17] https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/solution/e-government-core-vocabularies/about

- The Core Public Organisation Vocabulary

Next, the first three models are described. The last two, are included in the Service model section (See section 5.2), since they define some concepts for the Core Public Service Vocabulary Application Profile which provides the base service model for the ACROSS platform.

5.1.1 Core Person Vocabulary

It is one of the e-Government Core Vocabulary that is a simplified, reusable and extensible data model that captures the fundamental characteristics of a person, e.g. the name, the gender, the date of birth, the location etc. The following figure shows the UML diagram of the model including the relationship with other standard models like foaf[18] or skos[19].

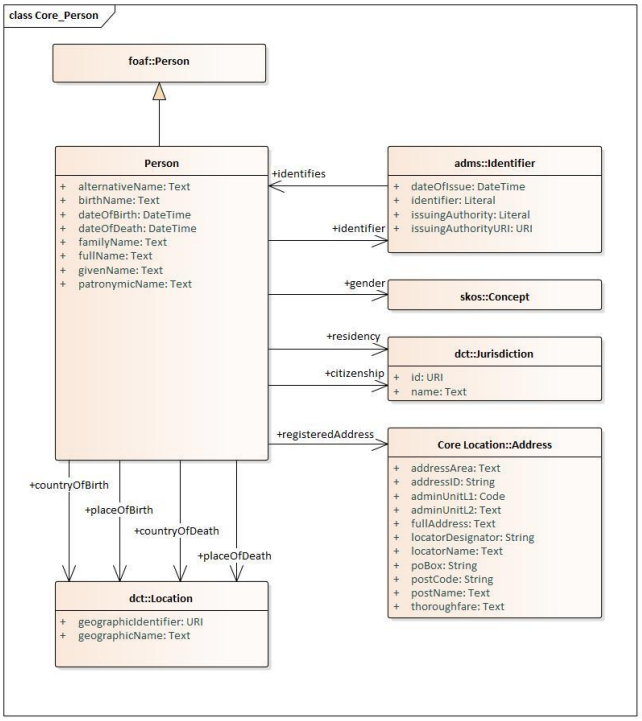

Figure 5 - Core Person Vocabulary

[18] https://en.wikipedia.org/wiki/FOAF_(ontology)
[19] https://www.w3.org/TR/2008/WD-skos-reference-20080829/skos.html

5.1.2 Core Location Vocabulary[20]

The Location Core Vocabulary provides a minimum set of classes and properties for describing a location represented as an address, a geographic name, or a geometry.

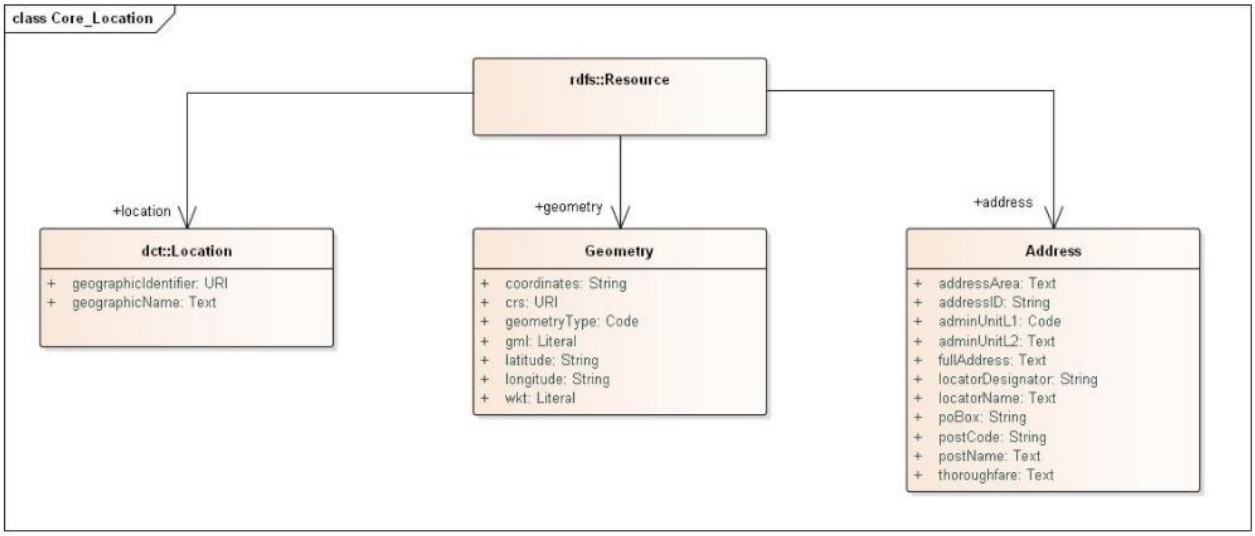

Figure 6 - Core Location Vocabulary

[20] https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/solution/e-government-core-vocabularies/core-location-vocabulary

5.1.3 Core Business Vocabulary[21]

Data model that captures the fundamental characteristics of a legal entity, e.g. the legal name, the activity, address, etc.

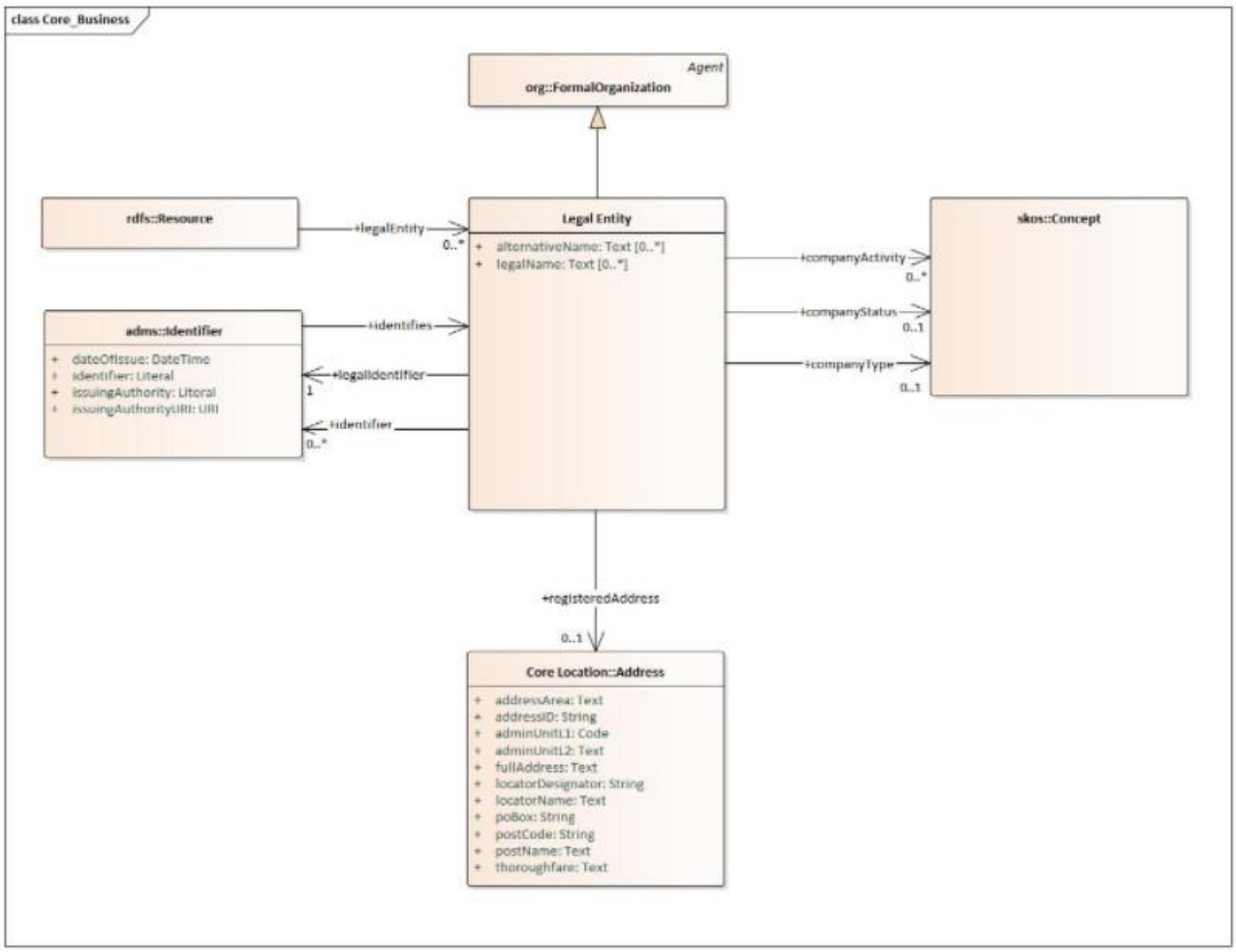

Figure 7 - Core Business Vocabulary

[21] https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/solution/e-government-core-vocabularies/core-business-vocabulary

There are some relationships are established between all of Core Vocabularies described above that it will see through next UML diagram:

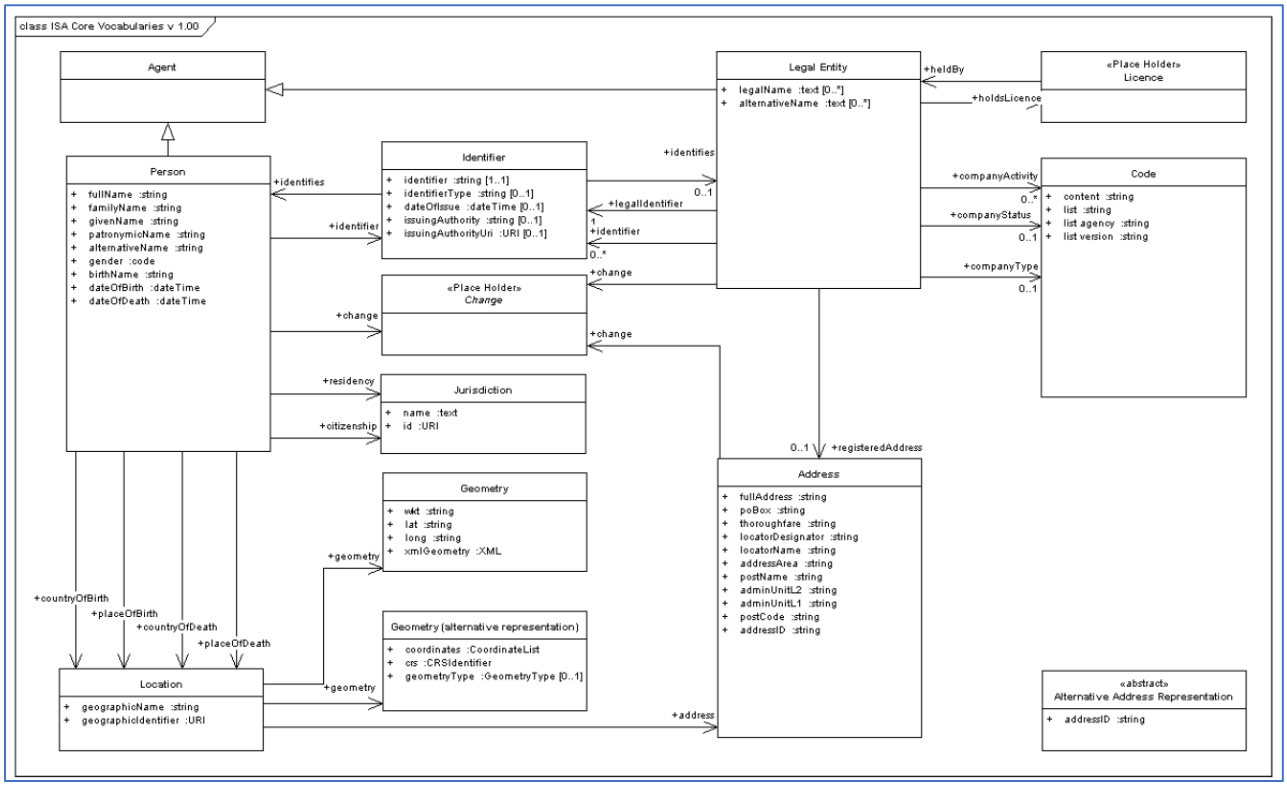

Figure 8 - Core Location-Person-Business Vocabulary

5.2 Service Model

The previous step before that analysis of a data model for describe public services, it is to know the model that is going to be use for describing public organization. ISA² define Core Public Organization Vocabulary, then the Core Public Service Vocabulary and the

5.2.1 Core Public Organization Vocabulary (CPOV)[22]

This vocabulary offers a common data model for describing public organizations in the European Union and is designed to support the exchange of basic information about individual public organizations.

22 https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/solution/e-government-core-vocabularies/core-public-organisation-vocabulary

5.2.2 The Core Public Service Vocabulary (CPSV)

CPSV is a simplified, reusable and extensible data model that captures the fundamental characteristics of a service offered by public administration. Such characteristics include the title, description, inputs, outputs, providers, locations, etc. of the public service. An application profile of the Core Public Service Vocabulary (**CPSV-AP**) has been developed for describing public services and grouping them in business events.

5.2.3 Core Public Service Vocabulary Application Profile (CPSV-AP)

The CPSV-AP[23] is a data model for describing public services and the associated life and business events. It standardises the semantics of personal milestones, including having a child, beginning education, looking for a new job, as well as professional changes such as starting or financing a company, hiring an employee. The descriptions will make data on these events structured, easier to capture and machine-readable. Public administrations and service providers can use this to guarantee a degree of cross-domain and cross-border interoperability between public service catalogues.

Next figure shows a set of tools that have been implemented supporting the application of CPSV-AP by the public administrations[24].

[23] Core Public Service Vocabulary Application Profile (CPSV-AP) | ISA² (europa.eu)
[24] https://github.com/catalogue-of-services-isa

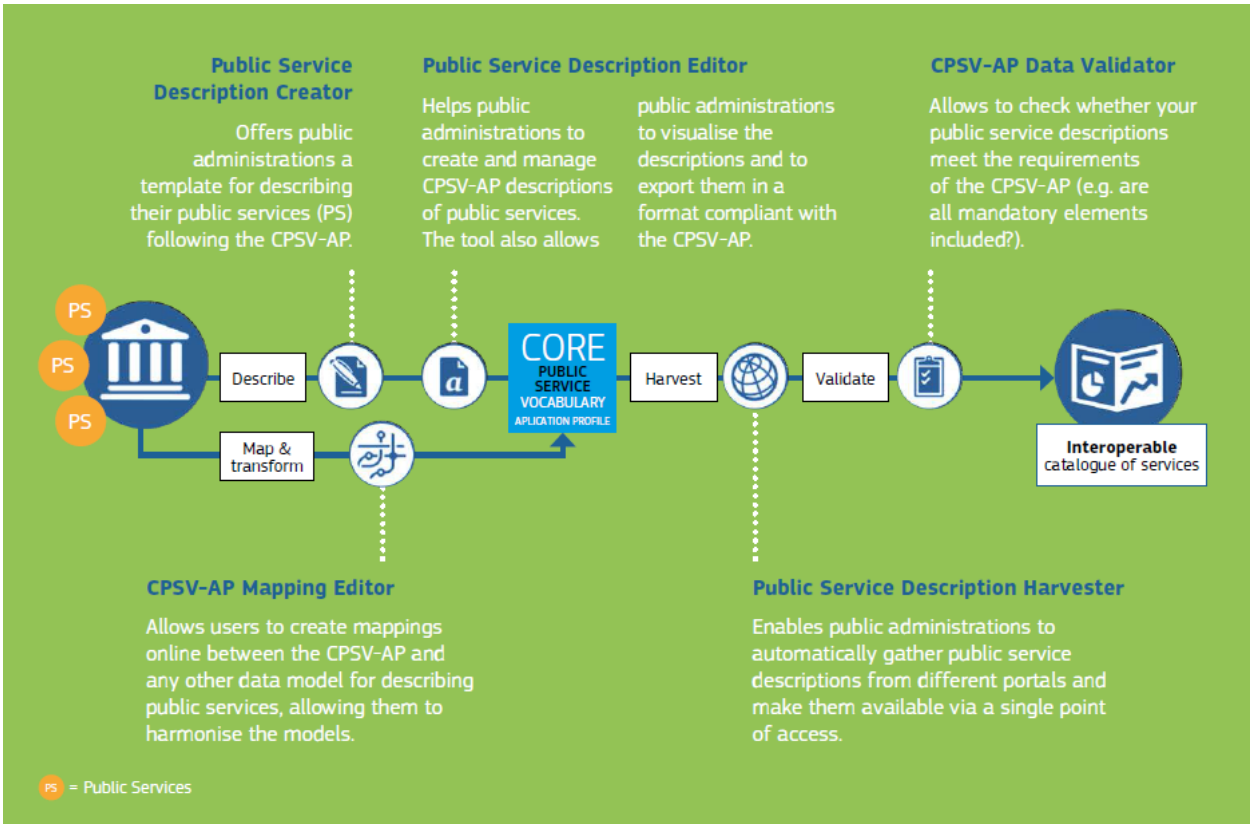

Figure 9 CPSV-AP supporting tools

Public administrations and service providers can take this model to describe their services and guarantee a level of cross-domain and cross-border interoperability at European, national and local level.

Attributes such as name, description, competent public organization are often used for the description of a public service.

The specification of the Core Public Service Vocabulary Application Profile is represented in a UML class diagram which includes:

- The classes and properties that define the service itself: the necessary inputs, possible outputs, the responsible public authority and the events that trigger service use.
- The classes and properties that describe the context in which the service is offered. This includes relevant legislation and rules of operation for the service; and
- The interface between the service and its users: how and when it can be accessed.

The following figure shows the UML diagram for CPSV-AP where it can be seen that the only mandatory classes (in blue color) are “Public Organization” and “Public Service”.

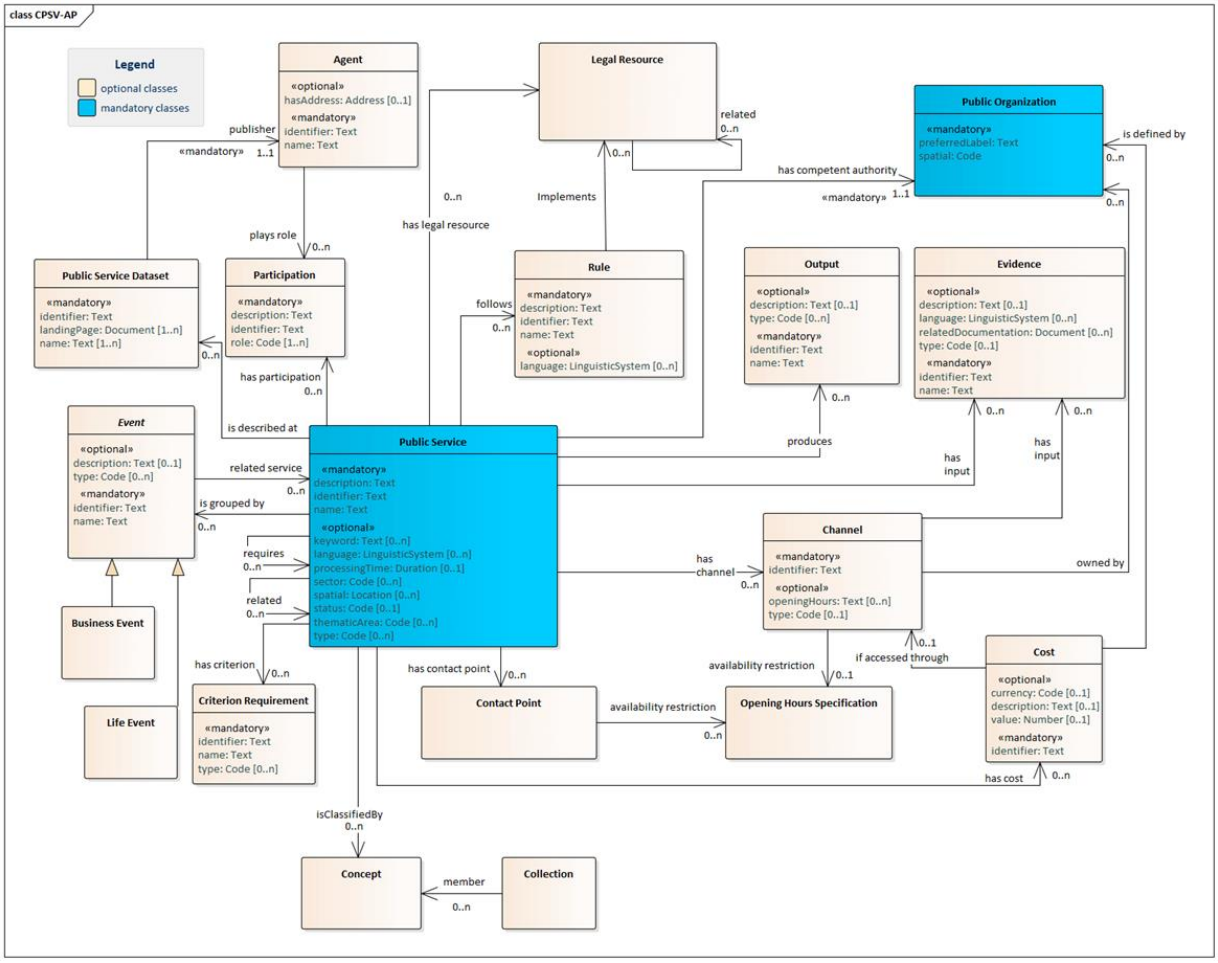

Figure 10 – CPSV-AP UML Diagram

Source: https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/solution/e-government-core-vocabularies/core-public-service-vocabulary-application-profile

The Channel class represents the medium through which an Agent provides, uses, or interacts in another way with a Public Service. Typical examples include online services, phone, walk-in centres etc. In Across to invoke to a specific public service is mandatory that public service define this property.

Link between CPOV and CPSV-AP

The Core Public Service Vocabulary Application Profile for Public Administrations in Europe (CPSV-AP) has defined an object property “has competent authority”. The relationship indicates how a public service and a formal organization, such as a public organization, are related.

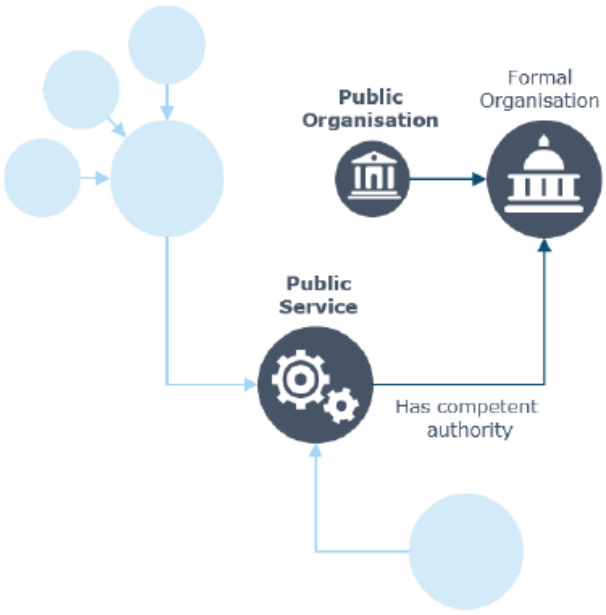

Figure 11 – Link between CPOV and CPSV-AP

5.2.4 European taxonomy for public services

The CPSV-AP defined some mandatory classes and properties to be used in any model compliant with it.

For some of its properties, the CPSV-AP recommends the user to select a value from a controlled list or taxonomy. This approach has the advantage of ensuring the consistency of the semantics between administrations.

Even though the CPSV-AP pushes for the harmonisation between catalogues of public services in Europe, many administrations follow different taxonomies. This results for example in differences in the naming of public services: multiple administrations calling the same service differently

Taxonomies and other types of controlled vocabularies are the preferred means to achieve such a common understanding by specifying the terms of the domain, disambiguating them from each other, controlling synonyms, and structuring the domain via term relationships. In the case of public services, a commonly agreed taxonomy of generic public services would help public administrations to harmonise their catalogues of services.

In the context of the Single Digital Gateway Regulation, all public administrations in the European Union have to exchange information about certain services at the European level. A common denomination between the services proposed would be a first step towards clear understanding of the services for the users.

Above paragraphs are an extract of a complete document regarding importance of a taxonomy of public services. You can obtain more detailed information in:

https://joinup.ec.europa.eu/sites/default/files/news/2019-09/ISA2_European%20taxonomy%20for%20public%20services.pdf

5.3 Data Usage Policy model

Access control allows access to certain digital resources. Usage Control allows to decide what can be done with a data asset through a series of previously defined restrictions or policies. At run time, the enforcement of data usage control prevents data from being mishandled.

Therefore, data usage control is a tool that prevents defined security mechanisms from being violated.

IDS (See Section 6.3.1 for a more detailed description) introduce a policy expression language called IDS Usage Policy Language based on a previous standard, Open Digital Rights Language (ORDL). ODRL is a policy expression language that provides a flexible and interoperable information model, vocabulary, and encoding mechanisms for representing statements about the usage of content and services.

5.3.1 Open Digital Rights Language (ODRL) Information Model

The Open Digital Rights Language (ODRL) is a policy expression language that provides a flexible and interoperable information model, vocabulary, and encoding mechanisms for representing statements about the usage of content and services. The ODRL Information Model describes the underlying concepts, entities, and relationships that form the foundational basis for the semantics of the ODRL policies.

Policies are used to represent permitted and prohibited actions over a certain asset, as well as the obligations required to be meet by stakeholders. In addition, policies may be limited by constraints (e.g., temporal, or spatial constraints) and duties (e.g. payments) may be imposed on permissions.[i]

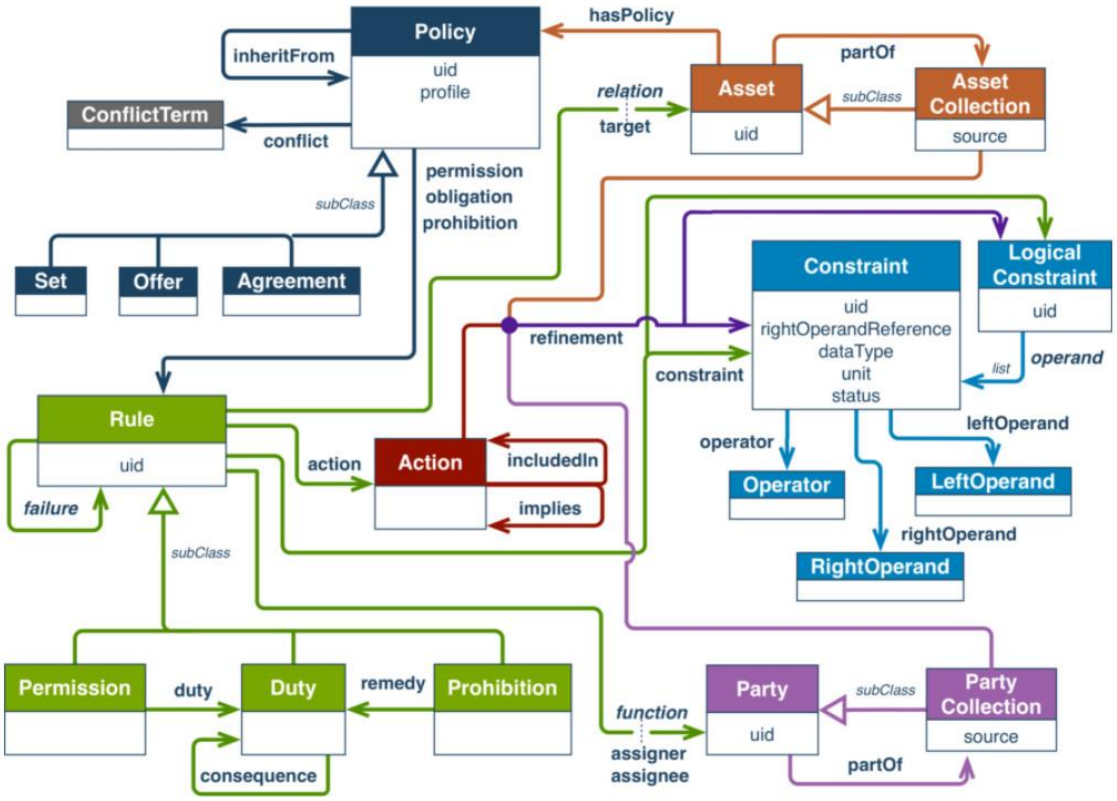

Figure 12 – ODRL Information Model

The next figure represents ODRL Information Model:

Example:

```
{
   "@context": "http://www.w3.org/ns/odrl.jsonld",
   "@type": "Set",
   "uid": "http://example.com/policy:1010",
   "permission": [{
      "target": "http://example.com/asset:9898.movie",
      "action": "use"
   }]
}
```

https://www.w3.org/TR/odrl-model/

5.3.2 IDS Usage Policy Language

It is based on ORDL.

From IDS perspective, a policy restricts the usage of the data to a specific Data Consumer.

Before proceeding with the data exchange, a bilateral communication is established between the data provider and the data consumer in which a series of messages are exchanged with the ultimate goal of reaching an agreement on the use of the data, if the data is to be used for a specific purpose, if they are to be used in a specific time interval, etc. This process is usually called Policy Negotiation.

Once this process is concluded, a ***ContractAgrement*** is obtained in which the policy of use of the data to be exchanged between the data provider and the data consumer is described. The following example is an IDS policy that allows usage of a specific during a concrete time interval.

In an IDS Policy, the main part are the Contracts. Contracts present the container of any usage control statement and come in three different realizations: Requests, Offers, and Agreements

https://industrialdataspace.jiveon.com/docs/DOC-2731

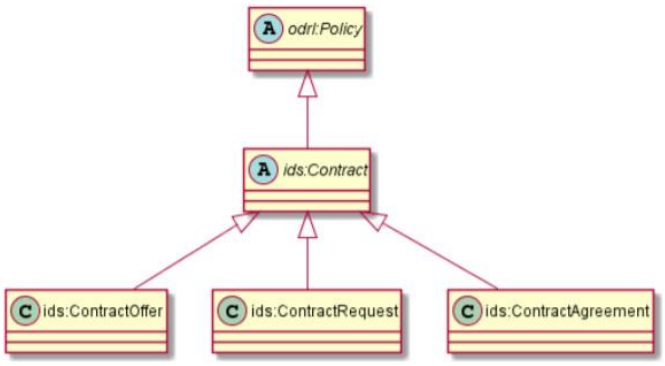

Figure 13 – IDS Contract Types

The IDS Information Model defines offers, requests and agreements as subclasses of the abstract contracts

Contract Type	Implication	Description and Interpretation
Contract Offer	The usage of a certain data resource might be possible regarding the stated constraints.	An offer is purely informative and voluntary metadata by the data provider. They give a rough idea on the usage restrictions and shall improve the discovery and selection process for Data Users. The Data Sovereign benefits by reaching a better visibility of its preferences.
Contract Request	The usage of a certain data resource is desired under the stated constraints.	The Data User indicates its interest, and may create the request relative to a previously exchanged Data Offer. The Data Sovereign gets to know acceptable constraints of the Data User while the later can further detail a contract.
Contact Agreement	Both participants conclude a contract and agree to the stated constraints. A later adjustment is not possible without a mutual consent.	The constraints have been fixed and accepted by both participants. The usage control systems import the agreement and enable or pre-vent access and usage accordingly.

Figure 14 – IDS Contract Types description

Source: https://industrialdataspace.jiveon.com/docs/DOC-2731

The following IDS Policy represents a ContractAgreement message type between a consumer and a provider for the use of a specific target/dataset in a time interval.

Usage During Interval

```
{
 "@context" : {
  "ids" : "https://w3id.org/idsa/core/",
        "idsc" : "https://w3id.org/idsa/code/"
 },
 "@type" : "ids:ContractAgreement",
 "@id" : "https://w3id.org/idsa/autogen/contractAgreement/52272512-dcbd-4b15-8f1f-f409327a4a9a",
 "ids:permission" : [ {
  "@type" : "ids:Permission",
  "@id" : "https://w3id.org/idsa/autogen/permission/59b0a20a-11bd-4276-8341-af40c8960e98",
  "ids:target" : {
   "@id" : "https://w3id.org/idsa/autogen/artifact/8e3a5056-1e46-42e1-a1c3-37aa08b2aedd"
  },
```

```
 "ids:title" : [ {
  "@value" : "Example Usage Policy",
  "@type" : "http://www.w3.org/2001/XMLSchema#string"
 } ],
 "ids:description" : [ {
  "@value" : "provide-access",
  "@type" : "http://www.w3.org/2001/XMLSchema#string"
 } ],
 "ids:action" : [ {
  "@id" : "idsc:USE"
 } ],
 "ids:constraint" : [ {
  "@type" : "ids:Constraint",
  "@id" : "https://w3id.org/idsa/autogen/constraint/0b7c4ca7-1f9e-4e30-8fa1-7551700c1980",
  "ids:rightOperand" : {
   "@value" : "2020-07-11T00:00:00Z",
   "@type" : "xsd:dateTimeStamp"
  },
  "ids:operator" : {
   "@id" : "idsc:AFTER"
  },
  "ids:leftOperand" : {
   "@id" : "idsc:POLICY_EVALUATION_TIME"
  }
 }, {
  "@type" : "ids:Constraint",
  "@id" : "https://w3id.org/idsa/autogen/constraint/9f2e0197-2ad9-442b-806b-5bb4951a2943",
  "ids:rightOperand" : {
   "@value" : "2021-07-11T00:00:00Z",
   "@type" : "xsd:dateTimeStamp"
  },
  "ids:operator" : {
   "@id" : "idsc:BEFORE"
  },
  "ids:leftOperand" : {
   "@id" : "idsc:POLICY_EVALUATION_TIME"
  }
 } ]
} ],
"ids:provider" : {
 "@id" : "https://w3id.org/idsa/autogen/baseConnector/7b934432-a85e-41c5-9f65-669219dde4ea"
},
"ids:consumer" : {
 "@id" : "https://w3id.org/idsa/autogen/baseConnector/7b934432-a85e-41c5-9f65-669219dde4ea"
},
"ids:contractDate" : {
```

```
    "@value" : "2021-02-18T10:15:21.137Z",
    "@type" : "http://www.w3.org/2001/XMLSchema#dateTimeStamp"
   },
   "ids:contractStart" : {
    "@value" : "2021-02-18T10:15:21.137Z",
    "@type" : "http://www.w3.org/2001/XMLSchema#dateTimeStamp"
   },
   "ids:contractEnd" : {
    "@value" : "2022-02-18T10:15:21.137Z",
    "@type" : "http://www.w3.org/2001/XMLSchema#dateTimeStamp"
   }
}
```


5.4 Consent model

From the data sharing point of view, we have a *consumer party* and a *provider party* exchanging data. If we look at IDS, there is an agreement (contract). This agreement authorizes the data exchange between data consumer and data provider. As we introduce in previous section 5.3 Policy Model, a contract agreement makes a reference to a data usage policy.

When a personal data is going to be share, the scenario is more complex. We have data consumer, data provider and now a data subject. It is the individual (or category of individuals) whose personal data is being processed. So now, usage rules are based on user and consent.

5.4.1 Data Privacy Vocabulary[25]

The Data Privacy Vocabulary (DPV) provides the structure of classes and properties to describe and represent information related to the processing of personal data, always relying on the General Data Protection Regulation of the EU (GDPR).

The vocabulary provides terms to describe:

- Personal Data Categories
- Purposes[26]
- Processing Categories
- Technical and Organizational Measures
- Legal Basis such as Consent

[25] https://dpvcg.github.io/dpv/

[26] https://dpvcg.github.io/dpv/#purposes-classes

- Entities such as Recipients, Data Controllers, Data Subjects
- Rights
- Risks

In this chapter we are going to intro how to manage Consents using Data Privacy Vocabulary.

Consent is one of the legal bases or legal justifications for the processing of personal data, whose legal validity is associated with requirements and obligations based on jurisdictional laws. DPV provides concepts to describe consent and its attributes to represent a 'record' of consent from a compliance perspective.

The module describing consent, illustrated in the next figure, provides the necessary terms to describe consent provision, withdrawal, and expiry.

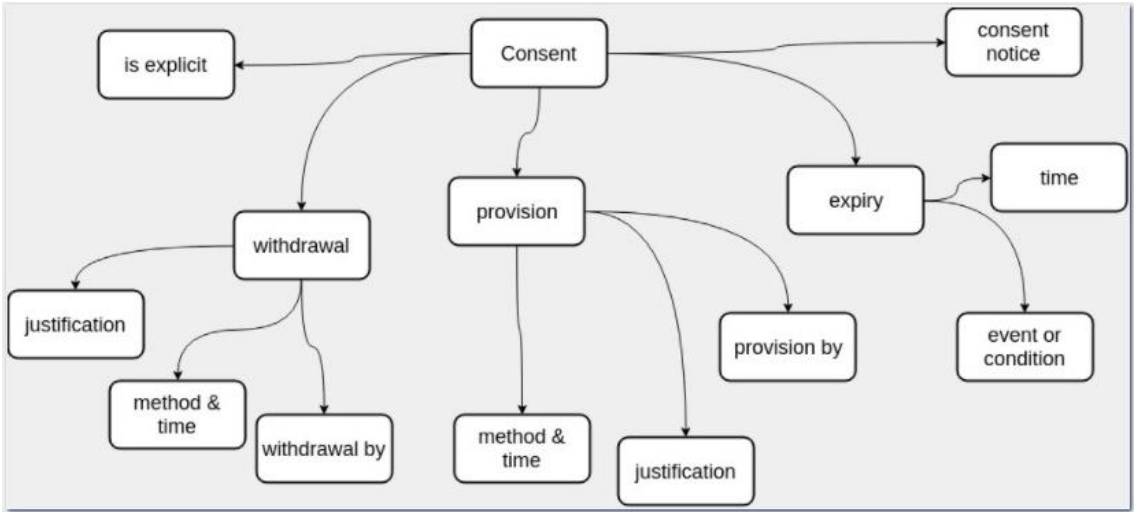

Figure 15 Consent model. Source: https://harshp.com/research/publications/032-creating-vocabulary-data-privacy

5.4.2 MyData Model[27]

MyData intends to build trust in personal data services through a combination of transparency, interchangeability, public governance, respectable companies, public awareness, and secure technology. Consent management is the primary mechanism for permitting and enforcing the legal use of data.

[27] https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/78439/MyData-nordic-model.pdf

Via MyData accounts individuals can instruct the services to fetch and process data in accordance with consents that the individual has granted to data services. In technical and legal terms, consent is equivalent to authorization.

In the MyData model, consents are dynamic, easy for people to comprehend, machine-readable, standardized, and managed in a coordinated way. A common format will make it possible for every individual to delegate data processing to third parties or to repurpose the use of data in new ways

Not all personal data usage requires the consent of data subjects. For example, public authorities are allowed to exchange data between each other without the consent of the data subject in certain circumstances. In such cases, the MyData infrastructure would not be used to enforce consent-based data management, but it would act instead as a transparency tool to notify end-users of the use of their data. It benefits everyone if public authorities are able to exchange personal data in a transparent way.

MyData is a model that equips individuals to control who uses their personal data, to stipulate for what purposes it can be used, and to give informed consent in accordance with personal data protection regulations. It makes data collection and processing more transparent and it helps companies or other organizations implement comprehensive privacy protections.

5.5 User rights model

5.5.1 Data Privacy Vocabulary

The Data Privacy Vocabulary provides terms (classes and properties) to annotate and categorize instances of legally compliant personal data handling. In particular, the vocabulary provides *LegalBasis* and *DataSubjectRight* as top-level concepts representing the various legal bases for justifying processing of personal data and rights provided to the data subject respectively.

Class	Description
Right	• The right(s) applicable, provided, or expected. • A 'right' is a legal, social, or ethical principle of freedom or entitlement which dictate the norms regarding what is allowed or owed. Rights as a concept encompass a broad area of norms and entities and are not specific to Individuals or Data Protection / Privacy.
DataSubjectRight	• The rights applicable or provided to a Data Subject.

	• Based on use of definitions, the notion of 'Data Subject Right' can be equivalent to 'Individual Right' or 'Right of a Person' • Subclass of dpv:Right

5.5.2 DPV-GDPR: GDPR Extension for Data Privacy Vocabulary[28]

This extension extends the DPV and provides concepts specific to the obligations and requirements of the General Data Protection Regulation (GDPR). More specifically, it provides a taxonomy of legal bases and rights as defined within the GDPR.

GDPR provides several rights to the data subject, whose applicability depends on the context and nature of processing taking place. DPV lists these rights at an abstract level as concepts along with their origin in specific clauses of the GDPR. The next table present all classes defined in this extension as subclass of dpv:DataSubjectRight:

Class	Description
A13 Right to be Informed	Information to be provided where personal data is directly collected from data subject
A14 Right to be Informed	information to be provided where personal data is collected from other sources
A15 Right of Access	Right of access
A16 Right to Rectification	Right to rectification
A17 Right to Erasure	Right to erasure ('Right to be forgotten')
A18 Right to Restrict Processing	Right to restriction of processing
A19 Right to Rectification	Right to be notified in case of rectification or erasure of personal data or restriction of processing
A20 Right to Data Portability	Right to data portability
A21 Right to object	Right to object to processing of personal data
A22 Right to object to automated decision making	Right not to be subject to a decision based solely on automated processing including profiling
A7-3 Right to Withdraw Consent	Right to withdraw consent
A77 Right to Complaint	Right to lodge a complaint with a supervisory authority

[28] https://w3c.github.io/dpv/dpv-gdpr/

6 Baseline technologies

This section contains information about three baseline technologies and its applicability in the context of the Personal Data Governance Framework:

- MyData operator
- Attribute Based Credentials (ABC)
- IDS Data usage control

The information about the technologies in this section has been gathered from the original sources.

6.1 MyData

MyData is a global non-profit organisation about digital rights awarded by the European Commission for being one of the most impactful initiatives contributing towards a human centric, trustworthy and sustainable internet. Ethical management of personal data lies at the core of MyData. The underpinning idea behind it is that ordinary people must have easy means to make sense of where data generated by their daily transactions in mobility, shopping, health records and so on go, to determine who can use it and modify these decisions in the course of time.

Interoperability is omnipresent throughout the human centric MyData logic. The aim is to align digital human rights and personal data protection standards with innovation in access to data and business opportunities. Having the people at the driving seat of their own personal data management reconciles ease of data use with data protection. This approach establishes trust among citizens and various types of public and private sector organisations, which in turn creates opportunities for setting up novel services based on the consensual further use of personal data.

In order for any organisational entity (public or private organisation, project, initiative, etc.) to fully understand MyData, it is of fundamental importance to accept the thinking model of user centricity. In this sense, the adoption and implementation of MyData signifies a departure from the traditional model of public service design and delivery that focuses primarily on internal operational capabilities of organisations and a progressive movement towards a model in which the infrastructure for data collection and use is designed with a mindset to serve the needs of the people.

In a nutshell MyData is:

- **a way of thinking** for empowering people to manage personal data, which includes different projects and initiatives internationally,
- **interoperable** in the sense that MyData operators can be hosted by several parties or even be self-hosted

- **a constantly evolving concept** aspiring to offer for everyone a globally accessible networking.

On the other hand, MyData is not:

- **a single project.** There are many projects and initiatives with varying focus, from technical to legal,
- **geographically limited** as national MyData hubs span across five continents, all working with the vision to promote and implement in practice the principles laid out in the Declaration.

6.1.1 ACROSS project and/or pilots as MyData operators

Depending on the extent in which the MyData model for personal data use and management is incorporated in the data governance framework of ACROSS, this would require a **Mydata operator.** In the context of ACROSS either in the level of the project as a whole or at the pilot countries, the MyData Operators will operate the infrastructure and provide tools for the person in a human-centric system of personal data exchange. Operators enable people to securely access, manage, and use personal data about themselves as well as to control the flow of personal data within and between data sources and data using services.

To implement the MyData principles for ACROSS as a project, in some or even all pilot countries, requires at minimum to provide people with access to data about themselves in a way that they can use it also elsewhere and for other purposes. However, it should be noted that providing full control that would allow, for example, editing or deleting that data from the original source is not a prerequisite for complying with MyData. The provision for ACROSS users of transparent access to their personal data in machine readable formats, allowing their re-use, is sufficient for a minimum compliance with Mydata. However, for innovative applications, it would be better to have continuous access to up-to-date data through standardised programming interfaces (APIs). In this way, data updates would not require users to visit the data provider's website, instead services could be used to automate this task. For example, in the case of online payments and their associated data, it would be most useful if citizens receive an electronic e-receipt automatically as soon as they pay for tuition fees.

A Mydata operator can be an organization that manages and utilizes open data, develops APIs and communicates with other organizations that also create and use data and APIs for their own operation. It can also exchange data with them as well as with individuals. To achieve this in a simple and sustainable way a Mydata operator must:

- Create and maintain the Data Standards and enforce that application and data source developers comply with them.
- Provide documentation for the API that grants access to the data.

- Develop login portals for the developers and the end users, using secure authentication and authorization protocols.

Data source developers maintain the data source and register it via the developer portal.

Application developers can then find the registered data sources and explore the given options in the API documentation.

End users can access the online applications, where they receive authorization to access the data.

MyData specification does not cover all the functionality included in the ACROSS Personal data governance framework:

- Data usage policies definition and enforcement
- Attribute based credentials to comply with data minimization principle
- Some of the GDPR user rights are not included in MyData strategy.

6.1.2 MyData operator

MyData Operators Thematic Group which is part of the MyData Global organisation has produced the whitepaper "Understanding MyData Operators" [3] which focuses on practical aspects of technology and governance to make the operation of infrastructures for personal data easier and more human-centric, with the goal of establishing full interoperability between operators.

According to the whitepaper, a MyData operator is an actor that provides infrastructure for human-centric personal data management and governance. In the paper the initial minimum requirements to be considered a MyData operator are presented.

One of the central ideas of the MyData operator model is that there will be a large number of actors providing personal data management services, and that those services should be interoperable and substitutable as well as technology agnostic as far as possible.

In the paper, four dimensions of the MyData Operator concept are considered:

- **Reference model**: The MyData operator reference model provides a structure within which to analyse operators' offerings and characterise their **functional elements**. The reference model creates a baseline for expectations for an operator from individuals, other operators, and other actors in the ecosystem.

- **Interoperability**: Interoperability is key to realising the many benefits of the MyData vision. The paper describes different aspects of interoperability, indicating the role that MyData can play in enhancing human-centric interoperability as ecosystems mature.
- **Governance**: The governance of human-centric data sharing ecosystems is discussed in the contexts of legal and voluntary frameworks. The paper considers how governance should be formulated and enacted, taking into account transparency, the responsibilities of operators towards individuals, and how the nature of who controls an operator impacts this relationship.
- **Business models:** The paper studies parameters of the business models options available to and currently used by some proto-operators, covering fundamental design criteria from the perspectives of human-centricity and financial sustainability.

The purpose of this section is to describe the functional requirements covered by a MyData operator comparing them with the ACROSS Personal Data Governance framework requirements.

The MyData operator reference model describes nine core functional elements of operators. These elements affect how easy it is to utilise personal data, how transparent and human-centric the utilisation of personal data is, and how well the infrastructure supports open competition.

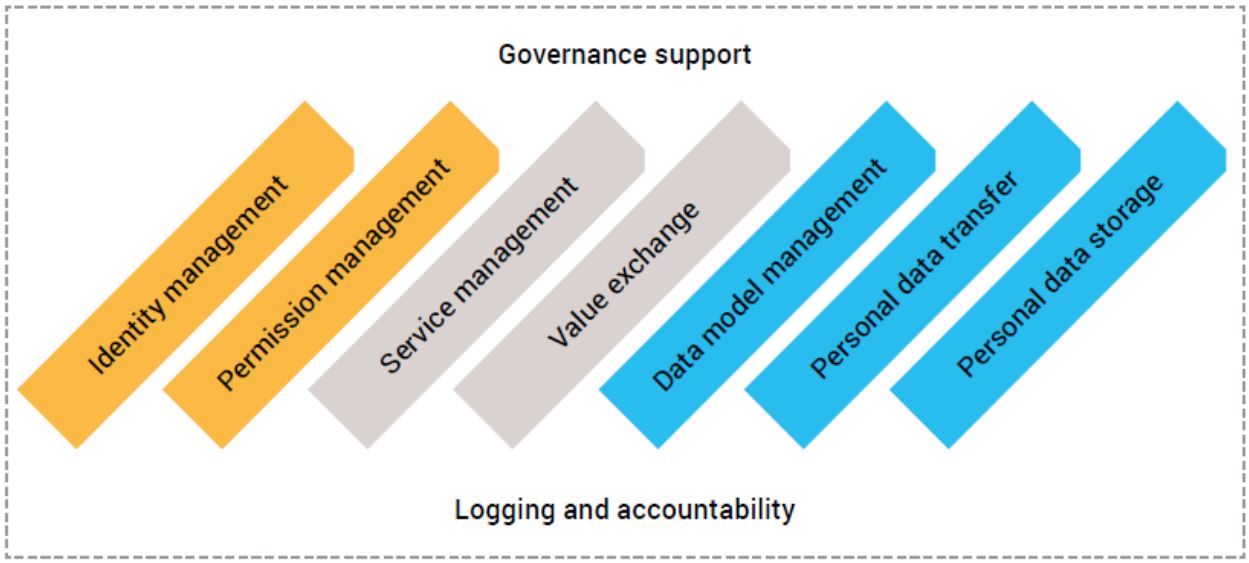

Figure 16 Functional elements of a MyData operator.

The first two (yellow) pillars mediate data transactions in terms of participants and permissions.

The middle two (grey) pillars describe what services are enabled in the ecosystem and how value can be exchanged between ecosystem participants.

The right-hand three (blue) pillars manage data, its meaning, its exchange, and its storage.

'Governance support' and 'Logging and accountability' provide context for the other functional elements and are critical for transparency and trust in the ecosystem.

Next, more information about the pillars are included along with a first mapping with the functional modules of the ACROSS personal data governance framework.

Identity management handles authentication and authorisation of individuals and organisations in different, linked identity domains and links identities to permissions. This module is also included in the ACROSS data governance framework and it is a common module used also by the ACROSS platform.

Permission management enables people to manage and have an overview of data transactions and connections and to execute their legal rights. It includes maintaining records (notices, consents, permissions, mandates, legal bases, purposes, preferences etc.) on data exchange. This is one of main ACROSS data governance framework modules along with the Service management and the Data model management module. The users of this module are the individuals in charge of managing their own personal data records.

Service management uses connection and relationship management tools to link operators, data sources, and data using services. Data can be available from different sources and can be used by multiple data using services.

Value exchange facilitates accounting and capturing value (monetary or other forms of credits or reputation) created in the exchange of data. The value exchange functionality is not within the scope of ACROSS.

Data model management is about managing the semantics (meaning) of data, including conversion from one data model to another. ACROSS data governance framework will use a "personal data model" to classify the personal data and to link the user permissions to the specific personal data categories.

Personal data transfer implements the interfaces (e.g. APIs) to enable data exchange between the ecosystem participants in a standardised and secure manner. This functionality is not covered by the ACROSS data governance framework, but it will be provided by the ACROSS platform to transfer data from the personal data storage to the Public/private services included in the defined user journeys.

Personal data storage allows data to be integrated from multiple sources (including data created by a person) in personal data storage (PDS) under the individuals' control. The personal data storage is an important functionality, but in ACROSS we will rely on existing solutions.

Governance support enables compliance with the underlying governance frameworks to establish trustworthy relationships between individuals and organisations.

Logging and accountability entails keeping track of all information exchanges taking place and creating transparency about who accessed what and when. ACROSS also provides this functionality.

Next a more detailed analysis about the two main functional elements and their implementation in the ACROSS data governance framework.

6.1.2.1 Permission management

According to MyData, permission management covers the technical functionalities required for human-centric *control* of personal data, such as the user interfaces and underlying data structures for individuals to view, understand, grant, revoke, and modify different kinds of permissions related to data flows.

The term 'permission' is used in a broad sense to cover the means that the individual has to take control of data flows. These means can be based on legislation (executing legal rights) or go beyond that. Part of the permission management functionality is that the operator only allows execution of such data transactions where the permission is valid.

ACROSS data governance framework will focus on permission management, providing a way for people to orchestrate the specific data that can be shared (or disclosed) between parties, for which purposes, and for how long. Furthermore, it will go beyond basic permission management including the following functionalities:

- IDS data usage policies for more fine-grained usage control
- Facilitate the GDPR minimization principle using Attribute Based Credentials or defining data transformation filters as a data usage policy.
- Provide the technical means to exercise the GDRP user rights, even after the data has been transferred.

6.1.2.2 Service management

MyData operators live in an ecosystem with data sources and data using services. Navigating this ecosystem requires the linking of actors through an operator: this is the purpose of the service management functionality. The human-centric manifestation of service management is the possibility for individuals to manage the relationships and connections to different data sources and data using services in the ecosystem.

Service management enables dynamic linking of data sources and data using services (permissioned by the individual) so that data can be available at different sources and can be used by multiple data using services.

In a multi-operator environment, it is a significant decision whether the operators use a shared service registry (potentially still distributed) or if each operator manages services separately. This is a topic that will evolve in future work; currently, there is limited standardisation or convergence in this field.

Service management encompasses both access control and technical connection management. However, the delivery of these functionalities is largely determined by the data sources. Operators may support these to a greater or less extent through, for example, key management services.

The ACROSS data governance framework will provide the service management functionality by designing and implementing a Graphical User Interface and providing a set of APIs to interact with the services.

The Graphical User Interface will allow the service providers to define the service metadata, including its characteristics from the point of view of personal data usage. On the other hand, the data governance framework APIs will provide the means to inform the services any change in the users´ consent or to exercise the GDRP user rights like data rectification or data portability.

6.1.2.3 ACROSS Data Governance Framework vs MyData operator: gap analysis

After a first analysis of the MyData operator requirements the conclusion is that ACROSS Data Governance Framework will share its main capabilities regarding permission and service management and it could be considered a proto-operator.

The main differences are:

- ACROSS data governance framework does not include data storage, data transfer and value exchange capabilities.
- ACROSS data governance extends the concept or permission with IDS data usage policies
- ACROSS will facilitates the GDPR minimization principle by leveraging the ABC technology (to be analysed)

6.1.3 ACROSS as a potential IHAN pilot[29]

In the current state of affairs, the citizens' information is fragmented - and often duplicated – between different branches of government, making acquiring and modifying it harder and more time-consuming

[29] https://www.sitra.fi/en/articles/ihan-fi-frequently-asked-questions/

than it should be. For example, the seemingly trivial task of updating your home address might require weeks of communication with public and private organizations, as each one of them must be informed about the change individually. In addition, most of them require additional proof, like an electricity bill, in order to proceed and update the information. The problem is magnified when moving abroad for work or studies, in which case the citizens have to be aware and compliant to each country's own system and procedures.

Our data operator's goal is to standardize the Data, so that:

1. Individuals will be able to access, verify and update their information.
2. Individuals will be able to control which organizations can access their data.
3. Organizations will be able to request access to data from other organizations or individuals.

This aims to help individuals in several actions that relate to two main transnational journeys: i) the case of studying abroad and ii) the case of working abroad. Some indicative actions in each scenario are listed in the following.

Study abroad

1. Commencement of the application process in an educational institution.
2. Certification of previous educational documents acquired (it is often necessary to demonstrate previous academic achievements etc.).
3. Submission of documents requested by an institution of higher education (which are similar in many universities but may differ).
4. Payment of the tuition fees (if any).
5. Access to different grants and other types of financial support (information may be scattered and incomplete).
6. Application for an academic recognition and certification for the under- or post-graduate studies awarded in a different country than that of destination. This is a common prerequisite of enrolment to an academic institution.
7. Dealing with relocation matters such as issuing a residence permit, searching for a residence (e.g. university-owned facilities/ residentials), and other practical issues affecting any person when moving to another country.

Work abroad

1. Commencement of the application process in an organisation, firm or institution (particularly if this is owned or controlled by a public administration).

2. Certification of educational documents acquired and/or application for academic recognition.
3. Submission of authentication documents (including proof of registration of birth), current address (proof or residence), work experience or social insurance (e.g. working stamps, health coverage, etc.), including the notification of changes in the personal or professional circumstances of the person receiving social security benefits,
4. Registration to local tax-office (where applicable) and/or submission of income tax declaration
5. Dealing with matters related to the relocation - residence permit, search for residence, and other practical issues affecting any person when moving to another country.

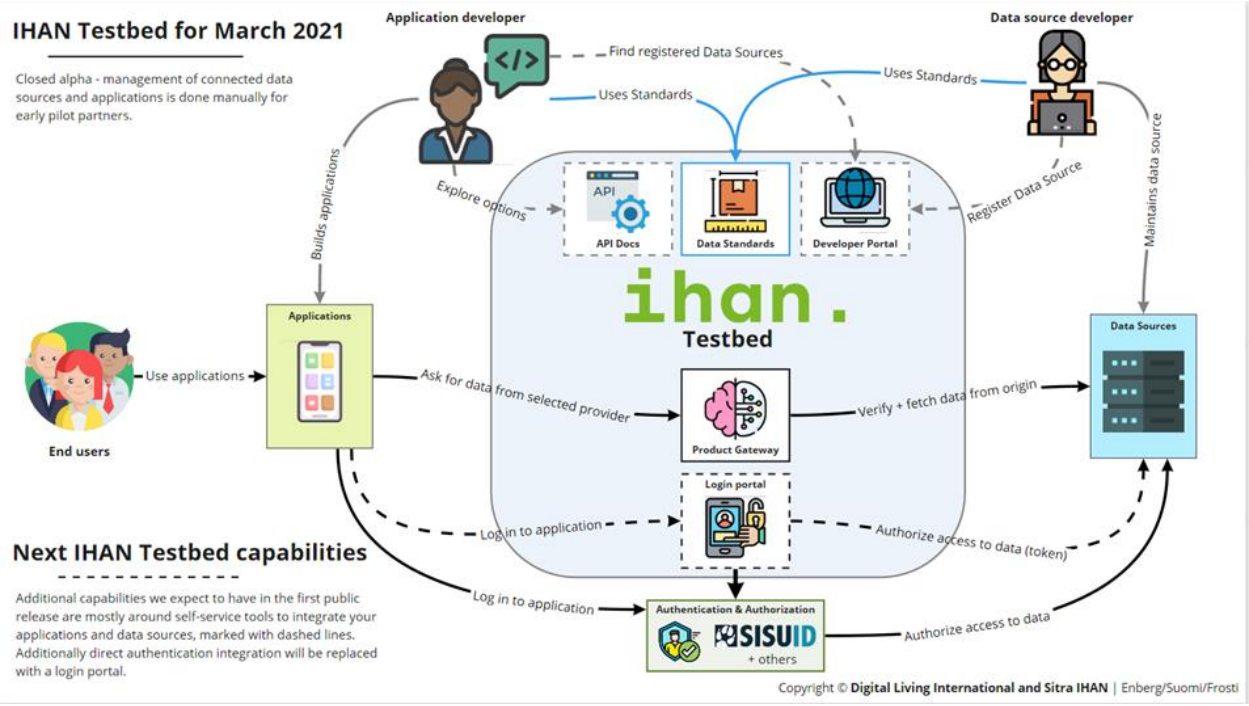

Figure 17 Capabilities of the IHAN Testbed

6.2 Attribute Based Credentials (ABC)

Attribute based credentials (ABCs) are a technology that could potentially be applied to enable privacy, granular control, and data minimisation in ACROSS. Put simply, ABCs are 'a way to have a trusted party 'vouch' for you in a situation where you don't want to give away any more information than is absolutely necessary' (Waag). PrivacyPatterns.org describes them as 'a form of authentication mechanism that allows to flexibly and selectively authenticate different attributes about an entity without revealing additional information about the entity (zero-knowledge property).'

The DECODE pilot in Amsterdam (in which Waag was a partner) developed a proof-of-concept to enable ABC in certain city services. One of the pilot's applications allowed people to generate a credential to

prove they are over 18 without disclosing all of the other information shown in the passport, including that person's specific age and date of birth. Further information on the ABC pilots in DECODE can be found at https://decodeproject.eu/publications/deployment-pilots-amsterdam.html .

There are a number of potential ways in which ABCs could potentially be applied to cross border services. At present, people moving across borders have to share a lot of information with a lot of different parties, ranging from public to private. One can imagine many examples where a person would want to exercise more granular control to minimise the amount of data they share – for example, a person may want to prove that they are an EU resident without disclosing their home address; or may want to prove they are eligible for an apartment without disclosing their income and employer; etc.

6.2.1 Applicability of ABC in ACROSS

Developments over the coming months will inform ACROSS partners as to whether and how ABCs may be a viable implementation in our own project. These developments include the completion of a preliminary gap analysis; the initial co-creation of the ACROSS governance framework; the development of specified use case scenarios; and a technical evaluation which considers this report, among others.

6.3 IDS and GAIA-X

6.3.1 IDS

The International Data Spaces objective is to create data spaces where businesses can exchange and exploit data in a secure manner. For the IDS as well as other data driven businesses, data sovereignty is a key success factor. Data sovereignty has the goal to provide a Data Owner with full control over his data. This includes being able to control the usage of his data by the Data Consumer.

Data usage control and data provenance are conceptual and technical solutions to cope with data sovereignty.

Nowadays, business is spurred by continuously exchanging information between business partners. However, data is typically secured by access control mechanisms only. After access to data has been granted by these mechanisms, data can be arbitrarily altered, copied and disseminated by the recipient. **Data usage control** offers possibilities to control future data usages beyond the initial access (also known as obligations).

In the age of Industry 4.0, there is more critical and sensitive data exchanged between business partners. In general, companies have intrinsic and extrinsic motivations to apply usage control: On the one hand, companies may use usage control to prevent misuse of their own data, to protect their intellectual

property, and to preserve the data value (intrinsic motivation). On the other hand, companies have to comply with legal obligations such as the European Union General Data Protection Regulation EU-GDPR (extrinsic motivation). Hence, companies have to prevent misuse of other persons or companies' data.

Usage control is an extension to traditional access control (see Figure 18). It is about the specification and enforcement of restrictions regulating what must (not) happen to data. Thus, usage control is concerned with requirements that pertain to data processing (obligations), rather than data access (provisions). Usage control is relevant in the context of intellectual property protection, compliance with regulations, and digital rights management.

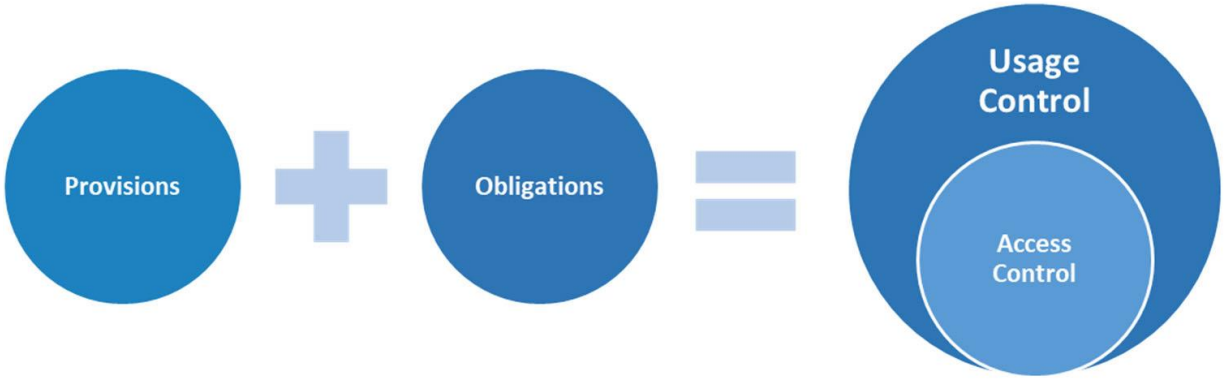

Figure 18 Usage Control consists of provisions and obligations

6.3.1.1 Usage Control

In contrast to access control, where access to specific resources (e.g., a service or a file) is restricted, the IDS architecture additionally supports data-centric usage control. In general, the overall goal is to enforce usage restrictions for data after access has been granted. Therefore, the purpose of usage control is to bind policies to data being exchanged and to continuously control the way how messages may be processed, aggregated, or forwarded to other endpoints. This data-centric perspective allows the user to continuously control data flows, rather than accesses to services. At configuration time, these policies support developers and administrators in setting up correct data flows.

At runtime, the usage control enforcement prevents IDS connectors from treating data in an undesired way, for example by forwarding personal data to public endpoints. Thus, usage control is both a tool for system integrators to ensure they are not building an architecture that violates security requirements, and an audit mechanism, which creates evidence of a compliant data usage.

It is important to note that the purpose of usage control is to allow the specification of such constraints and enforcing them in the running system.

6.3.1.2 Enforcement

For enforcing usage restrictions, data flows need to be monitored and potentially intercepted by control points (i.e., PEPs - Policy Enforcement Point). These intercepted data flows are given to the decision engine (i.e., the PDP – Policy Decision Point) for requesting permission or denial of the data flow. In addition to just allowing or denying the data flow, the decision can also require a modification of data. A PEP component encapsulates the enforcement.

For example, a data owner demands the deletion of data after a certain time or that only a limited audience can access the sensitive data. Hence, we have to intercept the data flow and check which audience (i.e., processing system) is using the data. For example, the data owner demands the data consumer that only the supplier management system can use the data.

6.3.1.3 Implementation of Usage Control in the IDS

There are two main activity streams within the IDS to implement data usage control:

- First, a **policy language** to express data usage restrictions is developed. The policy language is descriptive, technology-independent and based on the Open Digital Rights Language (ODRL) and further detailed in the form of IDS contracts. To express usage restrictions within the IDS, there are several predefined classes that express the most commonly data usage restrictions.
- Second, **usage control technologies** are developed to enforce these usage restrictions at technical level.

An IDS Contract is implicitly divided to two main sections: the contract specific metadata and the IDS Usage Control Policy of the contract. The contract specific information (e.g., date when the contract has been issued or references to the sensitive information about the involved parties) has no effect on the enforcement. However, the IDS Usage Control Policy is the key motive of organizational and technical Usage Control enforcement.

Furthermore, an IDS Usage Control Policy contains several Data Usage Control statements (e.g., permissions, prohibitions and obligations) called IDS Rules and is specified in the IDS Usage Control Language which is a technology independent language. The technically enforceable rules shall be transformed to a technology dependent policy (e.g., MYDATA[30]) to facilitate the Usage Control enforcement of data sovereignty.

[30] https://www.dataspaces.fraunhofer.de/en/software/usage-control/mydata.html

The inner part of the usage control is the IDS connector. Depending on the usage restrictions, they are applied at the data provider connector or at the data consumer connector.

At the data provider connector, usage control enforces policies such as how often data can be accessed, at what times (e.g., only within business hours), or that data must be filtered or masked (e.g., anonymized) before leaving the company. The usage restrictions at data provider connector are usually provisions that are technically handled by a PEP.

At the data consumer connector, usage control enforces policies that are usually obligations for the data consumer such as "data can only be used for fourteen days" or "data can only be used for the purpose of predictive maintenance". The technical enforcement is handled by a PEP or PXP (Policy Execution Point), depending on the usage restriction. Limiting data flowing to a specific target system to ensure the correct usage purpose is handled by a PEP, the deletion of data in storage infrastructure outside the connector is handled by a PXP that performs the delete operation.

6.3.2 GAIA-X

The GAIA-X Ecosystem is formed by an Infrastructure Ecosystem plus a Data Ecosystem, both connected via Federation services while the whole architecture bases upon Policy Rules and an Architecture of Standards. Both follow the idea of a shared economy where you can share your data and services while applying policies and maintaining sovereignty over them. The two ecosystems cannot be viewed separately. Within the Infrastructure Ecosystem infrastructure services are provided, connected or consumed, while the Data Ecosystems deal with data as the main business asset. Similar to the IDS, ecosystem participants are classified into the general roles Provider and Consumer. According to the activity, an entity can have both roles at the same time. To realize this architecture, GAIA-X aims at leveraging existing standards as well as open technologies and concepts. By combining existing solutions GAIA-X acts as orchestrator and integrator.

The combined architecture of GAIA-X and IDS supports and enables data spaces and builds advanced smart services in industry verticals. GAIA-X focuses on sovereign cloud services and cloud infrastructure, while IDS focuses on data and data sovereignty. The interaction of GAIAX and IDS has three main tasks: self-sovereign data storage, trustworthy data usage and interoperable data exchange. This way, GAIA-X is developed in accordance with the European Data Strategy and supports smart data applications and innovations across industry sectors. For this purpose, GAIA-X and IDS complement each other to ensure cloud and data sovereignty for end-to-end data value chains in federated ecosystems.

6.3.3 Applicability of IDS Data Usage Control in ACROSS

IDS Data usage control aims at exchanging of information between business partners and it does not explicitly cover personal/private data issues. However, some of the usage policies defined by IDS are also applicable to personal data transfers, giving the individual the possibility to control the way in which personal data is used in more fine-grained manner.

An IDS policy could be used to constraint the location in which data is used, to enforce the deletion of the personal data after a specific period or to modify the personal during the data transfer process.

In fact, some of the defined data usage policies can be used to enforce the some GDPR rights and principles. For example, the rule that allows to modify the data in transit could be used to ensure the data minimization principle.

However, in IDS, the policy enforcement functionality is performed by the IDS connectors and some of the rules can only be applied by the consumer connector, so it can be used only if the services (both public and private services) deploy the IDS connectors technology for data transfer.

ACROSS personal data framework will provide IDS policy enforcement capabilities without the use of IDS connectors as an added value functionality, as part of the Usage control module (See 4.2.2).

7 Conclusions and next steps

This section presents some conclusions gathered from the requirements analysis performed in this deliverable. These conclusions will drive the evolution of the data governance framework design and its implementation.

The concept of Personal Data Governance framework defined in ACROSS is perfectly aligned with several initiatives in the field of Personal and Private data management, including:

- **MyData initiative**: MyData operator
- **Data Governance Act**: Personal data sharing intermediary
- **Tech Dispatch published by the European Data Protection Supervisor**: Personal Information Management System (PIMS).

The ACROSS Personal Data Governance Framework is part of the so-called ACROSS platform, useful in the context of cross-border public/private services for the citizens. However, it has been designed to be an independent framework useful also for the individuals to manage their personal data according to the GDPR in any other context.

ACROSS data governance framework does not cover the following functionalities included in most of the initiatives:

- Secure Data Storage
- Secure Data transfer among services

ACROSS will analyse how to extend the MyData operator with the following functionalities:

- Data minimization via the ABC technology
- Data usage policies

The decisions presented in this deliverable will be a subject to refinements and modifications, based on the progress of the other work packages, as well as the validation and evaluation phases. The following tasks will be performed in order to design and deploy the first version of the Personal Data Governance Framework:

- Analyse CAPE[31] open source implementation of the MyData Operator concept
- Define the Data usage policies applicable to ACROSS use cases and provide the data usage enforcement functionality

[31] https://github.com/OPSILab/Cape

- Analyse DECODE ABC technology applicability
- Analyse the applicability of the existent models and modify/extend them to define the ACROSS data model.

8 References

[1] Data governance act: https://digital-strategy.ec.europa.eu/en/policies/data-governance-act

[2] Regulation (EU) 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. https://ec.europa.eu/info/law/law-topic/data-protection/data-protection-eu_en

[3] Understanding MyData Operators https://mydata.org/wp-content/uploads/sites/5/2020/04/Understanding-Mydata-Operators-pages.pdf

[4] TechDispatch #3/2020 – Personal Information Management Systems, 6 January 2021, from the European Data Protection Supervisor; see https://edps.europa.eu/data-protection/our-work/publications/techdispatch/techdispatch-32020-personal-information_en

[5] Opinion 9/2016 – EDPS Opinion on Personal Information Management Systems – Towards more user empowerment in managing and processing personal data; see https://edps.europa.eu/sites/default/files/publication/16-10-20_pims_opinion_en.pdf

[6] Regulation (EU) No 910/2014 of the European Parliament and of the Council of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC, OJ L 257, http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014R0910&from=EN

9 Annex I – WP5 Requirements

		WP5 Requirement Collection Sheet			
Id	**Title**	**Description**	**Type**	**Category**	**Priority**
Req_01	Semantic and technical interoperability with SDG	The system should ensure an alignment of semantic and technical interoperability with SDG IT Tools	non functional	Platform architecture and interoperability	MUST HAVE
Req_07	REQ-DevOps Processes Setup	A full end-to-end pipeline of processes should be set up to ensure the successful integration, deployment, testing and delivery of the services. The DevOps processes, development and operations should be integrated into a single-minded entity with common goals: high-quality software, faster releases and improved users' satisfaction.	non functional	Platform architecture and interoperability	SHOULD HAVE
Req_08	REQ-Continuous Platform Integration	The DevOps approach should be followed in ACROSS Platform to include the continuous integration tools.	non functional	Platform architecture and interoperability	SHOULD HAVE
Req_12	Scalability	The ACROSS platform should be designed to be scalable in terms of computational load, number of users accessing applications and amount of data storage. In particular the platform should be able to scale horizontally (e.g. add more nodes to a computational network) and vertically (e.g. add resources such as Memory, CPU to a single node in a system).	non functional	Platform architecture and interoperability	SHOULD HAVE
Req_13	Interoperability with legacy systems	It has to be possible to connect the ACROSS platform with the existent PA legacy systems (e.g. databases, web services). Secure and reliable communication with the existing public administration information systems have to be provided without requiring changes in these systems. The platform should also provide tools and predefined components to facilitate the interoperability.	non functional	Platform architecture and interoperability	MUST HAVE

Req_19	Reliability and Integrity	The implementation of ACROSS should follow open standards and use well-known and widely accepted technologies in order to ensure integrity. The ACROSS platform has to be reliable assuring integrity of the components/tools that are part of it.	non functional	Platform architecture and interoperability	MUST HAVE
Req_22	Privacy and Data Protection	The ACROSS platform has to be compliant with the EU legislation regarding privacy and data protection. It should adopt all the necessary technologies, standards and methods to protect privacy of the users of the platform services and to secure stored information that could be considered private.	non functional	Security and Privacy	MUST HAVE
Req_25	OpenID Connect - Client-Registration	It is needed to register the clients that want to use the ACROSS-Platform	non functional	Security and Privacy	SHOULD HAVE
Req_29	No vendor lock-in	I want the ACROSS reference architecture to be technologically agnostic to avoid vendor lock-in.	non functional	Platform architecture and interoperability	SHOULD HAVE
Req_30	Open source	I want the ACROSS reference architecture to reuse already available open source solutions and only create or improve those aspects that are not covered by the existing solutions	non functional	Platform architecture and interoperability	MUST HAVE
Req_33	Accessibility	The front-ends of the system should comply with the current Web Accessibility Directives and in particular with EN301549 (included in WCAG-2.1)	non functional	Web&Mobile applications	SHOULD HAVE
Req_34	Confidentiality	The platform has to follow the 'privacy-by-design' and 'security-by-design' approaches and in particular should comply with the principle that users shouldprovide only the information that is absolutely necessary.	non functional	Security and Privacy	MUST HAVE

Req_35	Usability and adaptability	The provided solutions in the platform should be user-friendly and easy to use and should be multilingual. No piece of text that might be displayed to a user shall reside in source code and solution and user should be able to select the preferred language . The implementation of the system should follow open standards and use well-known and widely accepted technologies in order to ensure ease of use.	non functional	Platform architecture and interoperability	MUST HAVE
Req_36	Minimal browser support.	The component user interface (where available e.g. dashboards, forms, ect..) should provide support for the wide range of widely used browsers.	non functional	Web&Mobile applications	MUST HAVE

[i] https://www.w3.org/TR/odrl-model/